Active Calculus - Multivariable

2nd Ed

Active Calculus - Multivariable

2nd Ed

Nicholas Long
Stephen F. Austin State University

Mitchel T. Keller
University of Wisconsin-Madison

Contributing Authors

Steve Schlicker
Grand Valley State University

David Austin
Grand Valley State University

Matt Boelkins
Grand Valley State University

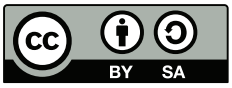
August 11, 2026

Edition: 2nd Edition

Website: https://activecalculus.org/

©2013–2026 Nicholas Long, Mitchel T. Keller, and Steven Schlicker

Permission is granted to copy and (re)distribute this material in any format and/or adapt it (even commercially) under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License. The work may be used for free in any way by any party so long as attribution is given to the author(s) and if the material is modified, the resulting contributions are distributed under the same license as this original. All trademarks™ are the registered® marks of their respective owners. The graphic



that may appear in other locations in the text shows that the work is licensed with the Creative Commons and that the work may be used for free by any party so long as attribution is given to the author(s) and if the material is modified, the resulting contributions are distributed under the same license as this original. Learn more at https://creativecommons.org/licenses/by-sa/4.0/.

Features of the Text

Similar to the presentation of the single-variable *Active Calculus*, instructors and students alike will find several consistent features in the presentation, including:

Motivating Questions At the start of each section, we list *motivating questions* that provide motivation for why the following material is of interest to us. One goal of each section is to answer each of the motivating questions.

Preview Activities Each section of the text begins with a short introduction, followed by a *preview activity*. This brief reading and the preview activity are designed to foreshadow the upcoming ideas in the remainder of the section; both the reading and preview activity are intended to be accessible to students *in advance* of class, and intended to be completed by students before a class meeting on which a particular section is to be considered.

Activities Every section in the text contains several *activities*. These are designed to engage students in an inquiry-based style that encourages them to construct solutions to key examples on their own, working either individually or in small groups.

Exercises There are dozens of calculus texts with (collectively) tens of thousands of exercises. Rather than repeat a large list of standard and routine exercises in this text, we recommend the use of WeBWorK with its access to the Open Problem Library and its many multivariable calculus problems. In this text, each section begins with several anonymous WeBWorK exercises, and follows with several challenging problems. The WeBWorK exercises are best completed in the .html version of the text. Almost every non-WeBWorK problem has multiple parts, requires the student to connect several key ideas, and expects that the student will do at least a modest amount of writing to answer the questions and explain their findings. For instructors interested in a more conventional source of exercises, consider the freely available APEX Calculus text by Greg Hartmann et al.

Our release of the 2nd Edition of ACM in Fall 2026 includes WeBWorK exercises in each of the sections, but we welcome suggestions of additional WeBWorK exercises that faculty would consider good inclusions and feedback on any aspect of this edition. The best way to provide feedback is through the Active Calculus - Multivariable Google Group or through this Google Form. If you are comfortable with GitHub and/or PreTeXt, you can submit your suggested changes as a Pull Request or Issue to the ACM repository. We will occasionally give some updates through the blog on activecalculus.org.

Graphics As much as possible, we strive to demonstrate fundamental ideas visually, and to encourage students to do the same. Throughout the text, we use full-color graphics to exemplify and magnify key ideas, and to use this graphical perspective alongside both numerical and algebraic representations of calculus. Because so much of vector calculus relies upon visualizing things in three dimensions, we elected to produce almost all of the 3D graphics using SageMath. This allows for a degree of consistency between the two-dimensional graphics, which are typically static, and the three-dimensional graphics, which are almost always interactive. Thus, you can grab a three-dimensional plot, rotate it, zoom in on it, etc.

- To *rotate* a three-dimensional graphic, (left) click and drag with your mouse.
- To *zoom* on an interactive three-dimensional graphic, use your mouse's scroll wheel or make your operating system's scroll gesture on your touchpad.
- To *move* a three-dimensional graphic instead of rotating, hold down the space bar while clicking (or use right click) and dragging.

For a more comprehensive description with examples and suggestions on other tools, please see our Preface on Graphics and Technology

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

Acknowledgments

2nd Edition: Nick Long. I want to first thank my family, Jane, Nixon, and Vivian, for their support of me during this and my too many other projects. I am also immensely thankful of my coauthor Mitch and our excellent working relationship which has served as a model for many others. I want to acknowledge my other colleagues, especially Jane Long, Jeremy Becnel, and Jacob Turner, who have encouraged me through this project and helped refine my crazy ideas. The Second Edition of ACM is much better because the authoring tools of PreTeXt have really delivered in helping authors write excellent materials without having to build everything themselves. In particular, I would like to thank Rob Beezer for his leadership of PreTeXt and Oscar Levin for the excellent work he has done on PreTeXt and the pretext-tools extension. Finally, I would like to thank SFA for giving me Faculty Development Leave to work on ACM half-time for the 24-25 academic year.

2nd Edition: Mitch Keller. Few journeys are straightforward, and my path to the release of the 2nd edition of *Active Calculus - Multivariable* has been particularly winding. (Perhaps appropriate for a text that features representing curves with vector-valued functions!) I started drafting the first sections of what is now Chapter 13 while a faculty member at Washington & Lee University because I needed to teach some vector calculus in multivariable calculus. I met Matt Boelkins at a Project NExT reception at the Joint Mathematics Meetings, and he encouraged me to apply to the American Institute of Mathematics workshop on open-source textbooks to work on *Applied Combinatorics*. Next came an invitation to work on finishing my work on vector calculus to supplement Steve Schlicker's work on the 1st edition. Tom Judson mentioned that he had this colleague Nick Long who was doing some interesting things with virtual reality for multivariable calculus and might make a good collaborator. That connection led us to releasing the vector calculus chapter in stages through a wonderfully productive collaboration. As Steve entered retirement and decided that further work on *ACM* was not high on his priority list, I was happy to have the opportunity to continue working with Nick on this 2nd edition, which has turned into a much larger revision than we originally imagined. I am thankful for how each of those folks helped move my role in this project forward, and especially to Nick for being a pleasure to work with.

I join Nick's heartfelt thanks to the PreTeXt community for providing the tools to make this text possible. Rob Beezer's foundational and ongoing work makes me wonder how little sleep a human needs to survive. Oscar Levin's enthusiasm for expanding features and providing tools that make it feasible for a wide range of authors to write in PreTeXt has been instrumental particularly in making collaborative efforts such as *ACM* possible. Although the 2nd edition

of *ACM* benefits only slightly from David Austin's PreFigure (static) graphics library, I am excited about how we will be able to bring a greater degree of consistency and polish to our static graphics before releasing the print version of the 2nd edition. I am also incredibly grateful to David Austin and Kathy Yoshiwara for regular meetings as an OER author group. Their generosity in reading (and sometimes re-reading) sections of *ACM* as we worked through them has made me a better writer and editor, and this book has improved considerably because of their feedback.

From the 1st Edition. This text is an extension of the single variable *Active Calculus* by Matt Boelkins. The initial drafts of this multivariable edition were written by Steve Schlicker; editing and revisions were made by David Austin and Matt Boelkins. David Austin is responsible for the beautiful full-color graphics in Chapters 9 through 11. Many people at GVSU have shared their ideas and resources, which undoubtedly had a significant influence on the product. We thank them for all of their support. Most importantly, we would like to thank the students who have used this text and offered helpful advice and suggestions.

In advance, we also thank our colleagues throughout the mathematical community who have or will read, edit, and use this book, and hence contribute to its improvement through ongoing discussion. The following people have used early drafts of this text and have generously offered suggestions that have improved the text.

Feryâl Alayont	Grand Valley State University
David Austin	Grand Valley State University
Jon Barker and students	St. Ignatius High School, Cleveland, OH
Matt Boelkins	Grand Valley State University
Brian Drake	Grand Valley State University
Brian Gleason	Nevada State College
Mitch Keller	University of Wisconsin-Madison

Early HTML versions of this text were made possible only because of the amazing work of Rob Beezer and his development of PreTeXt. Dr. Schlicker's ability to take advantage of Rob's work is largely due to the support of the American Institute of Mathematics, which funded him for a weeklong workshop in Mathbook XML in San Jose, CA, in April 2016. David Farmer also deserves credit for the original conversion of the text from LaTeX to PreTeXt. The PreTeXt community continues to provide enormous support and capabilities to all of the authors as we continually strive to provide an excellent text while updating the utility of this resource.

We are interested in your suggestions for improvments to this text, including any errors or inconsistencies you may find. Please join the Active Calculus - Multivariable Google Group to see updates and provide feedback to the authors.

David Austin, Matt Boelkins, Steven Schlicker, Mitch Keller, Nick Long -2023.

Active Calculus - Multivariable: our goals

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

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

In *Active Calculus - Multivariable*, we endeavor to actively engage students in learning the subject through an activity-driven approach in which the vast majority of the examples are completed by students. Where many texts present a general theory of calculus followed by substantial collections of worked examples, we instead pose problems or situations, consider possibilities, and then ask students to investigate and explore. Following key activities or examples, the presentation normally includes some overall perspective and a brief synopsis of general trends or properties, followed by formal statements of rules or theorems. While we often offer plausibility arguments for such results, rarely do we include formal proofs; We will often present a geometric or measurement based motivation for topics and provide an introduction to both nomenclature

and the critical definitions. It is not the intent of this text for the instructor or author to *demonstrate* to students that the ideas of calculus are coherent and true, but rather for students to *encounter* these ideas in a supportive, leading manner that enables them to begin to understand for themselves why calculus is both coherent and true.

This approach is consistent with the following goals:

- To have students engage in an active, inquiry-driven approach, where learners strive to construct solutions and approaches to ideas on their own, with appropriate support through questions posed, hints, and guidance from the instructor with proper motivation and context.
- To build students intuition for why the main ideas in multivariable calculus arise naturally and to motivate the related measurements that students will encounter in future math and science courses. A particular focus of this is to present the geometric version of each key idea and typically use either vector tools/intuition or a classic calculus approach. We strive to accomplish this by using specific cases to highlight the ideas for the general situation using contexts that are common and familiar.
- To challenge students to acquire deep, personal understanding of multivariable calculus through reading the text and completing preview activities on their own, through working on activities in small groups in class, and through doing substantial exercises outside of class time.
- To strengthen students' written and oral communication skills by having them write about, explain aloud, and consider other's descriptions of the key ideas of multivariable calculus.

How to Use this Text

Because the text is free, any professor or student may use the electronic version of the text for no charge. For reading on laptops or mobile devices, the best electronic version to use is the HTML version of the text, but you can find links to a pdf and hard copy of the text at https://activecalculus.org/. Furthermore, because the text is open-source, any instructor may acquire the full set of source files, which are available on GitHub.

As of August 2026, you can find the official release of the second edition of ACM at activecaclulus.org. We will strive to not make changes to this edition during the academic years. You can find a preview version of an updated version (typos and other fixes) of the Second Edition of ACM at https://nelong.github.io/active-calculus-vector/multi/frontmatter.html and the instructor version at https://nelong.github.io/active-calculus-vector/multi-instructor/frontmatter.html. For a complete description of the differences between the first and second editions see https://activecalculus.org/acm/acm-1e-vs-2e/

This text may be used as a stand-alone textbook for a standard multivariable calculus course or as a supplement to a more traditional text. Chapter 9 covers the precalculus ideas necessary for multivariable calculus: coordinates, vectors and common graphs (planes, lines, quadric surfaces, and cylinder surfaces) in three dimensions. Chapter 10 covers vector-valued functions of one variable and introduces how to use vector tools and calculus on these functions. Chapter 11 introduces functions of more than one variable and studies differentiation of functions of several independent variables in detail, addressing the typical topics including limits, partial derivatives, and optimization. Chapter 12 motivates and defines double and triple integrals.

Chapter 13 covers vector fields and vector calculus topics through Stokes Theorem and the Divergence Theorem. The 2nd edition of ACM has evolved because in reality, "multivariable calculus" can mean one of three different courses: one that follows chapters 9-12 of the 2nd edition and omits vector calculus entirely; one that introduces students to line integrals, the Fundamental Theorem of Calculus for Line Integrals, and Green's Theorem; and one that proceeds through all of that material plus Stokes's Theorem and Gauss's Divergence Theorem. As such, the extent to which certain topics in chapters

9-12 are necessary varies depending on the course's final destination.

Electronic Edition

Because students and instructors alike have access to the book in electronic format, there are several advantages to the text over a traditional print text. One is that the text may be projected on a screen in the classroom (or even better, on a whiteboard) and the instructor may reference ideas in the text directly, add comments or notation or features to graphs, and indeed write right on the projected text itself. Students can do the same when working at the board. In addition, students can choose to print only whatever portions of the text are needed for them. Also, the electronic versions of the text includes links to on-line programs, so student and instructor alike may follow those links to additional resources that lie outside the text itself. Finally, students can have access to a copy of the text anywhere they have a computer.

Beginning with the 2nd Edition, we have written activities and the presentation of material with the html as the primary focus. For example, almost every three-dimensional plot is interactive in the HTML version of the text, allowing readers to rotate, zoom, or use interactive tools like sliders and buttons. This offers students the opportunity to literally look at the objects from a different perspective.

Note: In the PDF version, there is sometimes not an obvious visual indicator of the HTML links, so some available information is suppressed and there are often links or QR codes to view an interactive plot. If you are using the text electronically in a setting with internet access, please know that it is assumed you are using the HTML version.

Activities Workbook

Each section of the text has a preview activity and at least three in-class activities embedded in the discussion. As it is the expectation that students will complete most of these activities, it is ideal for them to have room to work on them adjacent to the problem statements themselves. A separate workbook of activities that includes only the individual activity prompts, along with space provided for students to write their responses, is in development and should be widely available in the 26-27 academic year.

Community of Users

Because this text is free and open-source, we hope that as people use the text, they will contribute corrections, suggestions, and new ideas for material. At this time, the best way to communicate such feedback is through the Active Calculus - Multivariable Google Group. We have also set up a Google Form to collect feedback. If you are comfortable with GitHub and/or PreTeXt, you can submit your suggested changes as a Pull Request or Issue to the ACM repository. We will occasionally give some updates through the blog on activecalculus.org.

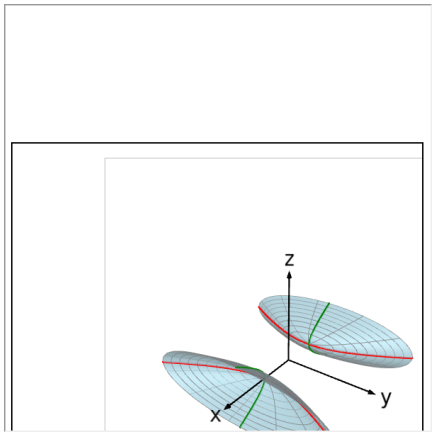
Graphics and Technology

As much as possible, we strive to demonstrate key fundamental ideas visually, and to encourage students to do the same. Throughout the text, we use full-color graphics to exemplify and magnify key ideas, and to use this graphical perspective alongside both numerical and algebraic representations of calculus. Because so much of vector calculus relies upon visualizing things in three dimensions, we elected to produce almost all of the 3D graphics using SageMath. This allows for a degree of consistency between the two-dimensional graphics, which are typically static, and the three-dimensional graphics, which are almost always interactive. Thus, you can grab a three-dimensional plot, rotate it, zoom in on it, etc.

- To *rotate* a three-dimensional graphic, left click (primary click) and drag with your mouse. On touchscreen devices, you can press (with a single finger or stylus) and drag to rotate a figure.
- To *zoom* on an interactive three-dimensional graphic, use can your mouse's scroll wheel to zoom in or out based on your current viewing position. On touchscreen or touchpad devices you can use the pinch and expand gestures to zoom in and out. On touchpad devices, you can also touch with two fingers and slide up or down to change the zoom.
- To *move* the location from which you are viewing a three-dimensional graphic (instead of rotating), you can right click (secondary click) and drag in the direction you would like to move. Additionally, you can hold Control and click and drag to change the viewing location. On touchscreen devices you can drag with two fingers to move the location from which you are viewing the plot.

Sometimes the interactable graphics can go to sleep and become unresponsive. If this happens you can reload the page and all interacts will refresh. Additionally, you may find it useful to click on the pop-out button

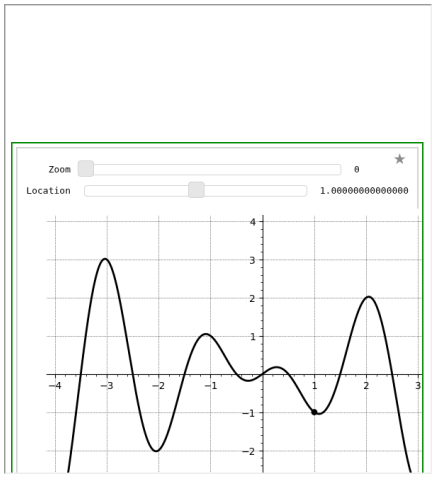
to the left of the figure number to open a separate page with only the selected interact. You can practice using these interactions on the surface plot below.



Standalone
In Context
Embed

Figure 8.0.1 An interactible three dimensional plot

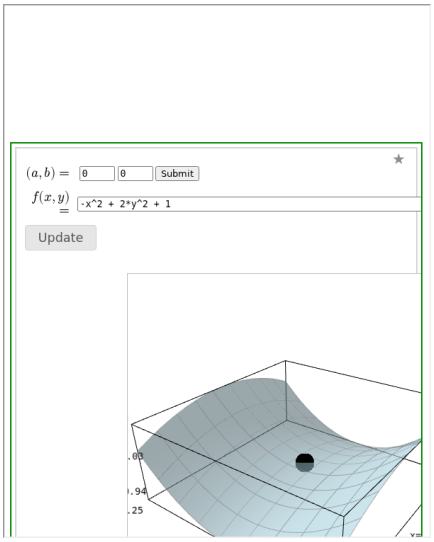
Some 2D and 3D plots will use other ways of allowing the user to change aspects of the figure or interact in different ways, such as allowing the user to specify thier own function, zoom in or out on different features, or add a new aspect to the plot. We try to explain in the context of the example or activity how to use these other features. We will most often use sliders and toggles to change features in plots. For instance, we use the following plot to allow users to explore the idea of locally linear functions using slider to change the location on the function we are investigating and another slider to control how closely we examine around the choosen point. Note that this plot will update when you change and release either of the sliders at a new value. A few plots will use and update button to allow the user to specify inputs before replotting.



Standalone
In Context
Embed

Figure 8.0.2 A plot with the ability to zoom in on a point of $y = f(x)$, where f is a nonlinear, differentiable function

In the following plot used in Section 11.10, the user can specify a function and a location that they think is a local max or min. This example plot used both the input structure for a point and a function.



Standalone
In Context
Embed

Figure 8.0.3 Plots of $z = f(x, y)$ around each of the critical points

Many of the ideas of multivariable calculus are best understood dynamically, and we encourage readers to make frequent use of technology to analyze graphs and data (maybe not on exams, though). Since technology changes so often, we refrain from indicating specific programs to use in the text. However, aside from computer algebra systems like *Sage* or *Mathematica*, there are many free graphing tools available for drawing three-dimensional surfaces or curves. These programs can be used by instructors and students to assist in the investigations and demonstrations. The use of these freely available tools is in accord with our philosophy that no one should be required to purchase materials to learn calculus. We are indebted to everyone who allows their expertise to be openly shared. Below is a list of a few of the technological tools that are available (links active at the writing of this edition). Of course, you can find your own by searching the web.

- CalcPlot3D, an excellent, all-purpose 3D graphing tool with numerous features built around multivariable calculus objects like curves in space, vector fields, surfaces (explicit, implicit, and parametric), and regions in 2D and 3D.
- Desmos a nice graphing tool with some computational capabilities and a separate plotter with a limited set of features for 3D.
- GeoGebra, all purpose graphing tool with some computer algebra capabilities and a 3D set of features. Clicking on the magnifying glass icon allows you to search a large database of GeoGebra applets.
- Sage Single Cell Server, a browser based way to run a single cell of Sage code to do calculations or produce plots that does not require a local installation of Sage. The authors also use JupyterLab notebooks with a Sage kernel for multiple purposes. You can find online installation instructions using WSL (on Windows systems) as well as Mac or Linux OS options
- Wolfram Alpha, an online tool for a variety of mathematical tasks, particularly useful for graphing surfaces in 2D and 3D, and for general calculations.

Contents

Features of the Text v

Acknowledgments ix

Active Calculus - Multivariable: our goals xi

How to Use this Text xiii

Graphics and Technology xv

9 Precalculus of Multivariable Functions 1

9.1 Why Multivariable? . . . 1
9.2 Three Dimensional Space . . . 2
9.3 Vectors . . . 16
9.4 The Dot Product . . . 33
9.5 The Cross Product . . . 48
9.6 Lines in Space . . . 62
9.7 Planes in Space . . . 74
9.8 Common Graphs in Three Dimensions . . . 85
9.9 Polar, Cylindrical, and Spherical Coordinates . . . 102

10 Vector-Valued Functions of One Variable 135

10.1 Classic Calculus Approach . . . 135
10.2 Vector-Valued Functions of One Variable . . . 140
10.3 Calculus of Vector-Valued Functions of One Variable . . . 156
10.4 Arc Length . . . 174
10.5 The TNB Frame . . . 188
10.6 Curvature . . . 200
10.7 Splitting the Acceleration Vector . . . 213

11 Derivatives of Multivariable Functions 227

11.1 Calculus of Several Variables . . . 227
11.2 Functions of Several Variables . . . 228
11.3 Limits . . . 253

11.4 First-Order Partial Derivatives 269
11.5 Second-Order Partial Derivatives 288
11.6 Linearization: Tangent Planes and Differentials 300
11.7 The Multivariable Chain Rule 318
11.8 Directional Derivatives and the Gradient 329
11.9 Higher Dimensions 353
11.10 Optimization 371
11.11 Constrained Optimization: Lagrange Multipliers 397

12 Multiple Integrals **409**

12.1 Integrating Functions of Several Variables 409
12.2 Double Riemann Sums and Double Integrals over Rectangles . .409
12.3 Iterated Integrals 425
12.4 Double Integrals over General Regions 433
12.5 Applications of Double Integrals 459
12.6 Double Integrals in Polar Coordinates 470
12.7 Triple Integrals. 486
12.8 Triple Integrals in Cylindrical and Spherical Coordinates . . .503
12.9 Change of Variables 513

13 Vector Calculus **529**

13.1 Integrating Multivariable Vector-Valued Functions 529
13.2 Vector Fields 529
13.3 The Idea of a Line Integral 539
13.4 Using Parameterizations to Calculate Line Integrals 554
13.5 Path-Independent Vector Fields and the Fundamental Theorem of Calculus for Line Integrals 565
13.6 Line Integrals of Scalar Functions 577
13.7 The Divergence of a Vector Field 594
13.8 The Curl of a Vector Field 607
13.9 Green's Theorem 626
13.10 Parameterizations of Surfaces and Surface Area 637
13.11 Flux Integrals 656
13.12 Surface Integrals of Scalar Valued Functions 670
13.13 Stokes' Theorem 681
13.14 The Divergence Theorem 692

Appendices

A. Supplementary Materials **705**

A.1 Conic Sections 705

Back Matter

Index **715**

Chapter 9

Precalculus of Multivariable Functions

9.1 Why Multivariable?

In your earlier calculus courses, you have explored how to compute a limit, how to measure the rate of change of a function, and how to calculate the area under the graph of a function. You likely also spent several years before calculus getting used to working algebraically and geometrically with a variety of different kinds of functions like polynomials and rational polynomials, as well as sinusoidal, exponential, and logarithmic functions. Nearly all this study involved functions that have one number as an input and one number as an output. In your first semester calculus, you considered how small changes in the input variable would lead to changes in the output variable.

In your previous mathematics and science courses, you probably also saw a few places where the math you were studying was not general enough to describe some reasonable scenarios. For instance, the calculus tools you learned in your earlier courses will not apply to graphs, like a circle or a hyperbola, where $y \neq f(x)$.

More generally, we will need to work with functions and expressions that depend on more than one input or involve outputs more complicated than a single scalar. For instance, when trying to find the optimal price for a product, you will need to consider fixed costs like a building lease, labor costs, and the varying costs of all the materials that are used in the product. Another example would be trying to measure the atmospheric temperature, which will vary over a three-dimensional space and will vary over time of day as well. If we wanted to use wind to help in modeling weather phenomena, we will need new tools since wind has both a direction and a strength associated with it, and thus cannot be measured by a single number. For each of the cases mentioned above, it is not entirely clear how to relate changes in the input variables to changes in the different kinds of outputs. A long-term goal of our work in this book will be to carefully define the relationship between changes of inputs and outputs for many new types of functions.

With our new expanded view of the world around us, we will also need to describe a variety of new examples of mathematical objects that are algebraically nice and exhibit different geometric features. Most students took several math classes to understand and apply to the calculus of functions with one-variable in and one-variable out; In this text, we will expand our tools in much less time because of the depth and variety of ideas that you have already

encountered. Before we look at the calculus of these new kinds of functions, we will need to do a bit of precalculus again.

9.2 Three Dimensional Space

Motivating Questions

- How can we define coordinates to describe a location in three-dimensional space?
- How can we draw and understand plots of points and basic figures in a three-dimensional space?
- How do you measure distance between locations in three or more dimensions?

9.2.1 Introduction

The functions you are most familiar with have the form $y = f(x)$, where x is the input or domain variable and y is viewed as the output variable. The graphs of these functions are plotted on a coordinate plane with the input expressed in terms of the horizontal coordinate and the output as the vertical coordinate. The following preview activity asks you to recall some algebraic and geometric ideas from before your single variable calculus course.

Preview Activity 9.2.1 In this Preview Activity, we will recall some ideas about measurements and important ideas in two dimensional space.

(a) What are the coordinates of the point drawn in Figure 9.2.1? Draw segments from the point P to the vertical and horizontal axes to demonstrate the measurements of the coordinates.

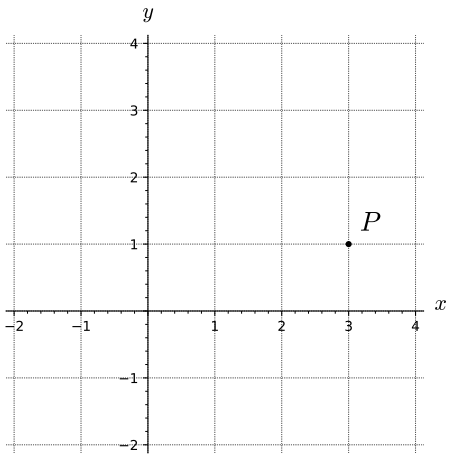


Figure 9.2.1 A 2D plot with point P labeled

(b) On the plot below, graph and label the following points: $P_1 = (0, 1)$, $P_2 = (1, 0)$, $P_3 = (2, -3)$, $P_4 = (3, -2)$, $P_5 = (-3, 2)$.

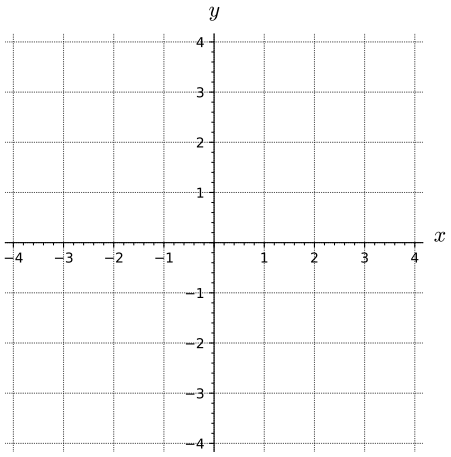


Figure 9.2.2 Two-dimensional axes for plotting points

(c) Give the coordinates of four points on the horizontal axis. What aspect do all of the points on the horizontal axis have in common? Use this idea to write an equation for the horizontal axis.

(d) On the axes below, draw a graph of each of the following equations: $x = 1$, $y = -2$, and $-x = y$

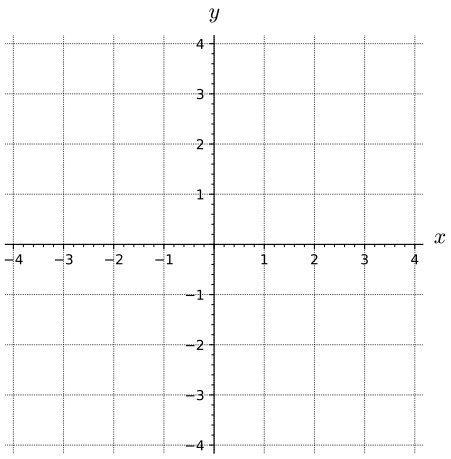


Figure 9.2.3 Two-dimensional axes for plotting equations

(e) Draw a plot of the points $(-1, 2)$ and $(7, -4)$. On your plot, draw the line segment that measures the distance between the given points and the segments that measure the horizontal and vertical changes. Your plotted segments should make a right triangle.

(f) Find the length of each of line segments in the plot that is your answer to part 9.2.1.e. Explain how the right triangle idea you drew can be generalized to find the distance between two points (x_1, y_1) and (x_2, y_2).

As you read in the introduction to this chapter, we will need to expand beyond the idea of measuring change involving expressions with one number as an input or output. One of the new kinds of functions we will examine is of the form $z = g(x, y)$, where the input is an ordered pair, (x, y), and the output of the function g is the real number z. In order to graph such a function, we will need to represent the basic ideas of coordinates and functions in a more

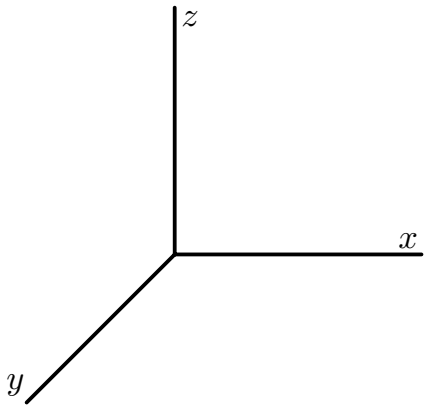
general setting to accommodate more dimensions.

The tasks in Preview Activity 9.2.1 connect ideas used in coordinates and graphing for two dimensions to specific measurements. We will use these connections to build useful tools for three or more dimensions. In the rest of this section, we will define the measurements for coordinates and basic measurements in three dimensions.

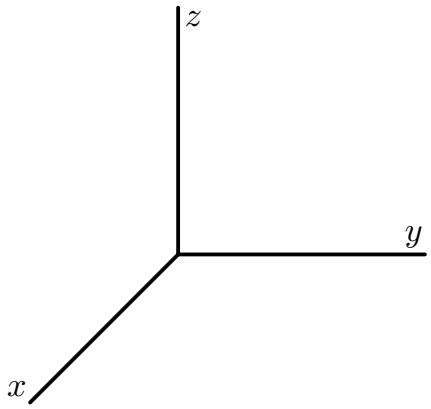
9.2.2 Three Dimensional Space and Coordinates

We let $\mathbb{R}^2$ denote the set of all ordered pairs of real numbers in the plane (two copies of the real number system) and let $\mathbb{R}^3$ represent the set of all ordered triples of real numbers, which constitutes three-dimensional space. For example, $(1, \pi)$ is in $\mathbb{R}^2$ and $(-e^2, \frac{\sqrt{7}}{4}, 0)$ is in $\mathbb{R}^3$, but $(\ln(3), 2i)$ is not in $\mathbb{R}^2$ because the second coordinate, $2i$, is an imaginary number and not a real number.

To plot points in three-dimensional space, we will set up a coordinate system with three mutually perpendicular axes—the x-axis, the y-axis, and the z-axis (called the **coordinate axes**). There are essentially two different ways we could set up a 3D coordinate system, as shown in Figure 9.2.4, which depicts only the positive portions of the axes. Thus, before we can proceed, we need to establish a convention.



(a) A left hand system



(b) A right hand system

Figure 9.2.4

The distinction between these two figures is subtle, but important. In the coordinate system shown in Figure 9.2.4(a), imagine that you are sitting on the positive z-axis next to the label "z." Looking down at the x- and y-axes, you see that the y-axis is obtained by rotating the x-axis by 90° in the *clockwise* direction. Again sitting on the positive z-axis in the coordinate system at right in Figure 9.2.4(b), you see that the y-axis is obtained by rotating the x-axis by 90° in the *counterclockwise* direction.

We call the coordinate system in Figure 9.2.4(b) a **right-hand system**; if we point the index finger of our *right* hand along the positive x-axis and our middle finger along the positive y-axis, then our thumb points in the direction of the positive z-axis. Following mathematical conventions, we choose to use a right-hand system throughout this book.

We let $\mathbb{R}^2$ denote the set of all ordered pairs of real numbers in the plane (two copies of the real number system) and let $\mathbb{R}^3$ represent the set of all ordered triples of real numbers (which constitutes three-space). For example,

$(1, \pi)$ is in $\mathbb{R}^2$ and $(-e^2, \frac{\sqrt{7}}{4}, 0)$ is in $\mathbb{R}^3$, but $(\ln(3), 2i)$ is not in $\mathbb{R}^2$ because the second coordinate is not a real number because $2i$ is imaginary.

In other disciplines or settings you may see the y-axis as being oriented vertically with the x-axis going left/right and the z-axis going in and out of the page like in the figure below. This orientation may be useful for you because you can think of the traditional orientation of the xy-plane (when drawn on paper) with the positive z-axis coming out of the page.

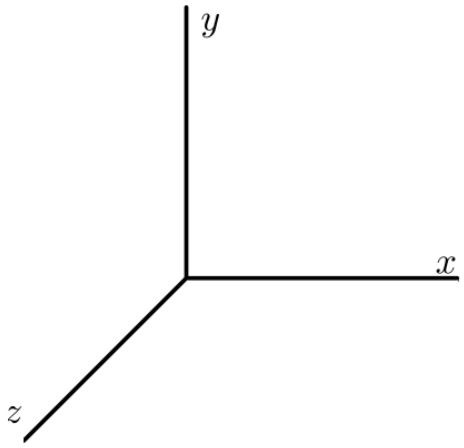


Figure 9.2.5 An alternate orientation of a right handed coordinate system for $\mathbb{R}^3$

Now that we have established how our axes will be defined, we can use these axes to define how to measure coordinates and make a few basic graphs.

In two dimensions, the horizontal coordinate was measured as the (perpendicular) distance to the vertical axis and the vertical coordinate was measured as the (perpendicular) distance to the horizontal axis. It is convenient to treat these measurements as signed distances with the measurement being positive when the point is above the given axis and negative when the measurement is below the given axis.

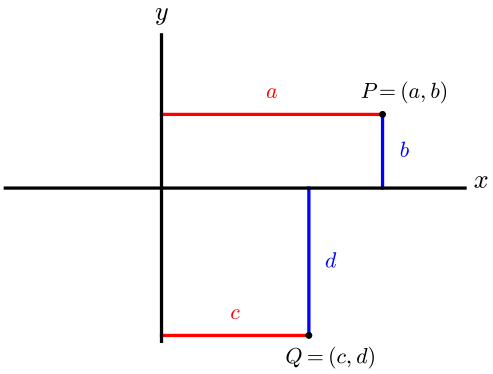


Figure 9.2.6 A plot with the 2D coordinate measurements labeled

Figure 9.2.6 shows the measurement of coordinates for P, a point in the first quadrant, and Q, a point in the fourth quadrant. Notice that the coordinates a, b, and c will be positive because each of the segments used to measure the distance are above the other coordinate axis. The coordinate d will be negative because the segment used to measure the vertical distance to the x-axis is below the x-axis. This is what we mean when we say that the coordinate

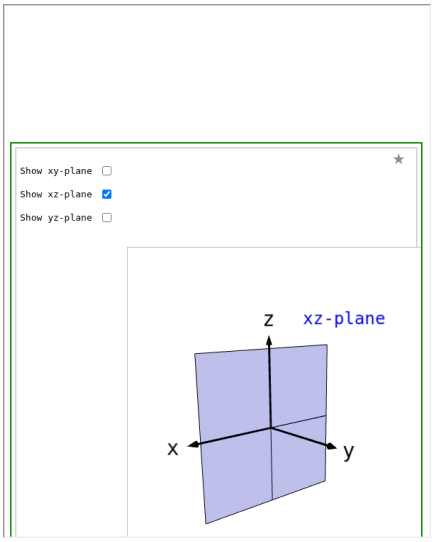
measurements are a ***signed distance***: the distance specifies how far the point is from each axis, while the sign specifies whether that distance is above or below the axis.

In a three dimensional space, we need to pay attention to not just the axes but also the **coordinate planes**. Figure 9.2.7 shows how the xz-plane is the plane that contains the x- and z-axes. We can define the xy-plane as the plane that contain the x-and y-axes and the yz-plane as the plane that contain the y-and z-axes.

In Figure 9.2.7, you can click on the buttons to either show or hide each of the coordinate planes. You should explore and turn each of the coordinate planes on and pay attention to how each plane is oriented with respect to the different axes. Figure 9.2.7 is the first interactable 3D plot, so we will take a moment to talk about other important ways you should interact with figures like this.

A wonderful feature of a figure like Figure 9.2.7 is that you can rotate the figure, zoom in or out, and change the point from which the figure is viewed. These will be very helpful actions to try when looking at interactive 3D graphs because this will help you visualize the fullest version of these graphs. There are often parts of these 3D figures that are "behind" other elements or might make more sense to you if viewed from a different perspective. By giving you control over how the graph is viewed, you will have the ability to look at features in whatever way helps you make the most sense at the time. We will always try to have the first orientation shown in these types of figure be helpful, but we will always be a bit limited by showing a 3D figure in a 2D medium.

To rotate a 3D interactable figure, you should left click and hold somewhere on the figure. While holding, you can rotate the figure by moving the point where you clicked up/down/left right. Additionally, you can move the point you are viewing the graph from by right clicking and holding while moving the cursor up/down/left/right. On touchscreen devices, you will need to touch with two fingers to act as a right click. The arrow keys on a keyboard will also move the point at which the figure is being viewed from. You can also zoom in or out having your cursor over the 3D interactable figure and using the scroll wheel on a mouse or using pinching/expanding motions with two fingers on touchscreen devices. Take a few minutes to explore these actions on the figure below and pay attention to how each of these actions will change your view of the graph.



Standalone
In Context
Embed

Figure 9.2.7 The coordinate planes in three dimensions

In many ways the coordinate planes in three dimensions are analogous to the axes in two dimensions. We will split up three-dimensional space based on whether a location is above or below different coordinate planes and we will measure coordinates in three dimensions as distances from the coordinate planes.

In two dimensions, the four regions that are separated by the coordinate axes are called **quadrants** and are typically labeled as in the plot below.

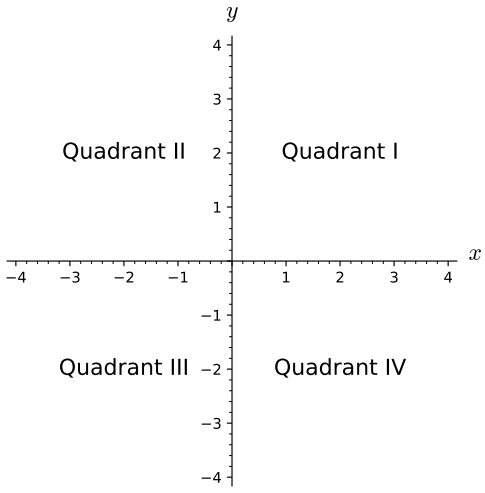
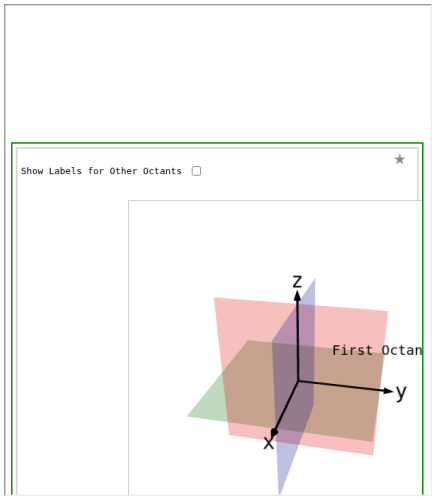


Figure 9.2.8 The four quadrants

The coordinate planes will divide our three dimensional space up into eight regions, which we will call **octants**. The first octant is the part of $\mathbb{R}^3$ that is in the positive $x-$, $y-$, and $z-$directions. You should rotate Figure 9.2.9 to see how there are eight regions that our 3D space is split into and how the orientation of the positive axis directions relates to the first octant.

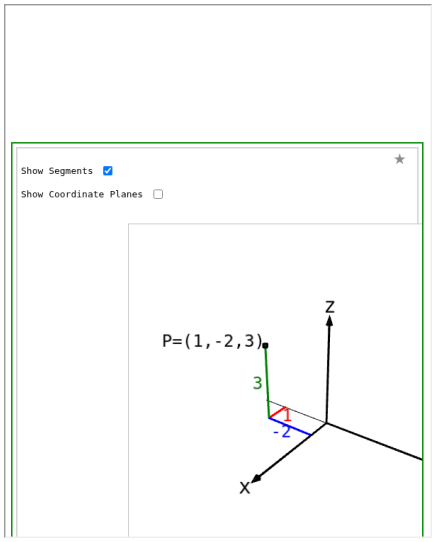


Standalone
In Context
Embed

Figure 9.2.9 A plot of the eight octants

The octants above the xy-plane are ordered just like the quadrants in two dimensions (when viewed looking down the $z-$axis) and octants 5 through 8 have the same relationships but are below the xy-plane. For example, octant 7 lies below quadrant III of the xy-plane. You can use the checkbox in Figure 9.2.9 to show the labels of the other octants.

The location of a point in a three-dimensional space is measured with three coordinates. The three-dimensional rectangular coordinates are measured as the signed distances to each of the coordinate planes. For example, the point $(1, -2, 3)$ is shown in Figure 9.2.10.



Standalone
In Context
Embed

Figure 9.2.10 A plot of the point $(1, -2, 3)$

When drawing points (or other plots) in three dimensions, it is often useful to draw segments or other features that are parallel to coordinate axes in order to help the viewer see the orientation and locations. For instance, we hope that you will find it easier to understand the location of the point $(1, -2, 3)$ in Figure 9.2.10 because of the parallel structures shown on the plot.

In Figure 9.2.10 you should click on the box labeled "Show Segments" to turn of the colored line segments. You should notice how hard it is to understand where the point is located without turning the axes or having the

colored line segments. You can rotate the figure with the point drawn without the segments and get some idea of where the point lies relative to the axes but this is a difficult visual task.

You should now click on the box labeled "Show Coordinate Planes" in Figure 9.2.10 to display the segments used to determine the coordinates in three dimensions as they are measured from the point to the coordinate planes. Most of our plots will not draw the coordinate planes. Framing the important measurements of your plot with structures that are parallel to the coordinate axes is a helpful way to make plots more easily understood.

If you select the "Show Segments" option and turn off the "Show Coordinate Planes" in Figure 9.2.10, you can see how it is easier to understand the location of the point $(1, -2, 3)$ because of the parallel structures that are shown on the plot. In particular, you can see the x- and y-coordinates measured in the xy-plane with the z-coordinate extended up. The segments drawn in this plot are usually reasonable to add to a drawing done by hand as well. Remember that interacting with and exploring different features you notice in the three-dimensional plots is an important part of building your understanding.

Activity 9.2.2 Drawing Points in 3D. In this problem to move "forward" or "backwards" is to move in the direction of positive x (the x-coordinate increases, the y and z remain the same) or negative x (the x-coordinate decreases, the y and z remain the same), respectively; "to the right" or "to the left" is moving in the positive or negative y direction, respectively; "up" or "down" is moving in the positive or negative z direction, respectively:

(a) Find the coordinates of the point A where one ends if one starts at point $(1, 2, 3)$ and moves 5 units forward, 4 units to the left, and 2 units up.

(b) Draw the point A (that is, your answer to (a)) on a set of three-dimensional axes and include the line segments that show that coordinate's points as a set of directions from the origin (like in Figure 9.2.10 with the "Show Segments" option).

(c) Find the coordinates of the point B where one ends if one starts at point $(3, -4, 2)$ and moves 4 units backwards, 4 units to the right, and 4 units down.

(d) Draw the point B (that is, your answer to (c)) on a set of three-dimensional axes and include the line segments that show that coordinate's points as a set of directions from the origin (like in Figure 9.2.10 with the "Show Segments" option).

9.2.3 Fundamental Planes

At this point, we are ready to understand graphs of some simple equations in three dimensions. For example, in $\mathbb{R}^2$, the graphs of the equations $x = a$ and $y = b$, where a and b are constants, are lines parallel to the coordinate axes. Remember that a graph of an equation is a plot of all points that satisfy this equation. This means that the graph of an equation is a visual representation of all locations whose coordinates will make the left side of your equation equal to the right side of the equation. The equation should give you a way to test whether a point is on the graph or not.

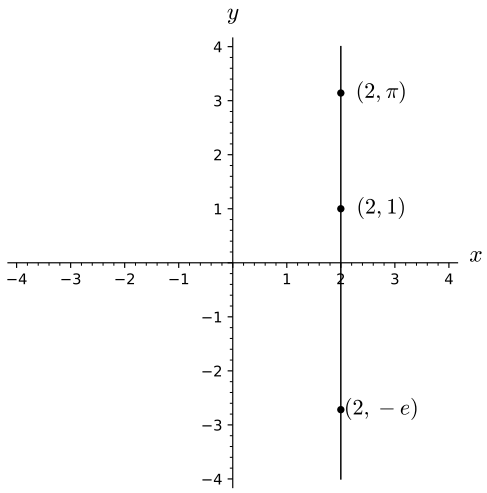


Figure 9.2.11 A 2D plot of $x = 2$

For instance, the graph of $x = 2$ in two dimensions will be a vertical line with x-intercept of 2, because the points of the form $(2, y)$ satisfy the equation $x = 2$ (for any choice of $y \in \mathbb{R}$) as shown in Figure 9.2.11. In the next activity we consider the three-dimensional analogs of these kinds equations.

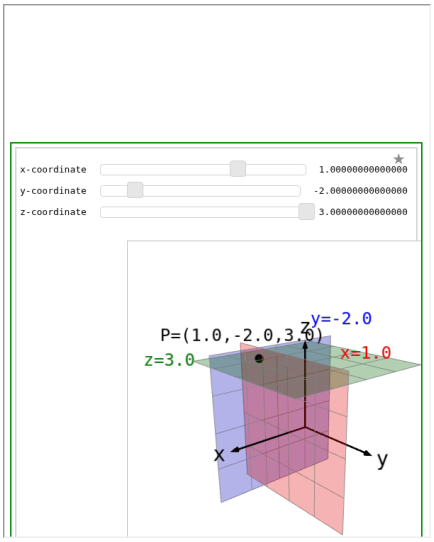
Activity 9.2.3

(a) Consider the set of points (x, y, z) that satisfy the equation $x = 2$. Write a sentence to describe this set as best as you can.

(b) Consider the set of points (x, y, z) that satisfy the equation $y = -1$. Write a sentence to describe this set as best as you can.

(c) Consider the set of points (x, y, z) that satisfy the equation $z = 0$. Write a sentence to describe this set as best as you can.

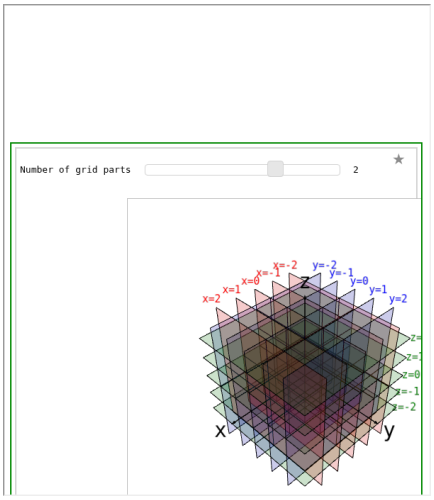
Activity 9.2.3 illustrates that equations where one of the rectangular coordinates is held constant lead to planes parallel to the coordinate planes. When we make the constant 0, we get the coordinate planes themselves . The xy-plane satisfies $z = 0$, the xz-plane satisfies $y = 0$, and the yz-plane satisfies $x = 0$ (see Figure 9.2.7). Planes of the form $x = a$, $y = b$, $z = c$ are called **fundamental planes** are useful in understanding and building structures in three dimensions. You can see how the intersection of fundamental planes $x = 1$, $y = -2$, and $z = 3$ corresponds to the point $(1, -2, 3)$ in Figure 9.2.12. As always, you should rotate and change the viewpoint of the figure to make sure you understand how each fundamental plane and the point P are oriented relative to the axes.



Standalone
In Context
Embed

Figure 9.2.12 A grid of fundamental planes in three dimensions

In a plot like Figure 9.2.3, you see a grid that helps measure coordinates of points in 2D. In Figure 9.2.13 you can look at how a different number of fundamental planes forming a grid looks. You can use the slider to increase or decrease the number of fundamental planes displayed in the grid. When you have more than one plane in each direction there will be a lot of visual clutter and very hard to distinguish fine features, especially something in the middle of the plot. Contrast this to how helpful the grid in Figure 9.2.15 is.



Standalone
In Context
Embed

Figure 9.2.13 A grid of fundamental planes in three dimensions

9.2.4 Distance in Three Dimensions

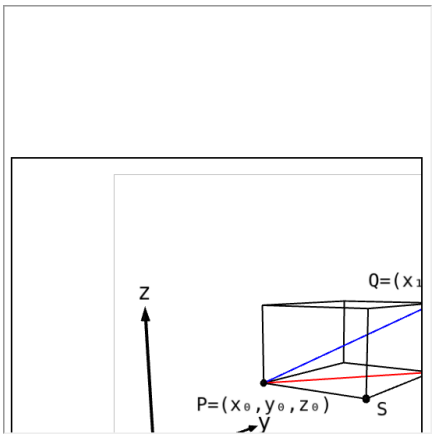
We conclude this section by using our knowledge of how to measure straight-line distance in $\mathbb{R}^2$ to find a formula for distance in three dimensions. On a related note, we define a circle in $\mathbb{R}^2$ as the set of all points equidistant from a fixed point. In $\mathbb{R}^3$, we call the set of all points equidistant from a fixed point a **sphere**. To find the equation of a sphere, we need to understand how to calculate the distance between two points in three-space, and we explore this idea in the next activity.

Activity 9.2.4 Remember that in two dimensions, the distance between $P = (x_0, y_0)$ and $Q = (x_1, y_1)$ is

$$\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}$$

and is related to the pythagorean theorem applied to a right triangle that measures changes in each coordinate direction (as demonstrated in part 9.2.1.f).

Let $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$ be two points in $\mathbb{R}^3$. These two points form opposite vertices of a rectangular box whose sides are fundamental planes as illustrated in Figure 9.2.14, and the distance between P and Q is the length of the blue diagonal shown in Figure 9.2.14.



Standalone
In Context
Embed

Figure 9.2.14 A plot of $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$ with connecting segments

(a) Consider the right triangle PRS in the base of the box whose hypotenuse is shown as the red line in Figure 9.2.14. What are the coordinates of R and S?

(b) Give the equation of the fundamental plane that contains the right triangle PRS.

Hint. One of the coordinates is the same for all three points P, R, S.

(c) Since the right triangle PRS lies in a plane, we can use the Pythagorean Theorem to find a formula for the length of the hypotenuse of this triangle. Find the length of the segment PR in terms of x_0, y_0, x_1, and y_1.

(d) Triangle PRQ has hypotenuse drawn with as blue segment connecting the points P and Q. Segment PR, which is the hypotenuse of triangle PRS that we considered earlier, is a leg of triangle PRQ. This triangle lies entirely in a plane, so we can again use the Pythagorean Theorem to find the length of its hypotenuse. Show that the length of PQ is

$$\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}.$$

The method used in Activity 9.2.4 does not depend on anything but the coordinates between the two points, so we can use the last result to measure the distance between any two points in $\mathbb{R}^3$.

The distance between points in three dimensions.

The distance between points $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$ (denoted as $|PQ|$) in $\mathbb{R}^3$ is given by the formula

$$|PQ| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2} \qquad (9.2.1)$$

As Activity 9.2.4 showed, the distance in two or three (or more!) dimensions depends on the change in each coordinate from one point to the other. Note that the distance does not depend on whether we consider P to Q or Q to P.

Equation (9.2.1) can be used to derive the equation for a **sphere** centered at a point (x_0, y_0, z_0) with radius R. Since the distance from any point (x, y, z) on such a sphere to the center (x_0, y_0, z_0) is R, the point (x, y, z) must satisfy the equation

$$\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = R$$

Squaring both sides, we come to the standard equation for a sphere.

The equation of a sphere.

The equation of a sphere with center (x_0, y_0, z_0) and radius R is

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2$$

We see a strong similarity when we compare this equation to its two-dimensional analogue, the equation of a circle of radius R in the plane centered at (x_0, y_0):

$$(x - x_0)^2 + (y - y_0)^2 = R^2.$$

Summary

- Coordinates in three-dimensional space are measured in terms of distances from a location to the coordinate planes.
- In $\mathbb{R}^3$, the distance between points $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$ (denoted as $|PQ|$) is given by the formula

  $$|PQ| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}.$$

  Consequently, the equation of a sphere with center (x_0, y_0, z_0) and radius R is

  $$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2.$$

9.2.5 Exercises

1. (A) If the positive z-axis points upward, an equation for a horizontal plane through the point $(3, 2, 2)$ is

_____.

(B) An equation for the plane perpendicular to the x-axis and passing through the point $(3, 2, 2)$ is

_____.

(C) An equation for the plane parallel to the xz-plane and passing through the point $(3, 2, 2)$ is

_____.

2. You are given the following points: $A = (10, 12, -16)$, $B = (16, 0, 5)$, $C = (-6, -8, 12)$.

Which point is closest to the yz-plane? (☐ A ☐ B ☐ C)
What is the distance from the yz-plane to this point? _____
Which point is farthest from the xy-plane? (☐ A ☐ B ☐ C)
What is the distance from the xy-plane to this point? _____
Which point lies on the xz-plane? (☐ A ☐ B ☐ C)

3. Find a formula for the shortest distance from a point (a, b, c) to the z-axis.
distance = _____

4. Find the distance from $(-9, 7, -10)$ to each of the following:

1. The xy-plane.
Answer: _____
2. The yz-plane.
Answer: _____
3. The xz-plane.
Answer: _____
4. The x-axis.
Answer: _____
5. The y-axis.
Answer: _____
6. The z-axis.
Answer: _____

5. Find the equation of the sphere of radius 3 centered at $(3, 3, 2)$.
This sphere is given by: _____

6. Find the center and radius of the sphere
$x^2 + 8x + y^2 - 18y + z^2 - 12z = -52$
Center: (_____, _____, _____)
Radius: _____

7. Consider the sphere $(x-4)^2 + (y-3)^2 + (z-5)^2 = 25$

(a) Does the sphere intersect each of the following planes at zero points, at one point, at two points, in a line, or in a circle?

The sphere intersects the yz-plane (☐ at one point ☐ in a circle ☐ in a line ☐ at two points ☐ at zero points)

The sphere intersects the xy-plane (☐ at zero points ☐ in a line ☐ at two points ☐ in a circle ☐ at one point)

The sphere intersects the xz-plane (☐ at two points ☐ in a line ☐ at zero points ☐ at one point ☐ in a circle)

(b) Does the sphere intersect each of the following coordinate axes at zero points, at one point, at two points, or in a line?

The sphere intersects the z-axis (☐ at two points ☐ in a line ☐ at zero points ☐ at one point)

The sphere intersects the x-axis (☐ at one point ☐ at two points ☐ in a line ☐ at zero points)

The sphere intersects the y-axis (☐ at two points ☐ in a line ☐ at zero points ☐ at one point)

8. In this problem, we will look at how the set of points that satisfies an equation, $x = 2$, will have a different description whether we think about our points being in one dimension (on a number line), in two dimensions (in the Cartesian plane), or in three dimensions (in space).

(a) Describe the set of points that satisfies $x = 2$ if you consider the space to be $\mathbb{R}$ (the number line). Your description should include a

sentence or two detailing the set and a plot of the graph of $x = 2$.

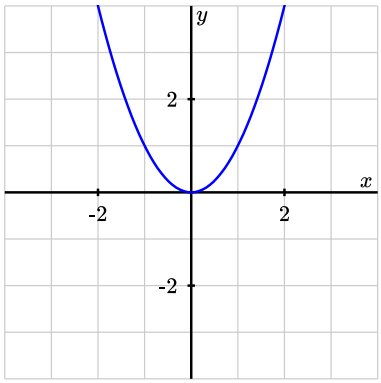
(b) Describe the set of points that satisfies $x = 2$ if you consider the space to be $\mathbb{R}^2$ (the cartesian plane). Your description should include a sentence or two detailing the set and a plot of the graph of $x = 2$.

(c) Describe the set of points that satisfies $x = 2$ if you consider the space to be $\mathbb{R}^3$ (3D space). Your description should include a sentence or two detailing the set and a plot of the graph of $x = 2$.

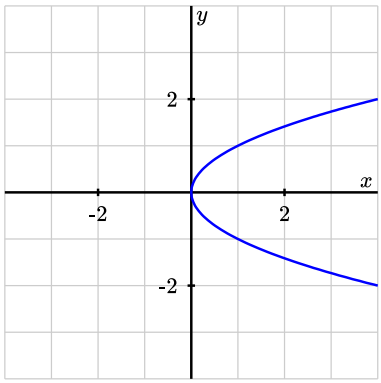
9. In this exercise, we want to graphically identify when a given plot can be expressed as:

- y as a function of x, but x cannot be written as a function of y
- x as a function of y, but y cannot be written as a function of x
- y as a function of x and x as a function of y
- neither y as a function of x nor x as a function of y

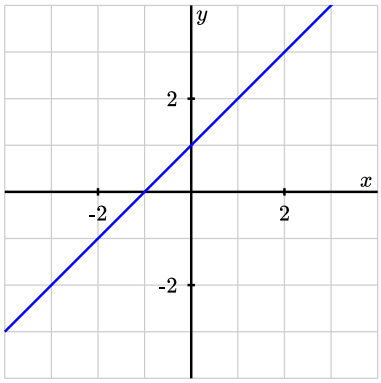
For each graph in Figure 9.2.15, state which of the four options above applies to whether the graph can be expressed with each coordinate being a function of the other or not. If the graph cannot be expressed with y as a function of x, write a sentence about what property you used to determine this. If the graph cannot be expressed with x as a function of y, write a sentence about what property you used to determine this.



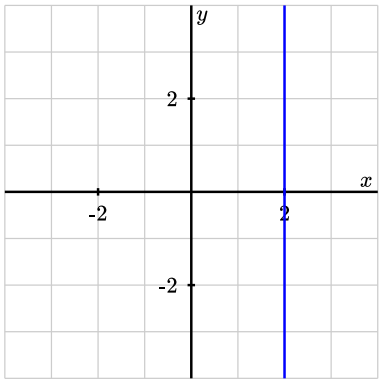
(a)



(b)



(c)



(d)

Figure 9.2.15 Four plots to determine if the graphs depict y as a function of x

10. In this exercise, we want to algebraically identify when the graph of a given equation can be expressed as:

- y as a function of x, but x cannot be written as a function of y
- x as a function of y, but y cannot be written as a function of x
- y as a function of x and x as a function of y
- neither y as a function of x nor x as a function of y

For each equation below, state which of the four options above applies to whether the graph can be expressed with each coordinate being a function of the other or not.

(a) $3x - 2y = 10$

(b) $3x^2 + 2y^2 = 10$

(c) $3y = 10x^2 - 10$

(d) $y = 5$

(e) $x = 2$

11. Find the equation of each of the following geometric objects.

(a) The plane parallel to the xy-plane that passes through the point $(-4, 5, -12)$.

(b) The plane parallel to the yz-plane that passes through the point $(7, -2, -3)$.

(c) The sphere centered at the point $(2, 1, 3)$ and has the point $(-1, 0, -1)$ on its surface.

(d) The sphere whose diameter has endpoints $(-3, 1, -5)$ and $(7, 9, -1)$.

9.3 Vectors

Motivating Questions

- What is a vector?
- What does it mean for two vectors to be equal?
- How do we add two vectors together or multiply a vector by a scalar?
- How do we determine the length of a vector? What is a unit vector, and how do we find a unit vector in the direction of a given vector?

9.3.1 Introduction

Quantities like length, speed, area, and mass are all measured using a single number (a **scalar**). Other quantities, like velocity, force, and displacement, have two attributes: magnitude (or length) and direction. These quantities are represented by **vectors** and are the focus of this section. For example, we will use vectors to calculate work done by a constant force, calculate torque, determine direction vectors for lines and normal vectors for planes, define curvature, and determine the direction of greatest increase on a surface. For most

of these applications, we will be interested in using vectors to measure the direction of some aspect of the application while needing to also pay attention to the magnitude of this measurement. Vectors will be an essential tool for describing the behavior of functions of several variables in Chapters 11 through 13.

Preview Activity 9.3.1 Postscript is a programming language whose primary purpose is to specify how to generate text or graphics. The following is a simple set of Postscript commands that produces the triangle in the plane with vertices $(0, 0)$, $(1, 1)$, and $(1, -1)$:

```
(0,0) moveto
(1,1) lineto stroke
(1,-1) lineto stroke
(0,0) lineto stroke
```

The process described by these commands is

1) tell Postscript to start at the point $(0, 0)$,
2) draw a line from the point $(0, 0)$ to the point $(1, 1)$ (this is what the `line to` and `stroke` commands do),
3) draw lines from $(1, 1)$ to $(1, -1)$, and
4) draw from $(1, -1)$ back to the origin.

Each of these commands encodes two important pieces of information: a direction in which to move and a distance to move. Mathematically, we can capture this information succinctly in a vector. To do so, we record the movement on the map in a pair $\langle x, y\rangle$, where x is the horizontal displacement and y the vertical displacement from one point to another. This pair $\langle x, y\rangle$ is a **vector**, and we call each number in the pair a **component**. As an example, the vector from the origin to the point $(1, 1)$ is represented by $\langle 1, 1\rangle$ while the vector from the point $(1, 1)$ to the origin is represented by $\langle -1, -1\rangle$.

(a) What is the vector $\vec{v}_1 = \langle x, y\rangle$ that describes the displacement from the point $(1, 1)$ to the point $(1, -1)$?

(b) How can we use the two components of $\vec{v}_1$ to determine the distance from the point $(1, 1)$ to the point $(1, -1)$?

(c) Suppose we want to draw the triangle with vertices $A = (2, 3)$, $B = (-3, 1)$, and $C = (4, -2)$. As a shorthand notation, we will denote the vector from the point A to the point B as $\overrightarrow{AB}$. Determine the vectors $\overrightarrow{AB}$, $\overrightarrow{BC}$, and $\overrightarrow{AC}$.

(d) How are the horizontal displacements of $\overrightarrow{AB}$ and $\overrightarrow{BC}$ related to the horizontal displacement of $\overrightarrow{AC}$? Does the same relationship follow for the vertical displacements? Write a couple of sentences to explain your reasoning in the context of this problem.

9.3.2 Representations of Vectors

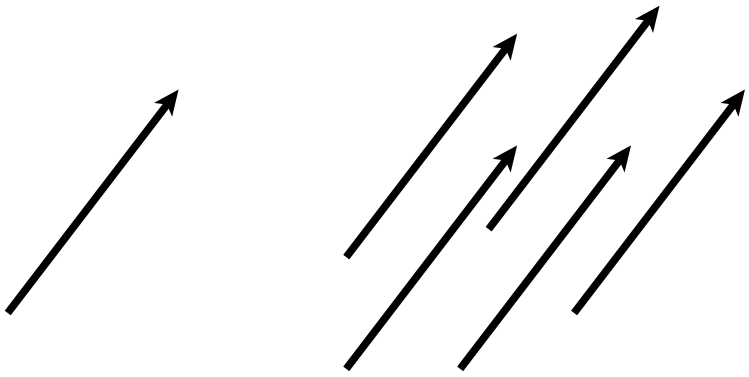
Preview Activity 9.3.1 shows how we can record the length and direction of a change in position using a pair of numbers $\langle x, y\rangle$. In particular, these numbers

measure the change in each coordinate direction separately. There are many quantities other than displacement, such as force and velocity, that possess the attributes of magnitude and direction. We will call such quantity a **vector**.

Definition 9.3.1 A **vector** is a quantity that possesses the attributes of magnitude and direction. ◆

This definition is a common geometric idea but some students will have seen an introduction to linear algebra at this point. In linear algebra, a vector is any kind of object that satisfies the definition of an element of a vector space (see the Wikipedia definition of a vector space). The idea of a vector can include A LOT of objects that do not look like arrows in two or three dimensions, but the ideas, like linear combinations and measuring length of a vector, are central to why linear algebra is a hugely impactful branch of mathematics.

We can represent a vector geometrically as a directed line segment, with the magnitude as the length of the segment and an arrowhead indicating direction, as shown at left in Figure 9.3.2.



(a) A vector

(b) Five different representations of the same vector

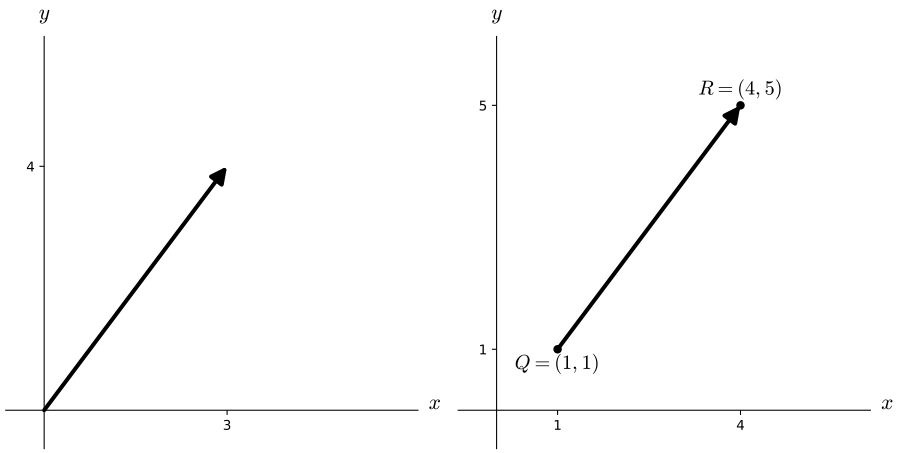
Figure 9.3.2

According to the definition, a vector possesses the attributes of magnitude (length) and direction; the vector's position, however, is not mentioned. Consequently, any two vectors having the same magnitude and direction are equal, as shown in Figure 9.3.2(b). In other words, two vectors are **equal** provided they have the same magnitude and direction. This means that the same vector may be drawn in the plane in many different ways.

Vectors do not have a particular starting or ending point. However, it will be useful to use directed line segments like $\overrightarrow{AB}$ as a specific representative of a vector. For instance, suppose that we would like to draw the vector $\langle 3, 4\rangle$, which represents a horizontal change of (positive) three units and a vertical change of (positive) four units. We may place the **tail** of the vector (the point from which the vector originates) at the origin and the **tip** (the terminal point of the vector) at $(3, 4)$, as illustrated in Figure 9.3.3(a). A vector with its tail at the origin is said to be in **standard position**.

The point at the tail of a vector is often called the initial point or starting point. The point at the end of the vector is often called the terminal point or final point. The entries of the vector measure the change in each coordinate separately. Remember that change is measured as "final minus initial", in

general.



(a) The vector $\langle 3, 4\rangle$ in standard position

(b) The vector $\langle 3, 4\rangle$ represented from $(1, 1)$ to $(4, 5)$

Figure 9.3.3

Alternatively, we may place the tail of the vector $\langle 3, 4\rangle$ at another point, such as $Q = (1, 1)$. After a displacement of three units to the right and four units up, the tip of the vector is at the point $R = (4, 5)$ (see the vector in Figure 9.3.3(b)).

In this example, the vector was represented by the directed line segment from Q to R, which we denote as $\overrightarrow{QR}$. We may also turn the situation around: given the two points Q and R, we obtain the vector $\langle 3, 4\rangle$ because we move horizontally three units and vertically four units to get from Q to R. In other words, $\overrightarrow{QR} = \langle 3, 4\rangle$. In general, the vector $\overrightarrow{QR}$ from the point $Q = (q_1, q_2)$ to $R = (r_1, r_2)$ is found by taking the difference of coordinates, so that

$$\overrightarrow{QR} = \langle r_1 - q_1, r_2 - q_2\rangle.$$

We will use arrows over the name/letters to represent vectors, such as $\vec{v} = \langle 3, 4\rangle$, to distinguish them from scalars. The entries of a vector are called its **components**; in the vector $\langle 3, 4\rangle$, the x-component is 3 and the y-component is 4. We use pointed brackets $\langle\ ,\ \rangle$ and the term *components* to distinguish a vector from a point $(\ ,\ )$ and its *coordinates*.

There is, however, a close connection between vectors and points. Given a point P, we will frequently consider the vector $\overrightarrow{OP}$ from the origin O to P. For instance, if $P = (3, 4)$, then $\overrightarrow{OP} = \langle 3, 4\rangle$ as in Figure 9.3.4. In this way, we think of a point P as defining a vector $\overrightarrow{OP}$ whose components agree with the coordinates of P. The vector $\overrightarrow{OP}$ is called the **position vector** of P.

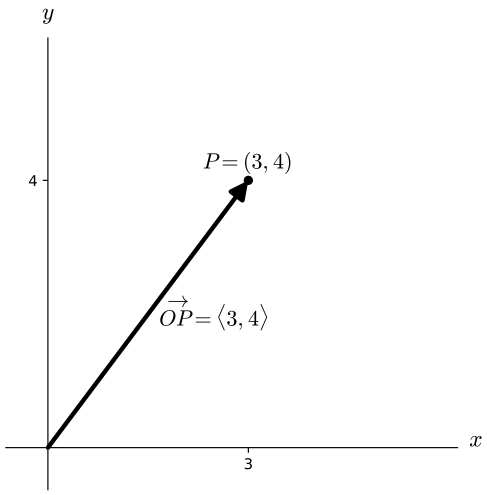


Figure 9.3.4 A point defines a vector in standard position

While we often illustrate vectors in the plane since pictures will be easier to draw and demonstrate the necessary characteristics, many situations will require the use of vectors in three or more dimensions. For instance, a vector $\vec{v}$ in n-dimensional space, $\mathbb{R}^n$, has n components and may be represented as

$$\vec{v} = \langle v_1, v_2, v_3, \ldots, v_n \rangle.$$

Our next activity will help us to become accustomed to vectors and their relationship to points in three dimensions.

Activity 9.3.2 After being bored in your previous calculus class (and definitely not in multivariable calculus) you count the floor tiles to see how large the room is. You now know that the room is 32 feet wide and 42 feet in length. Additionally, your tall friend Chuck can barely jump and touch the ceiling which means the ceiling is 10 feet high. Your calculus professor notices you doing these measurements and decides to create a classroom coordinate system. Your professor walks to the center of the room and notices that their head is five feet above the ground and sets their head as the origin of the classroom coordinate system. Your friend Alice is sitting at $A = (9, -6, -2.5)$, a projector is located at position $B = (0, 1, 3)$, and your friend Carlos is working while sitting on the floor at $C = (-2, 20, -5)$. All distances are measured in feet.

For each of the directed line segments given below, determine the components of the indicated vectors and explain in context what each represents.

(a) $\overrightarrow{OA} = \langle \, , \, , \, \rangle$

(b) $\overrightarrow{OB} = \langle \, , \, , \, \rangle$

(c) $\overrightarrow{OC} = \langle \, , \, , \, \rangle$

(d) $\overrightarrow{AB} = \langle \, , \, , \, \rangle$

(e) $\overrightarrow{AC} = \langle \, , \, , \, \rangle$

(f) $\overrightarrow{BC} = \langle \, , \, , \, \rangle$

9.3.3 Equality of Vectors

Because initial and final locations are not mentioned in the definition of a vector, any two vectors that have the same magnitude and direction are equal. It is helpful to have an algebraic way to determine when this occurs. That is, if we know the components of two vectors $\vec{u}$ and $\vec{v}$, we will want to be able to determine algebraically when $\vec{u}$ and $\vec{v}$ are equal. There is an straightforward set of conditions that we use in terms of components.

Equality of Vectors.

Two vectors $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$ in $\mathbb{R}^2$ are equal if and only if their corresponding components are equal: $u_1 = v_1$ and $u_2 = v_2$.

More generally, two vectors $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ and $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ in $\mathbb{R}^n$ are equal if and only if $u_i = v_i$ for every i from 1 to n.

9.3.4 Operations on Vectors

Vectors are not numbers (scalars), but we can now represent vectors with components that are real numbers. As such, we naturally wonder if it is possible to add two vectors together, multiply two vectors, or combine vectors in any other ways. In this section, we will study two operations on vectors: vector addition and scalar multiplication. To begin, we investigate a natural way to think about combining vectors in the context of measuring the change of location.

Activity 9.3.3 In this problem, we will be navigating to help find a friend who is lost at a local state park. We will be navigating using traditional map coordinates with east going in the positive x-direction and north going in the positive y-direction.

(a) Your friend told you they would be staying 3 km east and 4 km north of the main parking lot. You drive to the parking lot and park next to your friend's car, then hike 3 km east and 4 km north. You get to the location you expected your friend to be at, but you don't find your friend and call the ranger station. The ranger station says they think your friend is 2 km west and 1 km north of your current location. We will denote the location of the car with point C, your friend's originally anticipated location as point A, and the ranger's suggested location as point S.

On the grid below, we have labeled point C. Draw the location of each of the other points A and S.

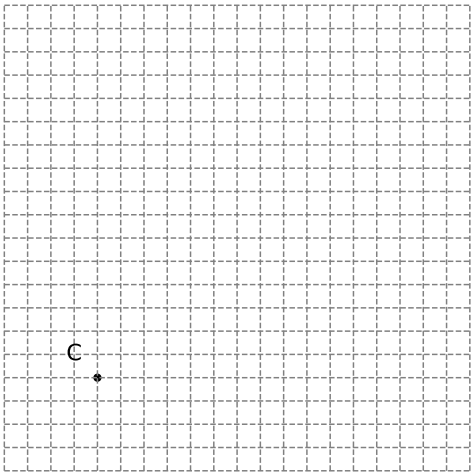


Figure 9.3.5 A grid for mapping locations needed to find your friend

(b) Using your picture from the previous task, give the components of the following vectors:

(a) $\overrightarrow{CA}$

(b) $\overrightarrow{CS}$

(c) $\overrightarrow{AS}$

(c) In the context of this problem, explain why you can add the horizontal components of $\overrightarrow{CA}$ and $\overrightarrow{AS}$ to get the horizontal component of $\overrightarrow{CS}$. Write a sentence or two about why this argument should work for the vertical components as well.

(d) After hiking to the location suggested by the ranger, you still don't see your friend and call the ranger station again. The regional manager of rangers answers this time. She says the first ranger made a mistake in their navigation. They sent a drone to the location of your car and the drone spotted your friend along the same direction from your car to point A, but your friend is three times as far away from your car as you were when you were at point A.

In order to avoid confusion or any other mistakes, you want to compute the vector from your car to the drone's suggested location, which we will call D. What are the components of $\overrightarrow{CD}$? Write a sentence or two to compare this vector to $\overrightarrow{CA}$.

(e) So we need to get from point S to point D. Explain how to use the components of $\overrightarrow{CS}$ and $\overrightarrow{CD}$ to find the components of $\overrightarrow{SD}$.

(f) You find your friend at point D and want to take them back to your car to go home. Explain how you can use the components of $\overrightarrow{CD}$ to give directions from point D back to your car.

In Activity 9.3.3 we saw how adding two vectors componentwise gave a vector that represented the total change of traveling along one vector followed by doing the change described by the other vector. Additionally, we saw that to go three times as far in a direction given by a vector, we can describe the total change by multiplying each component by three. Finally, we saw that we can

find the difference between two vectors by componentwise subtraction. This difference vector completed a triangle formed by positioning the two subtracted vectors with the same initial point. The ideas of adding vectors or multiplying a scalar by a vector generalizes beyond the context of Activity 9.3.3.

In general, we can add vectors $\vec{v} = \langle v_1, \ldots, v_n \rangle$ and $\vec{w} = \langle w_1, \ldots, w_n \rangle$ componentwise to get $\vec{v} + \vec{w} = \langle v_1 + w_1, \ldots, v_n + w_n \rangle$, where $\vec{v} + \vec{w}$ gives the total change described by $\vec{v}$ then $\vec{w}$. We define the scalar multiplication of vectors componentwise as well. If k is a real number, then $k\vec{v} = \langle kv_1, \ldots, kv_n \rangle$. Geometrically, the scalar multiplication of k by $\vec{v}$ will scale (stretch/shrink) $\vec{v}$ by a factor of k. We will discuss the way different values of k scale vectors in Subsection 9.3.6.

We can now add vectors and multiply vectors by scalars, and thus we can add together scalar multiples of vectors. This allows us to define *vector subtraction*, $\vec{v} - \vec{u}$, as the sum of $\vec{v}$ and -1 times $\vec{u}$, so that

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

You can alternatively think of of $\vec{v} - \vec{u}$ as the vector you need to add to $\vec{u}$ to get a result of $\vec{v}$ (as you did in part e of Activity 9.3.3).

Using vector addition and scalar multiplication, we will often represent vectors in terms of the special vectors $\hat{\imath} = \langle 1, 0 \rangle$ and $\hat{\jmath} = \langle 0, 1 \rangle$. For instance, we can write the vector $\langle a, b \rangle$ in $\mathbb{R}^2$ as

$$\langle a, b \rangle = a\langle 1, 0 \rangle + b\langle 0, 1 \rangle = a\hat{\imath} + b\hat{\jmath},$$

which means that

$$\langle 2, -3 \rangle = 2\hat{\imath} - 3\hat{\jmath}.$$

In $\mathbb{R}^3$, we let $\hat{\imath} = \langle 1, 0, 0 \rangle$, $\hat{\jmath} = \langle 0, 1, 0 \rangle$, and $\hat{k} = \langle 0, 0, 1 \rangle$, and we can write the vector $\langle a, b, c \rangle$ in $\mathbb{R}^3$ as

$$\langle a, b, c \rangle = a\langle 1, 0, 0 \rangle + b\langle 0, 1, 0 \rangle + c\langle 0, 0, 1 \rangle = a\hat{\imath} + b\hat{\jmath} + c\hat{k}.$$

The vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are called the **standard unit vectors**, and are commonly used in the physical sciences.

9.3.5 Properties of Vector Operations

We know that the sum (of scalars) $1 + 2$ is equal to the sum $2 + 1$. This is called the **commutative** property of scalar addition. Any time we define operations on objects (like addition of vectors) we usually want to know what kinds of properties the operations have. For example, is addition of vectors a commutative operation? To answer this question we take two *arbitrary* vectors $\vec{v}$ and $\vec{u}$ and add them together and see what happens. Let $\vec{v} = \langle v_1, v_2 \rangle$ and $\vec{u} = \langle u_1, u_2 \rangle$. Now we use the fact that v_1, v_2, u_1, and u_2 are scalars, and that the addition of scalars is commutative to see that

$$\begin{aligned}
\vec{v} + \vec{u} &= \langle v_1, v_2 \rangle + \langle u_1, u_2 \rangle \\
&= \langle v_1 + u_1, v_2 + u_2 \rangle \\
&= \langle u_1 + v_1, u_2 + v_2 \rangle \\
&= \langle u_1, u_2 \rangle + \langle v_1, v_2 \rangle \\
&= \vec{u} + \vec{v}.
\end{aligned}$$

So the vector sum is a commutative operation. Similar arguments can be used to show the following properties of vector addition and scalar multiplication.

Properties of vector operations.

Let $\vec{v}$, $\vec{u}$, and $\vec{w}$ be vectors in $\mathbb{R}^n$ and let a and b be scalars. Then

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

(2) $(\vec{v} + \vec{u}) + \vec{w} = \vec{v} + (\vec{u} + \vec{w})$

(3) The vector $\vec{0} = \langle 0, 0, \ldots, 0 \rangle$ has the property that $\vec{v} + \vec{0} = \vec{v}$. The vector $\vec{0}$ is called the **zero vector**.

(4) $(-1)\vec{v} + \vec{v} = \vec{0}$. The vector $(-1)\vec{v} = -\vec{v}$ is called the **additive inverse** of the vector $\vec{v}$.

(5) $(a + b)\vec{v} = a\vec{v} + b\vec{v}$

(6) $a(\vec{v} + \vec{u}) = a\vec{v} + a\vec{u}$

(7) $(ab)\vec{v} = a(b\vec{v})$

(8) $1\vec{v} = \vec{v}$.

We verified the first property for vectors in $\mathbb{R}^2$. The other properties can be verified to hold for all vectors in $\mathbb{R}^n$ in a similar manner because the operations are done componentwise on the vectors.

Activity 9.3.4 In this activity, use Properties of vector operations to simplify each of the following expressions into a single vector of the form $\langle a, b, c \rangle$. For this activity, let $\vec{u} = \langle 1, -2, 3 \rangle$, $\vec{v} = \langle 2, -2, 5 \rangle$, and $\vec{w} = \langle -3, 0, 3 \rangle$.

(a) $\vec{u} + 2\vec{v} + 3\vec{w}$

(b) $2\vec{w} - 3\vec{u}$

(c) $-4(\vec{v} - \vec{w})$

(d) $-2(3\vec{u})$

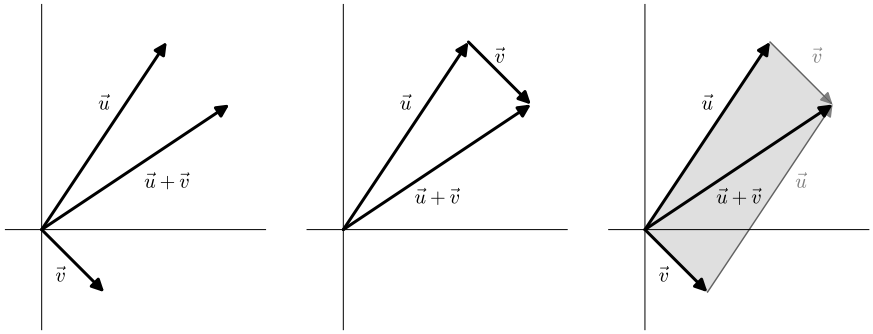
(e) Give the components of the vector that fills in the blank: $(2\vec{w} - 3\vec{u}) - _____ = \vec{0}$

(f) Explain why the following expression does not make sense:

$$2\langle 2, 4, -7 \rangle - 3\langle -\pi, \sqrt{7} \rangle$$

9.3.6 Geometric Interpretation of Vector Operations

Next, we explore the geometric representations of vector addition and scalar multiplication which will allow us to visualize these operations. Let $\vec{u} = \langle 4, 6 \rangle$ and $\vec{v} = \langle 3, -2 \rangle$. Then $\vec{w} = \vec{u} + \vec{v} = \langle 7, 4 \rangle$, as shown in Figure 9.3.6(a).



(a) A vector sum

(b) Summing displacements

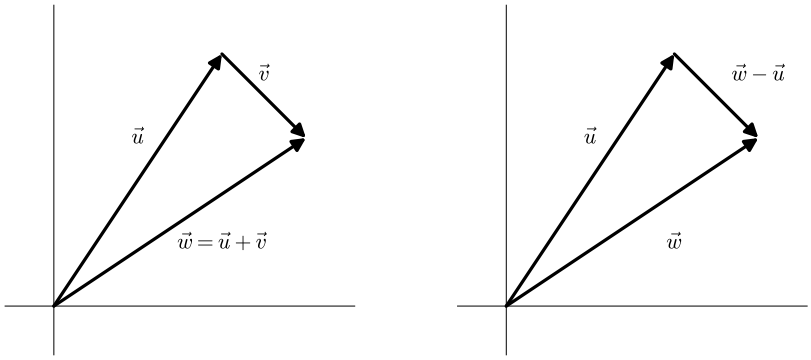
(c) The parallelogram law

Figure 9.3.6

If we think of these vectors as displacements measuring change in position in the plane, we can see a geometric way to envision vector addition. For instance, the vector $\vec{u}+\vec{v}$ will represent the displacement obtained by following the displacement $\vec{u}$ with the displacement $\vec{v}$. We may picture this by placing the tail of $\vec{v}$ at the tip of $\vec{u}$, as seen in Figure 9.3.6(b). In other words, the change described by $\vec{u}+\vec{v}$ is given by going along $\vec{u}$ then immediately along $\vec{v}$.

Vector addition is commutative. This means we obtain the same sum if we place the tail of $\vec{u}$ at the tip of $\vec{v}$ as if we place the tail of $\vec{v}$ at the tip of $\vec{u}$. We therefore see that $\vec{u}+\vec{v}$ appears as the diagonal of the parallelogram determined by $\vec{u}$ and $\vec{v}$, as shown in Figure 9.3.6(c).

Vector subtraction has a similar interpretation. In Figure 9.3.7(a) we see vectors $\vec{u}$, $\vec{v}$, and $\vec{w}=\vec{u}+\vec{v}$. If we rewrite $\vec{v}=\vec{w}-\vec{u}$, we have the arrangement shown at right in Figure 9.3.7(b). In other words, to form the difference $\vec{w}-\vec{u}$, we draw a vector from the tip of $\vec{u}$ to the tip of $\vec{w}$. This is consistent with our other interpretation of vector subtraction as well: the vector $\vec{w}-\vec{u}$ is what we need to add to $\vec{u}$ to get $\vec{w}$. This is stated algebraically as $\vec{u}+\underline{\vec{w}-\vec{u}}=\vec{w}$.



(a) Vector addition

(b) Vector subtraction

Figure 9.3.7

In a similar way, we may geometrically represent a scalar multiple of a vector. For instance, if $\vec{v}=\langle 2,3\rangle$, then $2\vec{v}=\langle 4,6\rangle$. As shown in Figure 9.3.8, multiplying $\vec{v}$ by 2 leaves the direction unchanged, but stretches $\vec{v}$ by 2. Remember that $2\vec{v}=\vec{v}+\vec{v}$, which describes the total change of doing $\vec{v}$ then $\vec{v}$ again and thus gives a result that will be in same direction as $\vec{v}$. Also, $-2\vec{v}=\langle -4,-6\rangle$, which shows that multiplying by a negative scalar gives a vector pointing in the opposite direction of $\vec{v}$.

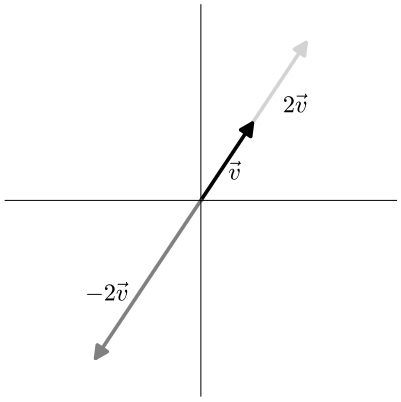
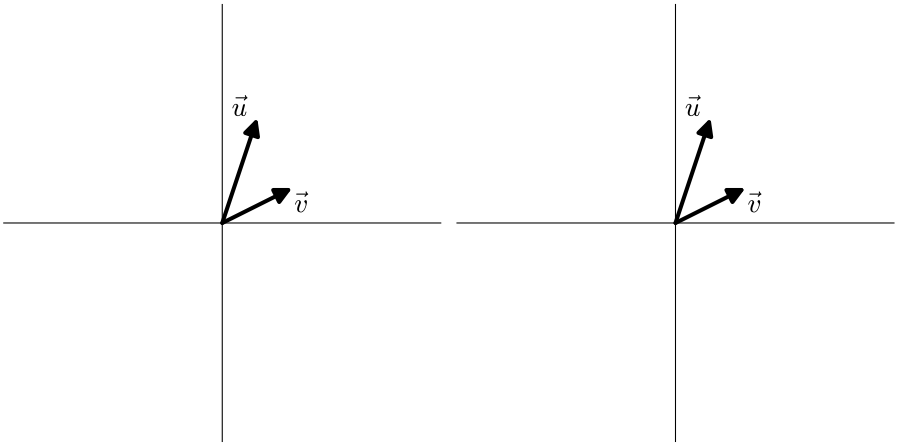


Figure 9.3.8 Scalar multiplication

Figure 9.3.8 motivates the following definition of parallel vectors: vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ are **parallel** provided that there exists a non-zero scalar $k \in \mathbb{R}$ such that $\vec{v} = k\vec{w}$. In other words, parallel vectors are (non-zero) scalar multiples of each other. There are a couple of things to note that are important about parallel vectors. First, the zero vector $\vec{0} = \langle 0, \ldots, 0 \rangle \in \mathbb{R}^n$ is ***not*** parallel to any other vector. Geometrically, this should make sense because the zero vector is the only vector that has no change and that will not correspond to any direction. Second, parallel vectors do not have to be in the same direction; in fact, parallel vectors can be in exactly opposite directions, as shown by $\vec{v}$ and $-2\vec{v}$ in Figure 9.3.8. Parallel vectors are important when trying to classify the orientation of lines, planes, and changes on surfaces.

Activity 9.3.5 Suppose that $\vec{u}$ and $\vec{v}$ are the vectors shown in Figure 9.3.9.



(a) Sketch sums **(b)** Sketch multiples

Figure 9.3.9

(a) On the axes in Figure 9.3.9(a), sketch the vectors $\vec{u} + \vec{v}$, $\vec{v} - \vec{u}$, $2\vec{u}$, $-2\vec{u}$, and $-3\vec{v}$.

(b) What are the components of $0\vec{v}$?

(c) On the axes in Figure 9.3.9(b), sketch the vectors $-3\vec{v}$, $-2\vec{v}$, $-1\vec{v}$, $2\vec{v}$, and $3\vec{v}$.

(d) Give a geometric description of the set of terminal points of the vectors $t\vec{v}$ where t is any scalar.

(e) On the set of axes in Figure 9.3.9(b), sketch the vectors $\vec{u} - 3\vec{v}$, $\vec{u} - 2\vec{v}$, $\vec{u} - \vec{v}$, $\vec{u} + \vec{v}$, and $\vec{u} + 2\vec{v}$.

(f) Give a geometric description of the set of terminal points of the vectors $\vec{u} + t\vec{v}$ where t is any scalar.

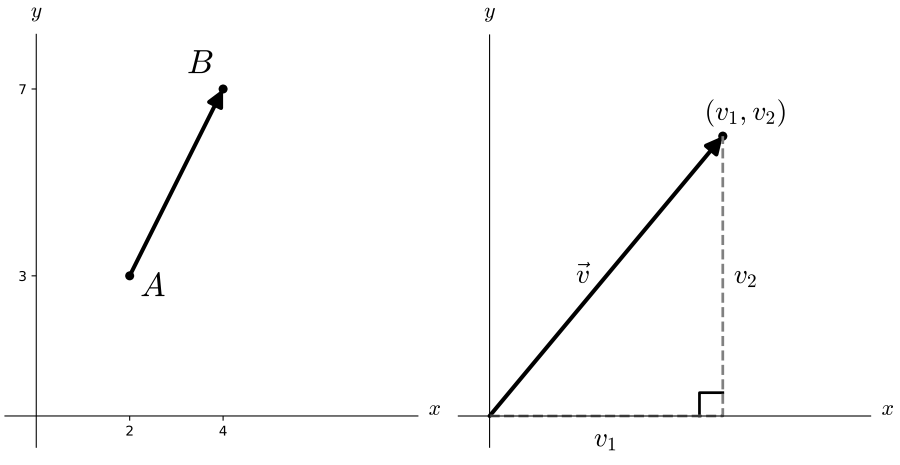
Before we move on to the magnitude of a vector, we will take a moment to point out a concept that permeates most the mathematics you will see for awhile. Our definitions of vector addition and scalar multiplication give us a way to take a few vectors and generate entire spaces with both familiar algebraic and geometric properties. We say that $a\vec{v}+b\vec{w}$ is a **linear combination** of the vectors $\vec{v}$ and $\vec{w}$ (where a and b are scalars). Earlier in this section we mentioned how every vector in $\mathbb{R}^3$ can be written as a linear combination of $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$. An important question in the next several math courses will be "Can you write any vector in $\mathbb{R}^n$ as a linear combination of $\vec{u}$, $\vec{v}$, and $\vec{w}$?" For now, we will point out that much of what has been done in this section is setting up the algebraic and geometric properties of linear combinations as a function of the vectors used.

9.3.7 The Magnitude of a Vector

By definition, vectors have both direction and magnitude (or length). We will now investigate how to calculate the magnitude of a vector. Since a vector $\vec{v}$ can be represented by a directed line segment, we can use the distance formula to calculate the length of the segment. This length is the **magnitude** of the vector $\vec{v}$ and is denoted $\|\vec{v}\|$.

Example 9.3.10 The magnitude of a vector is measured as the length of the directed line segments that represent the vector. For example, the directed line segment $\overrightarrow{AB}$ in Figure 9.3.11(a) will represent a vector with components of 2 and 4, which represent the horizontal and vertical changes from A to B. Notice that when we apply the distance formula to find the length of the segment from A to B, we compute $(x_B - x_A)$ and $(y_B - y_A)$, which are the components of the vector we are finding the magnitude of.

$$\left\|\overrightarrow{AB}\right\| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2} \qquad = \sqrt{(2)^2 + (4)^2}$$
$$= \sqrt{(v_1)^2 + (v_2)^2} \qquad = \sqrt{20}$$



(a) $\overrightarrow{AB}$ **(b)** An arbitrary vector, $\vec{v}$

Figure 9.3.11

Because the distance formula in $\mathbb{R}^n$ uses the sum of the squares of the change in each coordinate, we can see that given a vector $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle \in \mathbb{R}^n$ will have magnitude $\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$. Vectors of length one will be useful for describing directions without a focus on the magnitude. A **unit vector** is a vector whose magnitude is 1. Scalar multiplication gives us a very useful tool for describing a unit vector in a particular direction because we will need to stretch/shrink a given vector to length 1. Our vector given by $\overrightarrow{AB} = \langle 2, 4 \rangle$ will have magnitude $\sqrt{20}$, so the vector $\frac{1}{\sqrt{20}}\langle 2, 4 \rangle$ will have length 1 and will be in the same direction as $\overrightarrow{AB}$.

$$\left\| \left\langle \frac{2}{\sqrt{20}}, \frac{4}{\sqrt{20}} \right\rangle \right\| = \sqrt{\left(\frac{2}{\sqrt{20}}\right)^2 + \left(\frac{4}{\sqrt{20}}\right)^2} = \sqrt{\frac{4}{20} + \frac{16}{20}} = 1$$

▲

Activity 9.3.6

(a) Let $\vec{u} = \langle 2, 3 \rangle$ and $\vec{v} = \langle -1, 2 \rangle$. Find $\|\vec{u}\|$, $\|\vec{v}\|$, and $\|\vec{u} + \vec{v}\|$. Is it true that $\|\vec{u} + \vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$?

(b) Under what conditions will $\|\vec{w}_1 + \vec{w}_2\| = \|\vec{w}_1\| + \|\vec{w}_2\|$?

Hint. Think about how $\vec{w}_1$, $\vec{w}_2$, and $\vec{w}_1 + \vec{w}_2$ form the sides of a triangle.

(c) With the vector $\vec{u} = \langle 2, 3 \rangle$, find the lengths of $2\vec{u}$, $3\vec{u}$, and $-2\vec{u}$, respectively, and use proper notation to label your results.

(d) In general, if t is any scalar, how will $\|t\vec{w}\|$ be related to $\|\vec{w}\|$?

(e) Of the vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{\imath} + \hat{\jmath}$, which are unit vectors?

(f) Find a unit vector $\vec{v}$ whose direction is the same as $\vec{u} = \langle -2, 3 \rangle$.

(g) Find a unit vector $\vec{v}$ in the opposite direction to $\vec{u} = \langle -2, 3 \rangle$.

Summary

- A vector is an object that possesses the attributes of magnitude and direction. Examples of vector quantities are position, velocity, acceleration, and force.
- Two vectors are equal if they have the same direction and magnitude. Notice that position is not considered, so a vector is independent of its location.
- If $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ and $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ are two vectors in $\mathbb{R}^n$, then their vector sum is the vector

  $$\vec{u} + \vec{v} = \langle u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n \rangle.$$

  If $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ is a vector in $\mathbb{R}^n$ and c is a scalar, then the scalar multiple $c\vec{u}$ is the vector

  $$c\vec{u} = \langle cu_1, cu_2, \ldots, cu_n \rangle.$$

- The magnitude of the vector $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ in $\mathbb{R}^n$ is the scalar

  $$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

  A vector $\vec{u}$ is a unit vector provided that $\|\vec{u}\| = 1$. If $\vec{v}$ is a nonzero vector, then the vector $\frac{\vec{v}}{\|\vec{v}\|}$ is a unit vector with the same direction as $\vec{v}$.

9.3.8 Exercises

Note that some of the WeBWorK problems use the notation of putting vectors in bold font and not using the arrow over the name.

1. For each of the following, perform the indicated computation.
(a) $(3\,\mathbf{i}+8\,\mathbf{j}-10\,\mathbf{k})-(-10\,\mathbf{i}+4\,\mathbf{j}-3\,\mathbf{k}) =$ _____
(b) $(5\,\mathbf{i}+\mathbf{j}-\mathbf{k})-3(10\,\mathbf{i}-9\,\mathbf{j}+9\,\mathbf{k}) =$ _____

2. Find a vector **a** that has the same direction as $\langle -10, 7, 10\rangle$ but has length 4.
Answer: $\mathbf{a} =$ _____

3. Let $\mathbf{a} = \langle 5, 2\rangle$ and $\mathbf{b} = \langle 5, -3\rangle$.
Show that there are scalars s and t so that $s\mathbf{a} + t\mathbf{b} = \langle 0, -10\rangle$
You might want to sketch the vectors to get some intuition.
$s =$ _____
$t =$ _____

4. Let $\mathbf{a} =< 5, 4, -3 >$ and $\mathbf{b} =< -3, 2, 4 >$.
Show that there are scalars s and t so that $s\mathbf{a} + t\mathbf{b} =< -17, 4, 19 >$.
You might want to sketch the vectors to get some intuition.
$s =$_____
$t =$ _____

5. Resolve the following vectors into components:
(a) The vector $\vec{v}$ in 2-space of length 7 pointing up at an angle of $\pi/6$ measured from the positive x-axis.
$\vec{v} =$ _____ $\vec{i}$ + _____ $\vec{j}$
(b) The vector $\vec{w}$ in 3-space of length 5 lying in the yz-plane pointing upward at an angle of $3\pi/4$ measured from the positive y-axis.
$\vec{v} =$ _____ $\vec{i}$ + _____ $\vec{j}$ + _____ $\vec{k}$

6. Find all vectors $\vec{v}$ in 2 dimensions having $||\vec{v}|| = 5$ where the $\mathbf{i}$-component of $\vec{v}$ is $3\,\mathbf{i}$.
vectors: _____
(If you find more than one vector, enter them in a comma-separated list.)

7. Which is traveling faster, a car whose velocity vector is $23\vec{i} + 28\vec{j}$, or a car whose velocity vector is $30\vec{i}$, assuming that the units are the same for both directions?
(☐ the first car ☐ the second car) is the faster car.
At what speed is the faster car traveling?
speed = _____

8. Let $\mathbf{a} = \langle -1, 0, 5\rangle$ and $\mathbf{b} = \langle -5, 1, 4\rangle$.
Compute:
$\mathbf{a} + \mathbf{b} =$ (_____,_____,_____)
$\mathbf{a} - \mathbf{b} =$ (_____,_____,_____)
$2\mathbf{a} =$ (_____,_____,_____)
$3\mathbf{a} + 4\mathbf{b} =$ (_____,_____,_____)
$|\mathbf{a}| =$ _____

9. Find the length of the vectors
(a) $\mathbf{i}-4\,\mathbf{j}-2\,\mathbf{k}$: length = _____
(b) $2\,\mathbf{i}+0.6\,\mathbf{j}-\mathbf{k}$: length = _____

10. For each of the following, perform the indicated operations on the vectors $\vec{a} = 4\,\mathbf{j}+4\,\mathbf{k}$, $\vec{b} = 3\,\mathbf{i}+\mathbf{j}+4\,\mathbf{k}$, $\vec{z} = \mathbf{i}+\mathbf{j}$.

(a) $4\vec{a} + 2\vec{b} =$ _____

(b) $5\vec{a} + 3\vec{b} - 5\vec{z} =$ _____

11. Find the value(s) of a making $\vec{v} = 6a\,\vec{i} - 2\,\vec{j}$ parallel to $\vec{w} = a^2\,\vec{i} + 6\,\vec{j}$.

$a =$ _____

(If there is more than one value of a, enter the values as a comma-separated list.)

12. ***(a)*** Find a unit vector from the point $P = (3, 1)$ and toward the point $Q = (15, 6)$.

$\vec{u} =$ _____

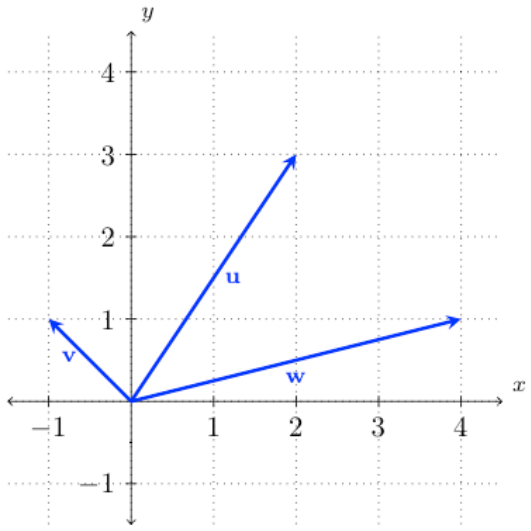
(b) Find a vector of length 39 pointing in the same direction.

$\vec{v} =$ _____

13. A truck is traveling due north at 75 km/hr approaching a crossroad. On a perpendicular road a police car is traveling west toward the intersection at 65 km/hr. Both vehicles will reach the crossroad in exactly one hour. Find the vector currently representing the displacement of the truck with respect to the police car.

displacement $\vec{d} =$ _____

14. Find the following expressions using the graph below of vectors **u**, **v**, and **w**.



1. $\mathbf{u} + \mathbf{v} =$ _____

2. $2\mathbf{u} + \mathbf{w} =$ _____

3. $3\mathbf{v} - 6\mathbf{w} =$ _____

4. $|\mathbf{w}| =$ _____

Note: *You can click on the graph to enlarge the image.*

15. Let $\mathbf{v} = \vec{PQ}$, where $P = (-3, -4), Q = (1, -8)$. Which of the vectors with the following given tails and heads are equivalent to **v**?

(a) (☐ Select ☐ Equivalent ☐ Not Equivalent) $(0, -2), (-4, -6)$

(b) (☐ Select ☐ Equivalent ☐ Not Equivalent) $(0, 0), (4, -4)$

(c) (☐ Select ☐ Equivalent ☐ Not Equivalent) $(5, -3), (9, -7)$

(d) (☐ Select ☐ Equivalent ☐ Not Equivalent) $(2, -2), (4, -6)$

16. Let $\mathbf{v} = \langle -2, 2, -3 \rangle$.

Find a unit vector in the same direction as **v**.

$\langle$ _____,_____,_____ $\rangle$

17. Which of the following vectors are parallel?

Please select the two vectors that are parallel.

☐ $< 24, -24, -12 >$

☐ $< -8, 2, 8 >$

☐ $< -12, -6, 12 >$

☐ $< -4, 2, 10 >$

☐ $< -12, -15, -6 >$

☐ $< -6, 6, 3 >$

Please select the two vectors that are parallel.

☐ $< 0, -32, -16 >$

☐ $< 3, -9, -12 >$

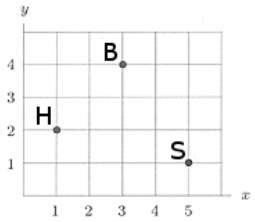
☐ $< 0, -8, -4 >$

☐ $< -20, -8, 0 >$

☐ $< 6, 10, 6 >$

☐ $< 0, 4, 8 >$

18. In the figure below the y-axis points north, the x-axis points east, and the xy-plane corresponds to the surface of the water. Suppose a boat is at point B, a submarine is 9 units below point S, and a helicopter is 12 units above point H.



Find the displacement vector and the distance from the submarine to the boat.

Displacement: _____

Distance: _____

Find the displacement vector and distance from the helicopter to the boat.

Displacement: _____

Distance: _____

Find the displacement vector and distance from the submarine to the helicopter.

Displacement: _____

Distance: _____

19. Let $\vec{v} = \langle 1, -2 \rangle$, $\vec{u} = \langle 0, 4 \rangle$, and $\vec{w} = \langle -5, 7 \rangle$.

(a) Determine the components of the vector $\vec{u} - \vec{v}$.

(b) Determine the components of the vector $2\vec{v} - 3\vec{u}$.

(c) Determine the components of the vector $\vec{v} + 2\vec{u} - 7\vec{w}$.

(d) Determine scalars a and b such that $a\vec{v} + b\vec{u} = \vec{w}$. Show all of your work in finding a and b.

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

(a) Determine the components and draw geometric representations of the vectors $2\vec{u}$, $\frac{1}{2}\vec{u}$, $(-1)\vec{u}$, and $(-3)\vec{u}$ on the same set of axes.

(b) Determine the components and draw geometric representations of the vectors $\vec{u} + \vec{v}$, $\vec{u} + 2\vec{v}$, and $\vec{u} + 3\vec{v}$ on the same set of axes.

(c) Determine the components and draw geometric representations of the vectors $\vec{u} - \vec{v}$, $\vec{u} - 2\vec{v}$, and $\vec{u} - 3\vec{v}$ on the same set of axes.

(d) Recall that $\vec{u} - \vec{v} = \vec{u} + (-1)\vec{v}$. Sketch the vectors $\vec{u}$, $\vec{v}$, $\vec{u} + \vec{v}$, and $\vec{u} - \vec{v}$ on the same set of axes. Use the "tip to tail" perspective for vector addition to explain the geometric relationship between $\vec{u}$, $\vec{v}$, and $\vec{u} - \vec{v}$.

21. Recall that given any vector $\vec{v}$, we can calculate its length, $\|\vec{v}\|$. Also, we say that two vectors that are (non-zero) scalar multiples of one another are **parallel**.

(a) Let $\vec{v} = \langle 3, 4\rangle$ in $\mathbb{R}^2$. Compute $\|\vec{v}\|$, and determine the components of the vector $\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v}$. What is the magnitude of the vector $\vec{u}$? How does its direction compare to $\vec{v}$?

(b) Let $\vec{w} = 3\hat{\imath} - 3\hat{\jmath}$ in $\mathbb{R}^2$. Determine a unit vector $\vec{u}$ in the same direction as $\vec{w}$.

(c) Let $\vec{v} = \langle 2, 3, 5\rangle$ in $\mathbb{R}^3$. Compute $\|\vec{v}\|$, and determine the components of the vector $\vec{u} = \frac{1}{\|\vec{v}\|}\vec{v}$. What is the magnitude of the vector $\vec{u}$? How does its direction compare to $\vec{v}$?

(d) Let $\vec{v}$ be an arbitrary nonzero vector in $\mathbb{R}^3$. Write a general formula for a unit vector that is parallel to $\vec{v}$.

22.

A force (like gravity) has both a magnitude and a direction. If two forces $\vec{u}$ and $\vec{v}$ are applied to an object at the same point, the resultant force on the object is the vector sum of the two forces. When a force is applied by a rope or a cable, we call that force *tension*. Vectors can be used to determine tension.

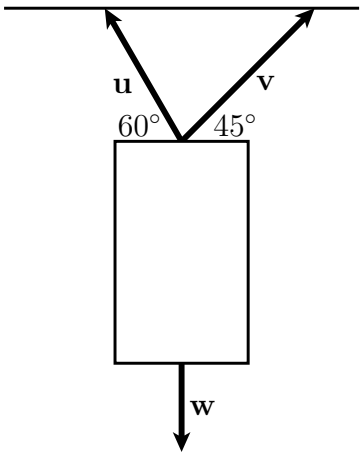


Figure 9.3.12 Forces acting on an object.

As an example, suppose a painting weighing 50 pounds is to be hung

from wires attached to the frame as illustrated in Figure 9.3.12. We need to know how much tension will be on the wires to know what kind of wire to use to hang the picture. Assume the wires are attached to the frame at point O. Let $\vec{u}$ be the vector emanating from point O to the left and $\vec{v}$ the vector emanating from point O to the right. Assume $\vec{u}$ makes a $60°$ angle with the horizontal at point O and $\vec{v}$ makes a $45°$ angle with the horizontal at point O. Our goal is to determine the vectors $\vec{u}$ and $\vec{v}$ in order to calculate their magnitudes.

(a) Treat point O as the origin. Use trigonometry to find the components u_1 and u_2 so that $\vec{u} = u_1\hat{\imath} + u_2\hat{\jmath}$. Since we don't know the magnitude of $\vec{u}$, your components will be in terms of $\|\vec{u}\|$ and the cosine and sine of some angle. Then find the components v_1 and v_2 so that $\vec{v} = v_1\hat{\imath} + v_2\hat{\jmath}$. Again, your components will be in terms of $\|\vec{v}\|$ and the cosine and sine of some angle.

(b) The total force holding the picture up is given by $\vec{u} + \vec{v}$. The force acting to pull the picture down is given by the weight of the picture. Find the force vector $\vec{w}$ acting to pull the picture down.

(c) The picture will hang in equilibrium when the force acting to hold it up is equal in magnitude and opposite in direction to the force acting to pull it down. Equate these forces to find the components of the vectors $\vec{u}$ and $\vec{v}$.

9.4 The Dot Product

Motivating Questions

- How is the dot product of two vectors defined and what geometric information does it tell us?
- How do we measure the angle between two vectors?
- How can we tell if two vectors in $\mathbb{R}^n$ are perpendicular?
- How can we find how much of one vector is parallel to another?

9.4.1 Introduction

In Section 9.3, we considered vector addition and scalar multiplication algebraically and interpreted the corresponding geometric meaning. In this section, we will introduce the first of two types of vector multiplication that we will study. This multiplication produces a scalar that tells us a remarkable amount about the geometric relationship between the vectors that are multiplied. We will start investigating these properties in the Preview Activity, but first, we provide the definition of this form of multiplication for two-dimensional vectors.

Consider the two-dimensional vectors $\vec{u} = \langle u_1, u_2\rangle$ and $\vec{v} = \langle v_1, v_2\rangle$. The **dot product** of $\vec{u}$ and $\vec{v}$ is denoted $\vec{u} \cdot \vec{v}$ and defined to be

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2.$$

For instance, we have $\langle 2, -1\rangle \cdot \langle 1, 4\rangle = 2 \cdot 1 + (-1) \cdot 4 = -2$. At first glance, this may feel like we've *lost* information, as we have taken two objects with both magnitude and direction, multiplied them, and obtained a scalar. However,

this scalar allows us to draw many useful conclusions, as you will now begin exploring.

Preview Activity 9.4.1 In this activity, we will use the following vectors:

$$\vec{u}_1 = \langle 1, 2\rangle, \vec{u}_2 = \langle -1, 1\rangle, \vec{u}_3 = \langle 2, -1\rangle, \vec{u}_4 = \langle 1, -2\rangle$$

(a) Draw representatives of each of $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ in standard position using the axes below.

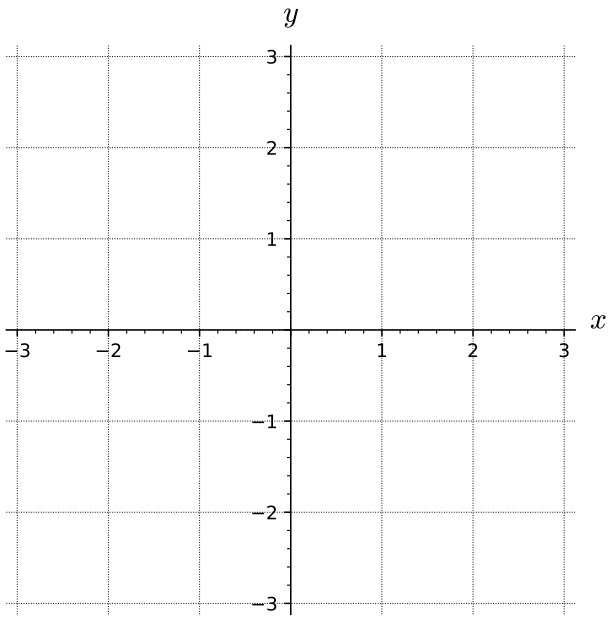


Figure 9.4.1 A plot of the 2D plane

(b) Compute each of the dot products listed in the table below. Fill in only the "Value" column at this point. The "Angle Classification" column will be completed in the next part.

Dot Product	Value	Angle Classification
$\vec{u}_1 \cdot \vec{u}_2$		
$\vec{u}_1 \cdot \vec{u}_3$		
$\vec{u}_1 \cdot \vec{u}_4$		
$\vec{u}_2 \cdot \vec{u}_4$		
$\vec{u}_3 \cdot \vec{u}_4$.		

(c) When we look at vectors drawn with the same initial point, as you have done with the vectors in Figure 9.4.1, we can consider the angle between two vectors to be the *smaller* angle formed by looking at them. This results in an angle between 0 and π if we measure in radians or between 0° and 180° if we measure in degrees. Rather than measuring the angles between the vectors you've drawn in Figure 9.4.1 precisely, classify the angle between the pairs of vectors in the table in part (b) by writing acute, right, or obtuse in the "Angle Classification" column.

(d) Look at the values of the dot products that you computed in part (b) and the classifications of the angles between the vectors that you made in part (c). Write a sentence or two to describe what you notice about the relationship between the *sign* of the dot product and the type of angle between the vectors.

(e) Does the value of the dot product or the classification of the angle change if we change the order in which we consider the vectors?

(f) Do you think that knowing only the value of the dot product $\vec{v} \cdot \vec{w}$ would be enough to determine the exact value of the angle between $\vec{v}$ and $\vec{w}$ if you didn't know the components of $\vec{v}$ and $\vec{w}$? Write a sentence to explain your reasoning.

Hint. Compare the pair of vectors $\vec{u}_1$ and $\vec{u}_4$ to $\vec{u}_3$ and $\vec{u}_4$.

(g) Although the angle between a vector and itself is 0, the dot product of a vector with itself can be computed. Compute $\vec{u}_1 \cdot \vec{u}_1$ and $\vec{u}_2 \cdot \vec{u}_2$ as well as the magnitude of $\vec{u}_1$ and $\vec{u}_2$. What do you notice about the relationship between these values? Write a sentence explaining why this seems reasonable based on the way the dot product of a vector with itself and the magnitude of that vector are computed.

In Preview Activity 9.4.1, you have started developing an understanding of how the dot product gives us information about the angle between two vectors in the plane. The rest of this section will formalize this idea about the dot product of two vectors and explore properties of this measurement.

9.4.2 The Dot Product

The idea of multiplying corresponding components of vectors and summing them can be extended to any dimension, which gives us the definition of the dot product.

Definition 9.4.2 The **dot product** of vectors $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ and $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$ in $\mathbb{R}^n$ is the scalar

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

◆

As we will see shortly, the dot product arises in physics to calculate the work done by a vector force in a given direction. It might be more natural to define the dot product in this context. However, it is more convenient from a mathematical perspective to define the dot product algebraically and then view work as an application of this definition.

To illustrate the idea of computing a dot product in more than three dimensions, we find that

$$\begin{aligned} \langle 3, 0, 1, 5 \rangle \cdot \langle -2, 1, 4, -2 \rangle &= 3 \cdot (-2) + 0 \cdot 1 + 1 \cdot 4 + 5 \cdot (-2) \\ &= -6 + 0 + 4 - 10 = -12. \end{aligned}$$

Notice that the resulting quantity is a *scalar*. In the next section, we will explore another product of vectors for which the resulting quantity is a *vector*. For this reason, you will sometimes see the phrase **scalar product** used in place of "dot product".

The dot product gives us a way to define a product of two vectors, and we will soon formalize the ideas from Preview Activity 9.4.1 about what the dot product tells us about the angle between vectors. First, however, we state some useful properties the dot product that show how it behaves in ways that are similar to the product of real numbers. In fact, all of these properties are consequences of properties of addition and multiplication of real numbers.

Properties of the dot product.

Let $\vec{u}$, $\vec{v}$, and $\vec{w}$ be vectors in $\mathbb{R}^n$. Then

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (The dot product is *commutative*.)
2. $(\vec{u} + \vec{v}) \cdot \vec{w} = (\vec{u} \cdot \vec{w}) + (\vec{v} \cdot \vec{w})$ (The dot product can be *distributed* across vector addition.)
3. if c is a scalar, then $(c\vec{u}) \cdot \vec{w} = c(\vec{u} \cdot \vec{w})$ (The dot product is *associative* with scalar multiplication.)

In Preview Activity 9.4.1, you observed that for two-dimensional vectors, the dot product of a vector with itself is the square of the vector's magnitude. This is so useful, that we state it below.

Key Idea 9.4.3 *If $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ is a vector in $\mathbb{R}^n$, then $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$. In other words, the dot product of a vector with itself gives the square of the length of the vector.*

Proof. We can do this computation directly and find

$$\vec{u} \cdot \vec{u} = \langle u_1, u_2, \ldots, u_n \rangle \cdot \langle u_1, u_2, \ldots, u_n \rangle = u_1^2 + u_2^2 + \ldots + u_n^2 = \|\vec{u}\|^2 .$$

■

The next activity has you practice computing dot products and applying properties of them.

Activity 9.4.2

(a) Determine each of the following.

(i) $\langle 1, 2, -3 \rangle \cdot \langle 4, -2, 0 \rangle$

(ii) $\langle 0, 3, -2, 1 \rangle \cdot \langle 5, -6, 0, 4 \rangle$

(b) Let $\vec{u}, \vec{v}, \vec{w}$ be vectors in $\mathbb{R}^n$. Suppose that you know that $\vec{u} \cdot \vec{w} = 10$ and $\vec{v} \cdot \vec{w} = -3$. Compute $(\vec{u} + \vec{v}) \cdot \vec{w}$.

(c) Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$. Suppose that you know that $\vec{u} \cdot \vec{v} = 3$, $\|\vec{u}\| = 4$, and $\|\vec{v}\| = 7$. Compute $(\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v})$.

(d) Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$. For what value of t is $(t\vec{u} + \vec{v}) \cdot \vec{v} = 0$?

9.4.3 The Angle between Vectors

As we saw in Preview Activity 9.4.1, the sign of the dot product gives us information about the type of angle formed by two vectors. We also saw that the dot product alone does not give us sufficient information to measure the angle. However, with a bit more information about the vectors, we are able to measure the angle, as we will soon see. If we are given two vectors $\vec{u}$ and $\vec{v}$, there are two angles that these vectors create, as depicted in Figure 9.4.4. We will call θ, the smaller of these angles, the **angle between these vectors**. Notice that θ lies between 0 and π because we would consider the angle from the other side if the angle was greater than π.

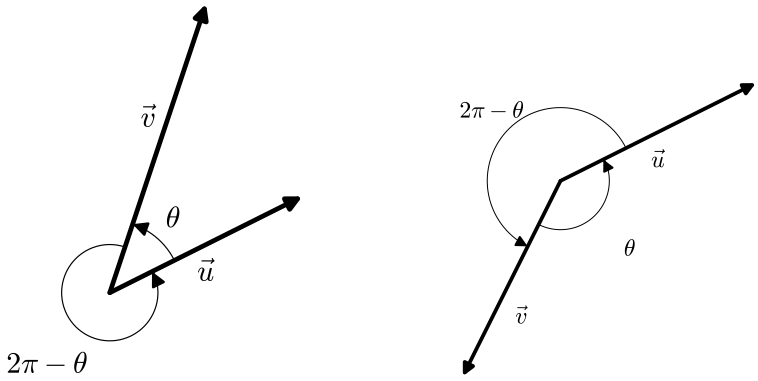


Figure 9.4.4 The angle between vectors for two different configurations of $\vec{u}$ and $\vec{v}$

Key Idea 9.4.5 *Given vectors $\vec{u}$ and $\vec{v}$, let θ be the smallest angle between $\vec{u}$ and $\vec{v}$. We have the following relationship:*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \tag{9.4.1}$$

Solving for the angle θ, we have:

$$\theta = \arccos\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|} \right) \tag{9.4.2}$$

Proof. Recall that the Law of Cosines says that if we consider any triangle with side lengths a, b, c and θ the measure of the angle formed by the sides of length a and b, then $c^2 = a^2 + b^2 - 2ab\cos(\theta)$. To determine the angle θ here, we may apply the Law of Cosines to the triangle shown in Figure 9.4.6.

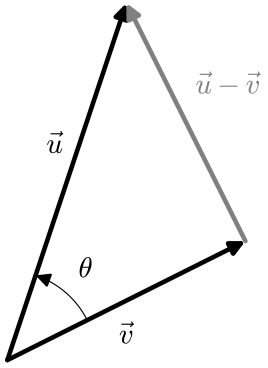


Figure 9.4.6 A triangle of vectors, $\vec{u}$, $\vec{v}$, and $\vec{u} - \vec{v}$

Using the fact that the dot product of a vector with itself gives us the square of its length, together with other properties of the dot product, we find:

$$\begin{aligned}
\|\vec{u} - \vec{v}\|^2 &= \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \\
(\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v}) &= \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{v} - 2\|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \\
\vec{u} \cdot (\vec{u} - \vec{v}) - \vec{v} \cdot (\vec{u} - \vec{v}) &= \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{v} - 2\|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \\
\vec{u} \cdot \vec{u} - 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v} &= \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{v} - 2\|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \\
-2\vec{u} \cdot \vec{v} &= -2\|\vec{u}\| \, \|\vec{v}\| \cos(\theta) \\
\vec{u} \cdot \vec{v} &= \|\vec{u}\| \, \|\vec{v}\| \cos(\theta)
\end{aligned}$$

■

The powerful conclusion from (9.4.1) is that the dot product is a simple algebraic calculation that provides us with important geometric information about the angle between the vectors that would be difficult to determine otherwise. While you normally need to consider if the angle that comes out of an inverse trigonometric function needs to be adjusted based on the context of a problem, we do not need to worry about that with (9.4.2) because the angle measurement and the range of arccosine are both from 0 to π.

Activity 9.4.3 Determine each of the following:

(a) The length of the vector $\vec{u} = \langle 1, 2, -3 \rangle$ using the dot product.

(b) The angle between the vectors $\vec{u} = \langle 1, 2 \rangle$ and $\vec{v} = \langle 4, -1 \rangle$ to the nearest tenth of a degree.

(c) The angle between the vectors $\vec{y} = \langle 1, 2, -3 \rangle$ and $\vec{z} = \langle -2, 1, 1 \rangle$ to the nearest tenth of a degree.

(d) If the angle between the vectors $\vec{u}$ and $\vec{v}$ is a right angle, what does the expression $\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta)$ say about their dot product?

(e) If the angle between the vectors $\vec{u}$ and $\vec{v}$ is acute—that is, less than $\pi/2$—what does the expression $\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta)$ say about their dot product?

(f) If the angle between the vectors $\vec{u}$ and $\vec{v}$ is obtuse—that is, greater than $\pi/2$—what does the expression $\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos(\theta)$ say about their dot product?

The last three parts of Activity 9.4.3 allow us to succinctly state the observations you arrived at in the Preview Activity. The next subsection will state these conclusions explicitly.

Many important relationships between vectors arise when the angle between two vectors is a right angle. In this case, we say the vectors are **orthogonal**. For instance, orthogonality often plays a role in optimization problems: to determine the shortest path from a point in $\mathbb{R}^3$ to a given plane, we move along a line orthogonal to the plane. From your prior mathematical experience, you are likely familiar with saying that two objects that form a right angle are *perpendicular*. When vectors are involved, we prefer to use the term *orthogonal* because it allows for one (or both) of the vectors to be $\vec{0}$, which is a case where we would not say the vectors are perpendicular.

As Activity 9.4.3 indicates, the dot product provides a simple means to determine whether two vectors are orthogonal to one another. In this case, $\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos(\pi/2) = 0$, so we make the following important observation.

The dot product and orthogonality.

Two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ are orthogonal to each other provided that $\vec{u} \cdot \vec{v} = 0$.

More generally, the sign of the dot product gives us useful information about the relative orientation of the vectors. Since the magnitudes in equation (9.4.1) are positive, recalling that

$$\begin{aligned} \cos(\theta) > 0 \quad & \text{if } \theta \text{ is an acute angle,} \\ \cos(\theta) = 0 \quad & \text{if } \theta \text{ is a right angle, and} \\ \cos(\theta) < 0 \quad & \text{if } \theta \text{ is an obtuse angle,} \end{aligned}$$

then we see that for nonzero vectors $\vec{u}$ and $\vec{v}$,

$\vec{u} \cdot \vec{v} > 0$	if θ is an acute angle,
$\vec{u} \cdot \vec{v} = 0$	if θ is a right angle, and
$\vec{u} \cdot \vec{v} < 0$	if θ is an obtuse angle.

This is illustrated in Figure 9.4.7.

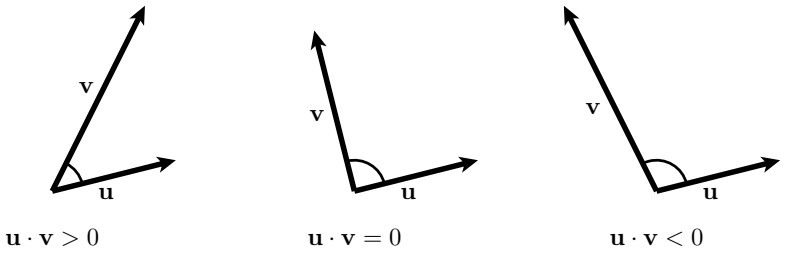


Figure 9.4.7 The orientation of vectors

These different orientations of vectors will lead us to seeking to understand how to construct a vector that represents the portion of one vector that is in the direction of (or parallel to) another vector. We examine this question in depth in the next subsection.

Activity 9.4.4 Return to Preview Activity 9.4.1 and use the results of this subsection to determine which pairs of vectors are orthogonal. Also classify the angles between non-orthogonal pairs of vectors as acute or obtuse.

9.4.4 Projections

Suppose we want to split the vector $\vec{u}$ into parts that are parallel to $\vec{v}$ and orthogonal to $\vec{v}$, which we will call $\vec{w}_1$ and $\vec{w}_2$ respectively. In other words, we want to write $\vec{u} = \vec{w}_1 + \vec{w}_2$ where $\vec{w}_1 = k\vec{v}$ and $\vec{w}_2 \cdot \vec{v} = 0$. Geometrically, this will look like one of the possibilities shown in Figure 9.4.8.

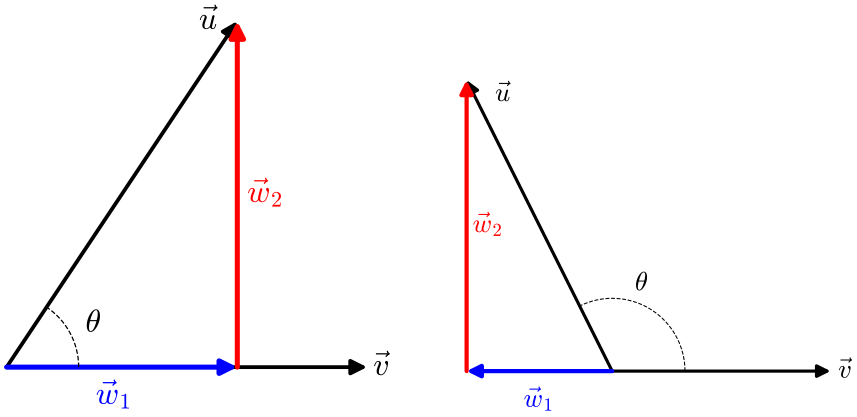


Figure 9.4.8 Different possibilities for splitting $\vec{u}$ into parts that are parallel and orthogonal to $\vec{v}$

Notice that if the angle between $\vec{u}$ and $\vec{v}$ is obtuse, then the amount of $\vec{u}$ that is parallel to $\vec{v}$ will be in the direction opposite $\vec{v}$. Before we look at how to calculate the vectors $\vec{w}_1$ and $\vec{w}_2$, we will do an activity about drawing the

graphical representation of this splitting of a vector into parts that are parallel to and orthogonal to another vector.

Activity 9.4.5 In Figure 9.4.8, you can see illustrations of splitting $\vec{u}$ into two parts: $\vec{w}_1$, which is parallel to $\vec{v}$, and $\vec{w}_2$, which is orthogonal to $\vec{v}$. Use this figure for reference as you do the following.

(a) We know from the previous subsection that there is a third configuration of vectors, which occurs when $\vec{u}$ and $\vec{v}$ are orthogonal. Suppose that $\vec{u}$ and $\vec{v}$ are nonzero orthogonal vectors. What would $\vec{w}_1$ and $\vec{w}_2$ be in this case?

(b) We want to switch the roles of $\vec{u}$ and $\vec{v}$ for the examples in the previous parts. Specifically, for these configuration of vectors, we want to split $\vec{v}$ into parts that are parallel to $\vec{u}$, which we will call $\vec{z}_1$, and orthogonal to $\vec{u}$, which we will call $\vec{z}_2$. On Figure 9.4.9, draw $\vec{z}_1$ and $\vec{z}_2$ for each configuration.

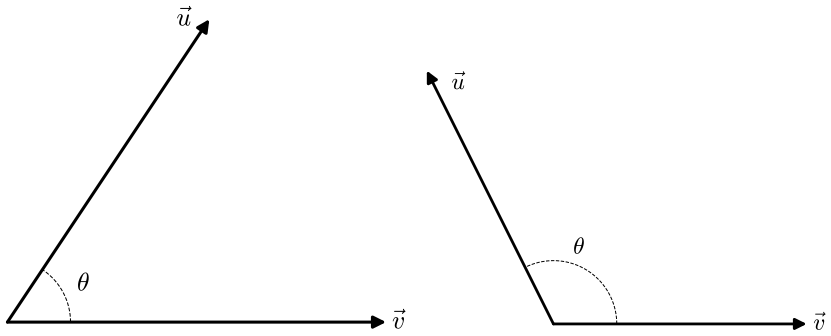


Figure 9.4.9 A plot of $\vec{u}$ and $\vec{v}$

Based on your drawing, is it the case that $\vec{w}_1 = \vec{z}_1$? What about $\vec{w}_2$ and $\vec{z}_2$?

Hint. You may find it useful to rotate your paper so that $\vec{u}$ appears horizontally.

Now that we have visually explored the relationship between $\vec{u}$, $\vec{v}$, $\vec{w}_1$, and $\vec{w}_2$, we will see how we can find formulas that allow us to compute $\vec{w}_1$ and $\vec{w}_2$ in terms of $\vec{u}$ and $\vec{v}$.

Activity 9.4.6 Suppose that $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$. We wish to find $\vec{w}_1$ and $\vec{w}_2$ so that $\vec{w}_1$ is parallel to $\vec{v}$ and $\vec{w}_2$ is orthogonal to $\vec{v}$, as shown in Figure 9.4.8.

(a) Explain why $\vec{u} = \vec{w}_1 + \vec{w}_2$.

(b) Compute $\vec{u} \cdot \vec{v}$ as $(\vec{w}_1 + \vec{w}_2) \cdot \vec{v}$. Simplify your answer as much as possible, using the fact that $\vec{w}_2$ is orthogonal to $\vec{v}$.

(c) Since $\vec{w}_1$ is parallel to $\vec{v}$, there is a scalar k so that $\vec{w}_1 = k\vec{v}$. Substitute $k\vec{v}$ for $\vec{w}_1$ in your answer to the previous part and then solve for k.

(d) Give a formula for $\vec{w}_1$ purely in terms of $\vec{u}$ and $\vec{v}$.

Hint. Your answer should be a scalar multiple of $\vec{v}$ and include two dot products and a fraction.

The vector $\vec{w}_1$ we have been considering in this subsection is so useful that it has specific notation. We use the previous activity to state the following

definition.

Definition 9.4.10 Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^n$. The **projection of $\vec{u}$ onto $\vec{v}$** is the vector

$$\text{proj}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v} \tag{9.4.3}$$

$$= \left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|}\right)\frac{\vec{v}}{\|\vec{v}\|}. \tag{9.4.4}$$

◆

Using the property that the dot prodcut of a vector with itself is the square of the vector's magnitude, we can write $\text{proj}_{\vec{v}}\vec{u} = \left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|}\right)\frac{\vec{v}}{\|\vec{v}\|}$. Since $\vec{v}/\|\vec{v}\|$ is a unit vector, we can see from this form that the magnitude of $\text{proj}_{\vec{v}}\vec{u}$ is $(\vec{u}\cdot\vec{v})/\|\vec{v}\|$. We call this scalar the **component of $\vec{u}$ in the direction of $\vec{v}$** and denote it by

$$\text{comp}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|}.$$

Recall from earlier in the section that $\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos(\theta)$, where θ is the angle between $\vec{u}$ and $\vec{v}$, we can also write

$$\text{comp}_{\vec{v}}\vec{u} = \frac{\|\vec{u}\|\,\|\vec{v}\|\cos(\theta)}{\|\vec{v}\|} = \|\vec{u}\|\cos(\theta).$$

Vector subtraction also gives us a way to solve for the piece $\vec{w}_2$ of $\vec{u}$ that is orthogonal to $\vec{v}$. We find $\vec{w}_2 = \vec{u} - \vec{w}_1 = \vec{u} - \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$. The **orthogonal projection of $\vec{u}$ onto $\vec{v}$** is

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

These relationships (for one configuration of $\vec{u}$ and $\vec{v}$) are shown in Figure 9.4.11. *Remember that the length of $proj_{\vec{v}}\vec{u}$ is $comp_{\vec{v}}\vec{u}$.* Also notice how the arrangement of the vectors provides an illustration of the fact that $\vec{u} = \text{proj}_{\vec{v}}\vec{u} + \text{proj}_{\perp\vec{v}}\vec{u}$.

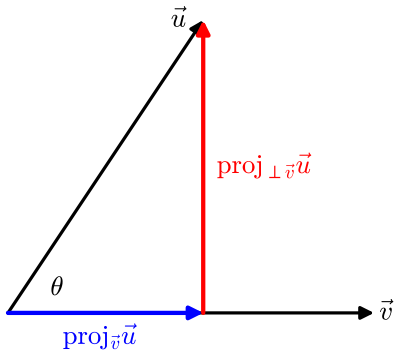


Figure 9.4.11 The relationship between $\vec{u}$, $\vec{v}$, the projection of $\vec{u}$ onto $\vec{v}$ and the orthogonal projection of $\vec{u}$ onto $\vec{v}$.

Notice in Figure 9.4.11 that you can also see

$$\cos(\theta) = \frac{\|\text{proj}_{\vec{v}}\vec{u}\|}{\|\vec{u}\|} = \frac{\text{comp}_{\vec{v}}\vec{u}}{\|\vec{u}\|},$$

which aligns with our earlier formula $\text{comp}_{\vec{v}}\vec{u} = \|\vec{u}\|\cos(\theta)$.

The component of $\vec{u}$ in the direction of $\vec{v}$ will be a scalar measure of how much of $\vec{u}$ is parallel to $\vec{v}$. In fact, this shows up in the second formula for $\text{proj}_{\vec{v}}\vec{u}$. Specifically, we can write

$$\text{proj}_{\vec{v}}\vec{u} = (\text{comp}_{\vec{v}}\vec{u})\frac{\vec{v}}{\|\vec{v}\|},$$

which separates the projection of $\vec{u}$ onto $\vec{v}$ into how much of $\vec{u}$ is parallel to $\vec{v}$ (a scalar) times the unit vector in the direction of $\vec{v}$. This useful idea of separating the magnitude of a vector and the direction of a vector (as described by a unit vector) will be used in the analysis of many vector calculations. The first expression for $\text{proj}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$ can be interpreted conceptually as saying the projection of $\vec{u}$ onto $\vec{v}$ must be parallel to $\vec{v}$ and $\frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}$ is closely related to the length of $\text{proj}_{\vec{v}}\vec{u}$.

Example 9.4.12 In this example we will look at calculating and making sense of the different parts of the projection of $\langle 4,3\rangle$ onto $\langle 0,2\rangle$. For notational purposes, we will let $\vec{u} = \langle 4,3\rangle$ and $\vec{v} = \langle 0,2\rangle$. Applying Definition 9.4.10, we find

$$\begin{aligned}\text{proj}_{\vec{v}}\vec{u} &= \left(\frac{\langle 4,3\rangle\cdot\langle 0,2\rangle}{\langle 0,2\rangle\cdot\langle 0,2\rangle}\right)\langle 0,2\rangle \\ &= \left(\frac{6}{4}\right)\langle 0,2\rangle \\ &= \left(\frac{3}{2}\right)\langle 0,2\rangle = \langle 0,3\rangle.\end{aligned}$$

Using the other formula for the projection, we have

$$\begin{aligned}\text{proj}_{\vec{v}}\vec{u} &= \left(\frac{\langle 4,3\rangle\cdot\langle 0,2\rangle}{\|\langle 0,2\rangle\|}\right)\frac{\langle 0,2\rangle}{\|\langle 0,2\rangle\|} \\ &= \left(\frac{6}{2}\right)\frac{\langle 0,2\rangle}{2} \\ &= (3)\langle 0,1\rangle = \langle 0,3\rangle.\end{aligned}$$

Note that the scalar (3) in the last equation is $\text{comp}_{\vec{v}}\vec{u}$ and the vector $\langle 0,1\rangle$ is the unit vector in the direction of $\vec{v}$.

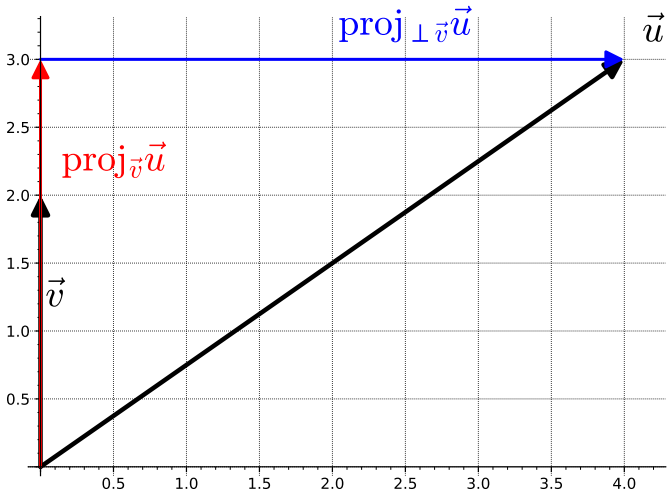


Figure 9.4.13 A plot of $\vec{u}$, $\vec{v}$, $\text{proj}_{\vec{v}}\vec{a}$, and $\text{proj}_{\perp\vec{v}}\vec{u}$

We can now compute the orthogonal projection of $\vec{u}$ onto $\vec{v}$ using $\text{proj}_{\vec{v}}\vec{u}$.

$$\text{proj}_{\perp\vec{v}}\vec{u} = \vec{u} - \text{proj}_{\vec{v}}\vec{u} = \langle 4, 3\rangle - \langle 0, 3\rangle = \langle 4, 0\rangle.$$

We can summarize these calculations as follows:

- The projection of $\vec{u}$ onto $\vec{v}$ is asking how much of $\langle 4, 3\rangle$ is in the vertical direction.
- The projection of $\vec{u}$ onto $\vec{v}$ can be split into the scalar 3, the component of $\vec{u}$ in the direction of $\vec{v}$, and the unit vector in the direction of $\vec{v}$, $\langle 0, 1\rangle$.
- The orthogonal projection of $\vec{u}$ onto $\vec{v}$ will be given by how much of $\vec{u}$ is horizontal because we subtract out all of the vertical component of $\vec{u}$.

▲

Going forward, we will use the projection more often than the dot product by itself. This is because the projection precisely measures how much of one vector is in a particular direction rather than the much vauger measurement of the type of angle between the vectors.

The following activity gives you an opportunity to practice the computations involved with finding parallel and orthogonal projections.

Activity 9.4.7 Let $\vec{u} = \langle 2, 6\rangle$.

(a) Let $\vec{v} = \langle 4, -8\rangle$. Find $\text{proj}_{\vec{v}}\vec{u}$, $\text{comp}_{\vec{v}}\vec{u}$, and $\text{proj}_{\perp\vec{v}}\vec{u}$. Draw a picture to illustrate the vectors involved. Finally, express $\vec{u}$ as the sum of two vectors where one is parallel to $\vec{v}$ and the other is perpendicular to $\vec{v}$.

(b) Now let $\vec{w} = \langle -2, 4\rangle$. Add $\vec{w}$ to the picture you drew in the previous part. Without doing any calculations, find $\text{proj}_{\vec{w}}\vec{u}$. Explain your reasoning.

(c) Find a vector $\vec{w}$ not parallel to $\vec{z} = \langle 3, 4\rangle$ such that $\text{proj}_{\vec{z}}\vec{w}$ has length 10. Note that there are infinitely many different answers!

It is important to note that in general $\text{proj}_{\vec{v}}\vec{w} \neq \text{proj}_{\vec{w}}\vec{v}$. You can explore in Exercise 9.4.6.19 how switching the vectors in the projection formula will give a different results.

9.4.5 Work, Force, and Displacement

In physics, work is a measure of the energy required to apply a force to an object through a displacement. In Figure 9.4.14, we can see a diagram showing the force applied to a refrigerator (in orange) that is used to move the refrigerator from point A to point B. The change in location, $\overrightarrow{AB}$, is called the **displacement**. Notice that not all of the force is in the direction of the displacement. The work done on the refrigerator only depends on the amount of force that is in the direction of $\overrightarrow{AB}$.

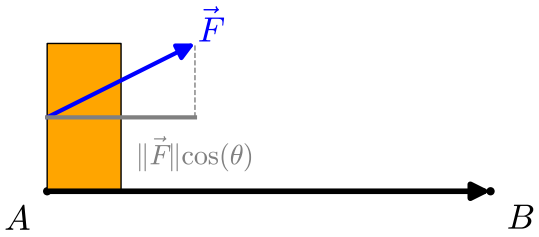


Figure 9.4.14 A diagram showing the force applied to the refrigerator as the refrigerator is moved from A to B

The work required to move the refrigerator is

$$W = \vec{F} \cdot \overrightarrow{AB} = \left\|\vec{F}\right\| \left\|\overrightarrow{AB}\right\| \cos(\theta).$$

This means that the work is determined only by the magnitude of the force applied parallel to the displacement and the length of the displacement. In fact, work is closely related to the projection of the force vector onto the displacement vector. (More precisely, it is the product of the component of the force vector in the direction of the displacement vector times the magnitude of the displacement vector.)

The units commonly used for work are Joules or Newton-meters in the metric system and foot-pounds in customary units. (In electrical settings, units such as coulomb-volts or kilowatt-hours may be used.) Note that these units come from work being the dot product of a force vector, using units of Newtons or pounds, and a displacement vector, using units of meters or feet.

Activity 9.4.8

(a) Determine the work done by a 25 pound force acting at a 30° angle to the direction of the object's motion, if the object is pulled 10 feet.

(b) Determine if more work or less work is done if the angle to the direction of the object's motion is 60°.

Summary

- The dot product of two vectors in $\mathbb{R}^n$, $\vec{u} = \langle u_1, u_2, \ldots, u_n \rangle$ and $\vec{v} = \langle v_1, v_2, \ldots, v_n \rangle$, is the scalar

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

  The dot product is related to the length of a vector since $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$.

- The dot product provides us with information about the angle between two vectors since

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

  where θ is the angle between $\vec{u}$ and $\vec{v}$.

- Two vectors are orthogonal if the angle between them is $\pi/2$. In terms of the dot product, the vectors $\vec{u}$ and $\vec{v}$ are orthogonal if and only if $\vec{u} \cdot \vec{v} = 0$.

- The projection of a vector $\vec{u}$ in $\mathbb{R}^n$ onto a vector $\vec{v}$ in $\mathbb{R}^n$ is the vector

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

9.4.6 Exercises

1. Find $\mathbf{a} \cdot \mathbf{b}$ if

$\mathbf{a} = \langle 3, -2, 5 \rangle$ and $\mathbf{b} = \langle 3, 0, 3 \rangle$

$\mathbf{a} \cdot \mathbf{b} =$_____

Is the angle between the vectors "acute", "obtuse" or "right"?

3. Determine if the pairs of vectors below are "parallel", "orthogonal", or "neither".

$\mathbf{a} = \langle 4, 5, -2 \rangle$ and $\mathbf{b} = \langle -12, -15, -123/2 \rangle$ are

$\mathbf{a} = \langle 4, 5, -2 \rangle$ and $\mathbf{b} = \langle -12, -15, 6 \rangle$ are

$\mathbf{a} = \langle 4, 5, -2 \rangle$ and $\mathbf{b} = \langle 12, 15, -6 \rangle$ are

4. Find $\mathbf{a} \cdot \mathbf{b}$ if $|\mathbf{a}| = 10$, $|\mathbf{b}| = 6$, and the angle between $\mathbf{a}$ and $\mathbf{b}$ is $\frac{\pi}{6}$ radians.

$\mathbf{a} \cdot \mathbf{b} =$ _____

5. Find the angle θ between the vectors $\mathbf{a} = 5\mathbf{i} - \mathbf{j} - 4\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

Answer (in radians): $\theta =$ _____

6. Find $\mathbf{a} \cdot \mathbf{b}$ if $|\mathbf{a}| = 3$, $|\mathbf{b}| = 9$, and the angle between $\mathbf{a}$ and $\mathbf{b}$ is $\frac{\pi}{2}$ radians.

$\mathbf{a} \cdot \mathbf{b} =$ _____

7. A rectangular box has a base 8 inches long and 12 inches wide, and the height of the box is 2 inches.

Find the angle between the diagonal of the box and the diagonal of its base. The angle should be measured in radians.

angle = _____

8. Find the measures of the three angles, in radians, of the triangle with the given vertices: $D(1, 1, 1)$, $E(1, -3, 2)$, and $F(-3, 2, 6)$.

1. $\angle D =$ _____

2. $\angle E =$ _____

3. $\angle F =$ _____

9. A constant force $\mathbf{F} = 2\mathbf{i} + 7\mathbf{j} + 7\mathbf{k}$ moves an object along a straight line from point $(5, -2, -4)$ to point $(2, -8, 11)$.

Find the work done if the distance is measured in meters and the magnitude of the force is measured in newtons.

Work: _____ Nm

10. A woman exerts a horizontal force of 7 pounds on a box as she pushes it up a ramp that is 6 feet long and inclined at an angle of 30 degrees above the horizontal.

Find the work done on the box.

Work: _____ ft-lb

11. If Yoda says to Luke Skywalker, "The Force be with you," then the dot product of the Force and Luke should be:

◯ negative

◯ positive

◯ zero

◯ any real number

12. Find the angle between the diagonal of a cube of side length 18 and the diagonal of one of its faces. The angle should be measured in radians.

13. Find the following expressions if $\mathbf{a} = 3\mathbf{i} - 4\mathbf{k}$, $\mathbf{b} = \mathbf{i} + \mathbf{j} - 4\mathbf{k}$, and $\mathbf{c} = 4\mathbf{i} + 6\mathbf{j} + 3\mathbf{k}$.

1. $\mathbf{a} \cdot \mathbf{b} =$ _____
2. $\mathbf{a} \cdot \mathbf{c} =$ _____
3. $\mathbf{b} \cdot \mathbf{c} =$ _____

14. A woman exerts a horizontal force of 1 pounds on a box as she pushes it up a ramp that is 2 feet long and inclined at an angle of 30 degrees above the horizontal.
The force vector is F = _____
The displacement vector is d = _____
Find the work done on the box.
Work = F·d = _____ ft-lb

15. An object moves along a straight track from the point $(-4, -2, -3)$ to the point $(-6, -10, 7)$. The only force acting on it is a constant $\mathbf{F} = -4\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}$ newtons. Find the work done if the distance is measured in meters.
Work = _____ joules.

16. Find the scalar and vector projections of $\mathbf{b}$ onto $\mathbf{a}$, where $\mathbf{a} = \langle -1, 1, 2\rangle$ and $\mathbf{b} = \langle -2, 7, 12\rangle$.
1. $\text{comp}_{\mathbf{a}}\mathbf{b} =$ _____
2. $\text{proj}_{\mathbf{a}}\mathbf{b} =$ _____

17. Find the scalar and vector projection of the vector $\mathbf{b} = \langle 1, -3, -3\rangle$ onto the vector $\mathbf{a} = \langle -1, -5, -3\rangle$.
Scalar projection (i.e., component): _____
Vector projection ⟨ _____, _____, _____⟩

18. Let $\vec{v} = \langle -2, 5\rangle$ in $\mathbb{R}^2$, and let $\vec{y} = \langle 0, 3, -2\rangle$ in $\mathbb{R}^3$.

(a) Is $\langle 2, -1\rangle$ perpendicular to $\vec{v}$? Why or why not?

(b) Find a unit vector $\vec{u}$ in $\mathbb{R}^2$ such that $\vec{u}$ is perpendicular to $\vec{v}$. How many such vectors are there? Justify your answers.

(c) Is $\langle 2, -1, -2\rangle$ perpendicular to $\vec{y}$? Why or why not?

(d) Find a unit vector $\vec{w}$ in $\mathbb{R}^3$ such that $\vec{w}$ is perpendicular to $\vec{y}$. How many such vectors are there?Justify your answers.

(e) Let $\vec{z} = \langle 2, 1, 0\rangle$. Find a unit vector $\vec{r}$ in $\mathbb{R}^3$ such that $\vec{r}$ is perpendicular to both $\vec{y}$ and $\vec{z}$. How many such vectors are there? Explain your process.

19. In this exercise, we will show that the roles of $\vec{u}$ and $\vec{v}$ are not symmetric in projection formula (Definition 9.4.10). We will use the same values as Example 9.4.12, namely $\vec{u} = \langle 4, 3\rangle$ and $\vec{v} = \langle 0, 2\rangle$, but we will be looking at the projection of $\vec{v}$ onto $\vec{u}$.

(a) Will $proj_{\vec{u}}\vec{v}$ be parallel to $\vec{u}$ or $\vec{v}$? Use this answer to explain how you know that the result of $proj_{\vec{u}}\vec{v}$ must be different than $proj_{\vec{v}}\vec{u}$ (which we calculated in Example 9.4.12).

(b) Use Definition 9.4.10 to calculate $comp_{\vec{u}}\vec{v}$, $proj_{\vec{u}}\vec{v}$, and $proj_{\perp\vec{u}}\vec{v}$.

(c) Draw a plot with $\vec{u}$, $\vec{v}$, $proj_{\vec{u}}\vec{v}$, and $proj_{\perp\vec{u}}\vec{v}$.

(d) Write a few sentences to explain how and why your plot for the previous part is different than Figure 9.4.13.

20. Consider the triangle in $\mathbb{R}^3$ given by $P = (3, 2, -1)$, $Q = (1, -2, 4)$, and $R = (4, 4, 0)$.

(a) Find the measure of each of the three angles in the triangle, accurate to 0.01 degrees.

(b) Choose two sides of the triangle, and call the vectors that form the sides (emanating from a common point) $\vec{a}$ and $\vec{b}$.

(a) Compute $\text{proj}_{\vec{b}}\vec{a}$, and $\text{proj}_{\perp\vec{b}}\vec{a}$.

(b) Explain why $\text{proj}_{\perp\vec{b}}\vec{a}$ can be considered a height of triangle PQR.

(c) Find the area of the given triangle.

21. Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^5$ with $\vec{u} \cdot \vec{v} = -1$, $\|\vec{u}\| = 2$, $\|\vec{v}\| = 3$. Use the properties of the dot product to find each of the following.

(a) $\vec{u} \cdot 2\vec{v}$

(b) $\vec{v} \cdot \vec{v}$

(c) $(\vec{u} + \vec{v}) \cdot \vec{v}$

(d) $(2\vec{u} + 4\vec{v}) \cdot (\vec{u} - 7\vec{v})$

(e) $\|\vec{u}\|\|\vec{v}\| \cos(\theta)$, where θ is the angle between $\vec{u}$ and $\vec{v}$

(f) θ

22. One of the properties of the dot product is that $(\vec{u}+\vec{v})\cdot\vec{w} = (\vec{u}\cdot\vec{w})+(\vec{v}\cdot\vec{w})$. That is, the dot product distributes over vector addition *on the right*. Here we investigate whether the dot product distributes over vector addition *on the left*.

(a) Let $\vec{u} = \langle 1, 2, -1\rangle$, $\vec{v} = \langle 4, -3, 6\rangle$, and $\vec{w} = \langle 4, 7, 2\rangle$. Calculate

$$\vec{u} \cdot (\vec{v} + \vec{w}) \quad \text{and} \quad (\vec{u} \cdot \vec{v}) + (\vec{u} \cdot \vec{w}).$$

What do you notice?

(b) Use the properties of the dot product to show that in general

$$\vec{x} \cdot (\vec{y} + \vec{z}) = (\vec{x} \cdot \vec{y}) + (\vec{x} \cdot \vec{z})$$

for any vectors $\vec{x}$, $\vec{y}$, and $\vec{z}$ in $\mathbb{R}^n$.

23. When running a sprint, the racers may be aided or slowed by the wind. The wind assistance is a measure of the wind speed that is helping push the runners down the track. It is much easier to run a very fast race if the wind is blowing hard in the direction of the race. So that world records aren't dependent on the weather conditions, times are only recorded as record times if the wind aiding the runners is less than or equal to 2 meters per second. Wind speed for a race is recorded by a wind gauge that is set up close to the track. It is important to note, however, that weather is not always as cooperative as we might like. The wind does not always blow exactly in the direction of the track, so the gauge must account for the angle the wind makes with the track. Suppose a 4 mile per hour wind is blowing to aid runners by making a 38° angle with the race track. Determine if any times set during such a race would qualify as records.

24.

Molecular geometry is the geometry determined by arrangements of atoms in molecules. Molecular geometry includes measurements like bond angle, bond length, and torsional angles. These attributes influence several properties of molecules, such as reactivity, color, and polarity.

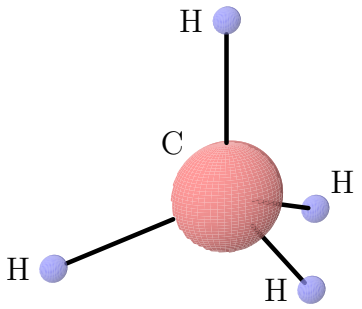


Figure 9.4.15 A methane molecule.

As an example of the molecular geometry of a molecule, consider the methane CH_4 molecule, as illustrated in Figure 9.4.15. According to the Valence Shell Electron Repulsion (VSEPR) model, atoms that surround single different atoms do so in a way that positions them as far apart as possible. This means that the hydrogen atoms in the methane molecule arrange themselves at the vertices of a regular tetrahedron. The *bond angle* for methane is the angle determined by two consecutive hydrogen atoms and the central carbon atom. To determine the bond angle for methane, we can place the center carbon atom at the point $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ and the hydrogen atoms at the points $(0,0,0)$, $(1,1,0)$, $(1,0,1)$, and $(0,1,1)$. Find the bond angle for methane to the nearest tenth of a degree.

9.5 The Cross Product

Motivating Questions

- How and when is the cross product of two vectors defined?
- What is the relationship between the direction of the cross product of two vectors to the direction of those vectors?
- What does the magnitude of the cross product represent?

9.5.1 Introduction

The last two sections have introduced some basic algebraic operations and useful geometric interpretations for vector addition, scalar multiplication, and the dot product. In this section, we will define and explore the **cross product**, another algebraic operation which measures useful geometric information related to pairs of vectors in $\mathbb{R}^3$.

Remember that we use a right-handed coordinate system, as described in Section 9.2. In particular, recall that the vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are oriented as shown below in Figure 9.5.1. We would like to think of a right-handed coordinate system in terms of vectors that are not just $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

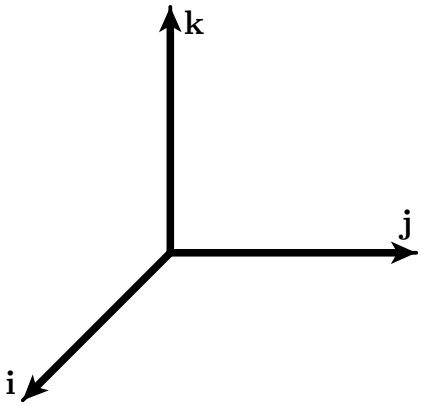


Figure 9.5.1 Basis vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

We will call an ordered list of vectors $\{\vec{u}, \vec{v}, \vec{w}\}$ a **right-handed coordinate system** (or right-handed orientation) provided that putting the fingers of your right hand in the direction of $\vec{u}$ and curling your fingers in the direction of $\vec{v}$ results in the thumb of your right hand pointing in the direction of $\vec{w}$. Remember that to get this orientation correct, you may need to rotate your right hand to get your fingers to curl in the correct direction, which will then give you the proper direction for the third vector $\vec{w}$.

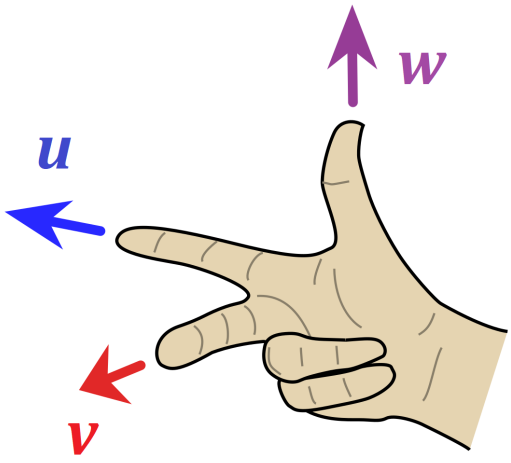


Figure 9.5.2 A right-handed orientation of vectors $\{\vec{u}, \vec{v}, \vec{w}\}$

We would like to create a product of two vectors, which we will call the cross product, that creates a right handed coordinate system with the input vectors as the first two vectors in our coordinate system. We will denote the cross product of $\vec{u}$ and $\vec{v}$ by $\vec{u} \times \vec{v}$. If $\vec{u}$ and $\vec{v}$ are not parallel, we want the cross product of $\vec{u}$ with $\vec{v}$ to make $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ a right-handed coordinate system. Additionally, we would like the cross product to work well with linear combinations of vectors. Specifically, we want the following algebraic property to be satisfied:

$$(k\vec{v} + \vec{u}) \times \vec{w} = k(\vec{v} \times \vec{w}) + (\vec{u} \times \vec{w})$$

An important difference to note here is that the cross product of two vectors will be a vector, whereas the dot product of two vectors gives a scalar.

In other sources, you may see the dot product referred to as the scalar product of vectors and the cross product referred to as the vector product. These terms comes from the result of the dot product being a scalar and the result of the cross product being a vector.

Preview Activity 9.5.1

(a) Our first task is to understand what it means to complete a right-handed coordinate system. For each of the cases below, you need to give a vector that fills in the blank to create a right-handed coordinate system. For example, the answer that would complete $\{\hat{\imath}, \hat{\jmath}, ???\}$ would be $\hat{k}$. You should be careful to when you have vectors that are in the opposite directions of $\hat{\imath}$ or $\hat{\jmath}$ or $\hat{k}$.

(i) $\{\hat{\jmath}, \hat{\imath}, ???\}$

(ii) $\{-\hat{\imath}, \hat{\jmath}, ???\}$

(iii) $\{\hat{k}, \hat{\imath}, ???\}$

(iv) $\{-\hat{\imath}, -\hat{\jmath}, ???\}$

(v) $\{\hat{k}, \hat{\jmath}, ???\}$

(b) Explain why there is not a way to complete $\{\hat{\imath}, \hat{\imath}, ???\}$ to be a right-handed coordinate system.

(c) As we said above, we would like the cross product to distribute over vector sums ($(\vec{v} + \vec{u}) \times \vec{w} = (\vec{v} \times \vec{w}) + (\vec{u} \times \vec{w})$). Complete each of the following right-handed coordinate systems geometrically. Focus on the direction of the vector needed to complete the right-handed coordinate system and do not worry about the length of the vector you choose.

(i) $\{\hat{\imath}, \hat{k}, ???\}$

(ii) $\{\hat{\jmath}, \hat{k}, ???\}$

(iii) $\{\hat{\imath} + \hat{\jmath}, \hat{k}, ???\}$

Be sure to note how your thumb is oriented in the last case and verify that the sum of the first two results gives you the last case.

(d) Finally, we want to address the relationship between the cross product and scalar multiplication.

(i) What direction should the third vector in the right-handed coordinate system $\{2\hat{\imath}, 3\hat{\jmath}, ???\}$ be?

(ii) What direction should the third vector in the right-handed coordinate system $\{2\hat{k}, \hat{\jmath}, ???\}$ be?

(iii) Note that $3\hat{k} = \hat{k} + \hat{k} + \hat{k}$, so $(3\hat{k}) \times \hat{\jmath} = (\hat{k} \times \hat{\jmath}) + (\hat{k} \times \hat{\jmath}) + (\hat{k} \times \hat{\jmath})$. What do you think the vector $(3\hat{k}) \times \hat{\jmath}$ will be?

(iv) What do you think the result of $(4\hat{\imath}) \times (-6\hat{\jmath})$ will be?

9.5.2 Computing the cross product

As we have seen in Preview Activity 9.5.1, the cross product $\vec{u} \times \vec{v}$ will be defined for two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$ and produces another vector in $\mathbb{R}^3$. Using the right-hand rule, we saw that

$$\hat{\imath} \times \hat{\jmath} = \hat{k} \qquad \hat{\imath} \times \hat{k} = -\hat{\jmath} \qquad \hat{\jmath} \times \hat{k} = \hat{\imath}$$

$$\hat{\jmath} \times \hat{\imath} = -\hat{k} \qquad \hat{k} \times \hat{\imath} = \hat{\jmath} \qquad \hat{k} \times \hat{\jmath} = -\hat{\imath}$$

We also want the cross product to distribute over linear combinations so that $(k\vec{v} + \vec{u}) \times \vec{w} = k(\vec{v} \times \vec{w}) + (\vec{u} \times \vec{w})$. Using the ideas directly above, we can compute the cross product in terms of the components of general vectors. Applying these ideas of right-handed coordinate systems and the cross product working with linear combinations of vectors gives us the following:

$$\begin{aligned}
\vec{u} \times \vec{v} &= (u_1\hat{\imath} + u_2\hat{\jmath} + u_3\hat{k}) \times (v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}) \\
&= u_1\hat{\imath} \times (v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}) + u_2\hat{\jmath} \times (v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}) \\
&\quad + u_3\hat{k} \times (v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}) \\
&= u_1v_1\hat{\imath} \times \hat{\imath} + u_1v_2\hat{\imath} \times \hat{\jmath} + u_1v_3\hat{\imath} \times \hat{k} + u_2v_1\hat{\jmath} \times \hat{\imath} + u_2v_2\hat{\jmath} \times \hat{\jmath} \\
&\quad + u_2v_3\hat{\jmath} \times \hat{k} + u_3v_1\hat{k} \times \hat{\imath} + u_3v_2\hat{k} \times \hat{\jmath} + u_3v_3\hat{k} \times \hat{k} \\
&= u_1v_2\hat{k} - u_1v_3\hat{\jmath} - u_2v_1\hat{k} + u_2v_3\hat{\imath} + u_3v_1\hat{\jmath} - u_3v_2\hat{\imath} \\
&= (u_2v_3 - u_3v_2)\hat{\imath} - (u_1v_3 - u_3v_1)\hat{\jmath} + (u_1v_2 - u_2v_1)\hat{k}.
\end{aligned}$$

Like the dot product, the cross product arises in physical applications, e.g., torque, but it is more convenient mathematically to begin from an algebraic perspective.

Definition 9.5.3 The **cross product** $\vec{u} \times \vec{v}$ of vectors $\vec{u} = u_1\hat{\imath} + u_2\hat{\jmath} + u_3\hat{k}$ and $\vec{v} = v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}$ in $\mathbb{R}^3$ is the vector

$$\vec{u} \times \vec{v} = (u_2v_3 - u_3v_2)\hat{\imath} - (u_1v_3 - u_3v_1)\hat{\jmath} + (u_1v_2 - u_2v_1)\hat{k}. \tag{9.5.1}$$

◆

As presented, the expression in (9.5.1) is complicated, and the variety of subscripts and minus signs may appear difficult to remember. However, we can rewrite this definition using determinants, which allows important structure to emerge. The determinant of a 2×2 matrix is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

It follows that we can thus rewrite Equation (9.5.1) in the form

$$\vec{u} \times \vec{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \hat{\imath} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \hat{\jmath} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \hat{k}.$$

For those familiar with the determinant of a 3×3 matrix, we write the mnemonic as

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

Example 9.5.4 We will use the definition of cross product to check some of our work from Preview Activity 9.5.1. First, we want to check a calculation like $\hat{k} \times \hat{\imath}$ from part (a). We will write both $\hat{k}$ and $\hat{\imath}$ in their component forms of $\langle 0, 0, 1\rangle$ and $\langle 1, 0, 0\rangle$. So $\hat{k} \times \hat{\imath}$ will be computed as

$$\begin{aligned}
\langle 0, 0, 1\rangle \times \langle 1, 0, 0\rangle &= \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \hat{\imath} - \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \hat{\jmath} + \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \hat{k} \\
&= 0\hat{\imath} - (-1)\hat{\jmath} + 0\hat{k} \\
&= \langle 0, 1, 0\rangle
\end{aligned}$$

which matches our answer of $\hat{\jmath}$ from the Preview Activity.

Next we want to check our work from part (c) when we found $(\hat{\imath}+\hat{\jmath})\times\hat{k}$ geometrically.

$$\begin{aligned}\langle 1,1,0\rangle \times \langle 0,0,1\rangle &= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\hat{\imath} - \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\hat{\jmath} + \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix}\hat{k} \\ &= 1\hat{\imath} - (1)\hat{\jmath} + 0\hat{k} \\ &= \langle 1,-1,0\rangle\end{aligned}$$

which matches our answer of $\hat{\imath}-\hat{\jmath}$ from part (c) of the Preview Activity.

For our last part of this example we will verify our calculation of $(4\hat{\imath})\times(-6\hat{\jmath})$ from part (d) of the Preview Activity.

$$\begin{aligned}\langle 4,0,0\rangle \times \langle 0,-6,0\rangle &= \begin{vmatrix} 0 & 0 \\ -6 & 0 \end{vmatrix}\hat{\imath} - \begin{vmatrix} 4 & 0 \\ 0 & 0 \end{vmatrix}\hat{\jmath} + \begin{vmatrix} 4 & 0 \\ 0 & -6 \end{vmatrix}\hat{k} \\ &= 0\hat{\imath} - 0\hat{\jmath} + (-24)\hat{k} \\ &= \langle 0,0,-24\rangle\end{aligned}$$

Which should match our work of $-24\hat{k}$ from the last part of the Preview Activity. ▲

The next activity gives you an opportunity to practice computing cross products and to start investigating its algebraic properties.

Activity 9.5.2 Suppose $\vec{u}=\langle 0,1,3\rangle$ and $\vec{v}=\langle 2,-1,0\rangle$. Use equation (9.5.1) for the following.

(a) Find the cross product $\vec{u}\times\vec{v}$.

(b) Evaluate the dot products $\vec{u}\cdot(\vec{u}\times\vec{v})$ and $\vec{v}\cdot(\vec{u}\times\vec{v})$. What does this tell you about the geometric relationship among $\vec{u}$, $\vec{v}$, and $\vec{u}\times\vec{v}$?

(c) Find the cross product $\vec{v}\times\hat{\imath}$.

(d) Recall that multiplication of real numbers is *associative*. For example, $(2\cdot 5)\cdot 3 = 2\cdot(5\cdot 3)$. Is it true that the cross product of vectors is associative? For instance, is it true that $(\vec{u}\times\vec{v})\times\hat{\imath} = \vec{u}\times(\vec{v}\times\hat{\imath})$?

(e) Find the cross product $\vec{u}\times\vec{u}$ and write a sentence or two to explain the meaning of your result.

Hint. Your ideas for part 9.5.1.b will apply here.

The cross product satisfies the following properties, some of which were illustrated in Preview Activity 9.5.1 or may be verified from the definition (9.5.1).

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. Then

1. $\vec{u}\times\vec{v} = -(\vec{v}\times\vec{u})$
2. $(\vec{u}+\vec{v})\times\vec{w} = (\vec{u}\times\vec{w}) + (\vec{v}\times\vec{w})$
3. $(c\vec{u})\times\vec{w} = c(\vec{u}\times\vec{w}) = \vec{u}\times(c\vec{w})$
4. $\vec{u}\times\vec{v} = \vec{0}$ if $\vec{u}$ and $\vec{v}$ are parallel.

5. The cross product is not associative; that is, in general

$$(\vec{u} \times \vec{v}) \times \vec{w} \neq \vec{u} \times (\vec{v} \times \vec{w}).$$

While some of these properties may not seem natural, we need to look at these from the perspective of what we desired the cross product to measure. 1) The result of the cross product should complete a right handed coordinate system and 2) the cross product should work well with linear combinations by distributing across vector addition and scalar multiplication. The first property above is a consequence of the cross product being the vector needed to complete the associated right-handed coordinate system. If you switch the order of the first two vectors in our right-handed coordinate system, then the third vector must switch direction in order for the orientation to follow the right hand rule. The second and third properties are what we meant for the cross product to work well with linear combinations and can be verified algebraically from (9.5.1), and the fifth should make sense after thinking about part d of Subsection 9.5.2.

9.5.3 The Vector Nature of the Cross Product

The output of the cross product is a vector and thus has both magnitude and direction. In this section, we will split up the cross product's geometric meaning into these two fundamental properties of vectors. We will begin by first mentioning that the direction of the cross product (as defined in equation (9.5.1)) satisfies the claim need from the introduction to this section.

Key Idea 9.5.5 *The cross product of $\vec{u}$ and $\vec{v}$ is oriented to make $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ a right-handed coordinate system. In particular, this means that $\vec{u} \times \vec{v}$ is orthogonal to both $\vec{u}$ and $\vec{v}$.*

Proof. Remember from Preview Activity 9.5.1 that the cross product of $\vec{u}$ and $\vec{v}$ will complete the right handed coordinate system, $\{\vec{u}, \vec{v}, ???\}$. In particular, the last vector in our right handed coordinate system will be at a $\frac{\pi}{2}$ or $90°$ angle to the other vectors because your thumb makes a $90°$ angle with the plane that your fingers can extend and curl in. We defined the cross product to satisfy the right-hand orientation relationship for $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ (and for linear combinations of these vectors), so we need only verify the orthogonal relationship algebraically by computing $\vec{u} \cdot (\vec{u} \times \vec{v})$.

$$\begin{aligned}\vec{u} \cdot (\vec{u} \times \vec{v}) &= u_1(u_2v_3 - u_3v_2) - u_2(u_1v_3 - u_3v_1) + u_3(u_1v_2 - u_2v_1) \\ &= u_1u_2v_3 - u_1u_3v_2 - u_2u_1v_3 + u_2u_3v_1 + u_3u_1v_2 - u_3u_2v_1 \\ &= 0\end{aligned}$$

To summarize, we have $\vec{u} \cdot (\vec{u} \times \vec{v}) = 0$, which implies that $\vec{u}$ is orthogonal to $\vec{u} \times \vec{v}$. A similar algebraic argument will show that $\vec{v}$ is orthogonal to $\vec{u} \times \vec{v}$, and thus the output of $\vec{u} \times \vec{v}$ is a vector that is perpendicular to both input vectors, $\vec{u}$ and $\vec{v}$. Because the cross product works well with linear combinations, $\vec{u} \times \vec{v}$ is perpendicular to any linear combination of $\vec{u}$ and $\vec{v}$ (to be shown in 23). ■

We now look at the magnitude of $\vec{u} \times \vec{v}$ in terms of the vector attributes (magnitude and direction) of the input vectors $\vec{u}$ and $\vec{v}$.

Key Idea 9.5.6 *The length of the cross product of $\vec{u}$ and $\vec{v}$ satisfies*

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin(\theta) \tag{9.5.2}$$

where θ is the angle between $\vec{u}$ and $\vec{v}$.

Proof. To investigate, we will compute the square of the length $\|\vec{u} \times \vec{v}\|^2$ and denote by θ the angle between $\vec{u}$ and $\vec{v}$, as in Subsection 9.4.3. The following simplification is not intuitive, but you can find through some significant algebra that

$$\begin{aligned}
\|\vec{u} \times \vec{v}\|^2 &= (u_2v_3 - u_3v_2)^2 + (u_1v_3 - u_3v_1)^2 + (u_1v_2 - u_2v_1)^2 \\
&= u_2^2v_3^2 - 2u_2u_3v_2v_3 + u_3^2v_2^2 + u_1^2v_3^2 - 2u_1u_3v_1v_3 + u_3^2v_1^2 \\
&\quad + u_1^2v_2^2 - 2u_1u_2v_1v_2 + u_2^2v_1^2 \\
&= u_1^2(v_2^2 + v_3^2) + u_2^2(v_1^2 + v_3^2) + u_3^2(v_1^2 + v_2^2) \\
&\quad - 2(u_1u_2v_1v_2 + u_1u_3v_1v_3 + u_2u_3v_2v_3) \\
&= u_1^2(v_1^2 + v_2^2 + v_3^2) + u_2^2(v_1^2 + v_2^2 + v_3^2) + u_3^2(v_1^2 + v_2^2 + v_3^2) \\
&\quad - (u_1^2v_1^2 + u_2^2v_2^2 + u_3^2v_3^2 + 2(u_1u_2v_1v_2 + u_1u_3v_1v_3 + u_2u_3v_2v_3)) \\
&= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\
&= \|\vec{u}\|^2 \|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2 \\
&= \|\vec{u}\|^2 \|\vec{v}\|^2 (1 - \cos^2(\theta)) \\
&= \|\vec{u}\|^2 \|\vec{v}\|^2 \sin^2(\theta).
\end{aligned}$$

Therefore, we have found $\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \|\vec{v}\|^2 \sin^2(\theta)$. ■

Note that Key Idea 9.5.6 stated above implies $\vec{u} \times \vec{v} = \vec{0}$ if $\vec{u}$ and $\vec{v}$ are parallel. If $\vec{u}$ and $\vec{v}$ are parallel, then $\theta = 0$ or π. Thus, equation (9.5.2) tells us that $\sin(\theta) = 0$ implies $\vec{u} \times \vec{v} = \vec{0}$.

9.5.4 Applications of the Cross Product

Equation (9.5.2) is also related to the parallelogram formed by two vectors $\vec{u}$ and $\vec{v}$, as shown in Figure 9.5.7. We say that $\vec{w}$ is the **outward direction** of the parallelogram formed by $\vec{u}$ and $\vec{v}$ provided that $\vec{w}$ is perpendicular (in $\mathbb{R}^3$) to the parallelogram and $\{\vec{u}, \vec{v}, \vec{w}\}$ has a right-handed orientation. You should now convince yourself that in Figure 9.5.7 the outward direction is coming directly out of the page (or screen). Remember that the order of our vectors matters for orientation purposes, so the outward direction for the parallelogram formed by $\vec{v}$ and $\vec{u}$ would be *into* the page.

The parallelogram with a horizontal segment drawn with a horizontal arrow going to the right, labeled $\vec{u}$ and the adjacent side drawn with an arrow that goes more up than to the right and is labeled $\vec{v}$. The angle from the first arrow to the second is drawn by a circular arc and labeled θ. A vertical line segment from the end of the second arrow down to the horizontal arrow is labeled $\|\vec{v}\| \sin(\theta)$.

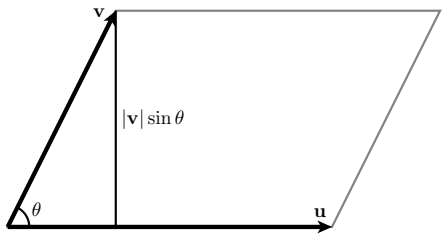


Figure 9.5.7 The parallelogram formed by $\vec{u}$ and $\vec{v}$

Key Idea 9.5.8 *The magnitude of the cross product, $\vec{u} \times \vec{v} = \|\vec{u}\|\|\vec{v}\| \sin(\theta)$, is the area of the parallelogram defined by $\vec{u}$ and $\vec{v}$. Additionally, $\vec{u} \times \vec{v}$ is the*

outward direction for the parallelogram defined by $\vec{u}$ and $\vec{v}$.

Proof. The area of a parallelogram is the product of its base and height. Based on Figure 9.5.7, we consider the base of the parallelogram to be $\|\vec{u}\|$ and the height to be $\|\vec{v}\|\sin(\theta)$. This means that the area of the parallelogram formed by $\vec{u}$ and $\vec{v}$ is

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

If $\vec{u} = u_1\hat{\imath} + u_2\hat{\jmath} + 0\hat{k}$ and $\vec{v} = v_1\hat{\imath} + v_2\hat{\jmath} + 0\hat{k}$ are vectors in the xy-plane, then Equation (9.5.1) shows that the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$ is $\|\vec{u}\times\vec{v}\| = |u_1v_2 - u_2v_1|$ is the absolute value of the 2×2 determinant $\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$. Thus, the absolute value of a determinant of a 2×2 matrix is also the area of a parallelogram.

The outward direction of the parallelogram is a consequence of Key Idea 9.5.5. ■

The next activity gives you a chance to practice thinking about how to use the cross product to answer questions about parallelograms.

Activity 9.5.3

(a) Find the area of the parallelogram formed by the vectors $\vec{u} = \langle 1, 3, -2\rangle$ and $\vec{v} = \langle 3, 0, 1\rangle$.

(b) Find the area of the parallelogram in $\mathbb{R}^3$ whose vertices are $(1, 0, 1)$, $(0, 0, 1)$, $(2, 1, 0)$, and $(1, 1, 0)$.

Hint. It might be helpful to draw a picture to see how the vertices are arranged so you can determine which vectors you might use.

There is yet one more geometric implication we may draw from Key Idea 9.5.8. Suppose $\vec{u}$, $\vec{v}$, and $\vec{w}$ are vectors in $\mathbb{R}^3$ that are not coplanar. (That is, $\vec{w}$ is not a linear combination of $\vec{u}$ and $\vec{v}$.) In this case, $\vec{u}$, $\vec{v}$, and $\vec{w}$ define a three-dimensional solid called a **parallelepiped** and depicted in Figure 9.5.9.

The parallelepiped is drawn with three edges shown starting from a common point. The first arrow is labeled $\vec{u}$ and is changing only horizontally. The second vector moves to the right and up and is labeled $\vec{v}$. The third arrow moves more up than right and is labeled $\vec{w}$. A vertical line segment is drawn from the end of the third vector to the parallelogram on the bottom of the figure. This vertical segment is labeled h. A vertical arrow is drawn from the same starting point as the other three vectors and is labeled $\vec{n}$. The angle from the third vector to the vertical arrow is highlighted with a curved arc and labeled α.

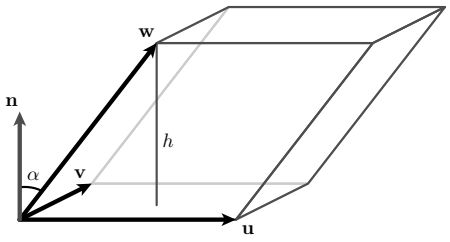


Figure 9.5.9 The parallelepiped determined by $\vec{u}$, $\vec{v}$, and $\vec{w}$

The volume of the parallelepiped is determined by multiplying the area A of the base by the height h. As we have just seen, the area of the base is $\|\vec{u}\times\vec{v}\|$. Moreover, $h = \|\vec{w}\|\cos(\alpha)$ where α is the angle between $\vec{w}$ and the vector $\vec{n}$ orthogonal to the plane formed by $\vec{u}$ and $\vec{v}$. Since $\vec{n}$ is parallel to $\vec{u}\times\vec{v}$,

the angle between $\vec{w}$ and $\vec{u} \times \vec{v}$ is also α. This shows that

$$|(\vec{u} \times \vec{v}) \cdot \vec{w}| = \|\vec{u} \times \vec{v}\| \, \|\vec{w}\| \cos(\alpha) = Ah.$$

We summarize this useful relationship below.

The cross product and the volume of a parallelepiped.

The volume of the parallelepiped determined by $\vec{u}$, $\vec{v}$, and $\vec{w}$ is $|(\vec{u} \times \vec{v}) \cdot \vec{w}|$.

As a dot product of two vectors, the quantity $(\vec{u} \times \vec{v}) \cdot \vec{w}$ is a scalar and is called the **scalar triple product**.

Activity 9.5.4 Suppose $\vec{u} = \langle 3, 5, -1 \rangle$ and $\vec{v} = \langle 2, -2, 1 \rangle$.

(a) Find two unit vectors orthogonal to both $\vec{u}$ and $\vec{v}$.

(b) Find the volume of the parallelepiped formed by the vectors $\vec{u}$, $\vec{v}$, and $\vec{w} = \langle 3, 3, 1 \rangle$.

(c) Find a vector orthogonal to the parallelogram containing the points $(0, 1, 2)$, $(4, 1, 0)$, and $(-2, 2, 2)$.

(d) Given the vectors $\vec{u}$ and $\vec{v}$ shown below in Figure 9.5.10, sketch the cross products $\vec{u} \times \vec{v}$ and $\vec{v} \times \vec{u}$.

A plot with three mutually perpendicular axes labeled with x going down and left, y going to the right and z going up. The first black arrow is shown going up and to the right and is labeled $\vec{v}$. A second black arrow going to the left and up and is labeled $\vec{u}$.

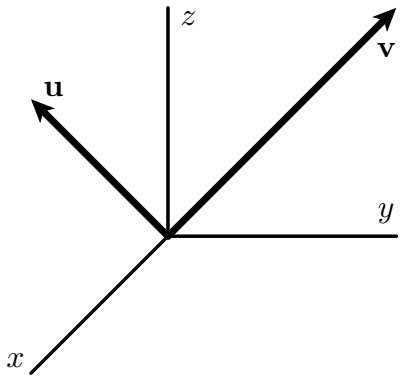


Figure 9.5.10 Vectors $\vec{u}$ and $\vec{v}$

(e) Are the vectors $\vec{a} = \langle 1, 3, -2 \rangle$,$\vec{b} = \langle 2, 1, -4 \rangle$, and $\vec{c} = \langle 0, 1, 0 \rangle$ in standard position coplanar? Use the concepts from this section to explain your answer.

We have seen that the cross product enables us to produce a vector perpendicular to two given vectors, to measure the area of a parallelogram, and to measure the volume of a parallelepiped. Besides these geometric applications, the cross product also enables us to describe a physical quantity called **torque**.

Suppose that we would like to turn a bolt using a wrench as shown in Figure 9.5.11. If a force $\vec{F}$ is applied to the wrench and $\vec{r}$ is the vector from the position on the wrench at which the force is applied to center of the bolt,

we define the **torque**, $\overrightarrow{\tau}$, to be

$$\overrightarrow{\tau} = \vec{F} \times \vec{r}.$$

A drawing of a wrench drawn vertically with an arrow through the middle of the wrench. Another black arrow is shown moving down and left with ending point being the bottom of the wrench. The angle between these arrows is highlighted by a curved arc and labeled θ. A horizontal segment is drawn in gray from the start of the second arrow to the middle of the wrench.

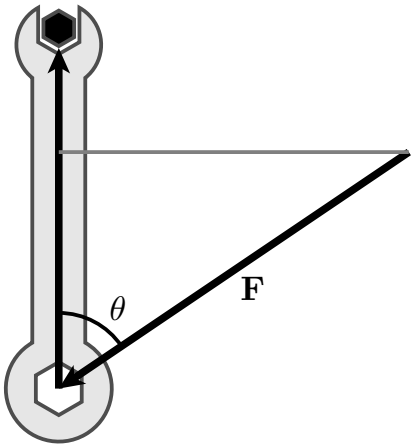


Figure 9.5.11 A force applied to a wrench

When a force is applied to an object, Newton's Second Law tells us that the force is equal to the rate of change of the object's linear momentum. Similarly, the torque applied to an object is equal to the rate of change of the object's angular momentum. In other words, torque will cause the bolt to rotate.

In many industrial applications, bolts are required to be tightened using a specified torque. Of course, the magnitude of the torque is $\|\overrightarrow{\tau}\| = \left\|\vec{F} \times \vec{r}\right\| = \left\|\vec{F}\right\| \|\vec{r}\| \sin(\theta)$. Thus, to produce a larger torque, we can increase either $\left\|\vec{F}\right\|$ or $\|\vec{r}\|$, which you may know if you have ever removed lug nuts when changing a flat tire. Increasing $\left\|\vec{F}\right\|$ means you exert more force and increasing $\|\vec{r}\|$ means you use a longer wrench. The ancient Greek scientist and mathematician Archimedes said: "Give me a lever long enough and a fulcrum on which to place it, and I shall move the world." A modern spin on this statement is: "Allow me to make $\|\vec{r}\|$ large enough, and I shall produce a torque large enough to move the world."

When opening a door, you use the knob or push bar as far from the hinge as possible. If you were to try to open a door by pushing on the middle of the door (rather than the edge farthest from the hinges), you would need to push twice as hard as normal because you have half of $\|\vec{r}\|$.

9.5.5 Comparing the dot and cross products

In this subsection, we will summarize the last few sections by comparing and contrasting several properties of the dot and cross products.

1. $\vec{u} \cdot \vec{v}$ is a scalar, while $\vec{u} \times \vec{v}$ is a vector.
2. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$, while $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$
3. $\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos(\theta)$, while $\|\vec{u} \times \vec{v}\| = \|\vec{u}\|\|\vec{v}\| \sin(\theta)$.

4. $\vec{u} \cdot \vec{v} = 0$ if $\vec{u}$ and $\vec{v}$ are perpendicular, while $\vec{u} \times \vec{v} = \vec{0}$ if $\vec{u}$ and $\vec{v}$ are parallel.

9.5.6 Object Types and Vector Notation

In this final subsection, we practice making sense of different expressions involving vectors by describing what kind object the result will be. It is easy to get caught up in all of the manipulations of symbols and new operations and write an expression that *looks* reasonable but is not defined. For example, $(\vec{u} \cdot \vec{v}) \times \vec{w}$ looks like a nice and reasonable calculation involving vectors (and looks a lot like a scalar triple product!). However, $(\vec{u} \cdot \vec{v}) \times \vec{w}$ is undefined because the quantity in the parentheses is a scalar and you can't take the cross product of a scalar and a vector.

Activity 9.5.5 In this activity, we are focused on the type of objects being used and whether the expression makes sense to do at all. We are *not* going to worry about interpreting or understanding what is being measured by these expressions.

(a) For each of the expressions below, state whether the result is a scalar, a vector, or undefined. You should write a sentence or two about each to explain your reasoning. (Assume that vectors are nonzero and not orthogonal.)

(i) $\dfrac{\vec{v}}{\vec{w}}$

(ii) $\dfrac{\vec{v}}{\vec{w} \cdot \vec{v}}$

(iii) $(\vec{u} \times \vec{w}) + \vec{v}$

(iv) $k(\vec{u} \cdot \vec{w}) + c\vec{v}$

(v) $k(\vec{u} \cdot \vec{v}) \cdot \vec{w}$

(vi) $\dfrac{\vec{v}}{\|\vec{w}\|}$

(vii) $(k + \vec{u}) \times \vec{v}$

(viii) $k + (\vec{u} \times \vec{v})$

(ix) $\vec{v} + (\vec{u} \cdot \vec{w})$

(x) $k + \|\vec{u} \times \vec{w}\|$

(b) Use the operations of dot product and vector subtraction to write an expression involving $\vec{u}$, $\vec{v}$, and $\vec{w}$ that evaluates to a scalar. You can use other operations if you want.

(c) Use the operations of cross product, scalar multiplication, and vector addition to write an expression involving $\vec{u}$, $\vec{v}$, and $\vec{w}$ that evaluates to a vector. You can use other operations if you want.

(d) Use the operations of dot product and vector addition to write an expression involving $\vec{u}$, $\vec{v}$, and $\vec{w}$ that is undefined. You can use other operations if you want.

Summary

- The cross product is defined *only* for vectors in $\mathbb{R}^3$. The cross product of vectors $\vec{u} = u_1\hat{\imath} + u_2\hat{\jmath} + u_3\hat{k}$ and $\vec{v} = v_1\hat{\imath} + v_2\hat{\jmath} + v_3\hat{k}$ in $\mathbb{R}^3$ is the vector

$$\vec{u} \times \vec{v} = (u_2v_3 - u_3v_2)\hat{\imath} - (u_1v_3 - u_3v_1)\hat{\jmath} + (u_1v_2 - u_2v_1)\hat{k}.$$

- Geometrically, the cross product is

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

where θ is the angle between $\vec{u}$ and $\vec{v}$ and $\vec{n}$ is a unit vector perpendicular to both $\vec{u}$ and $\vec{v}$ as determined by the right-hand rule.

- The cross product of vectors $\vec{u}$ and $\vec{v}$ is a vector perpendicular to both $\vec{u}$ and $\vec{v}$.
- The magnitude $\|\vec{u} \times \vec{v}\|$ of the cross product of the vectors $\vec{u}$ and $\vec{v}$ gives the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$. Also, the scalar triple product $|(\vec{u} \times \vec{v}) \cdot \vec{w}|$ gives the volume of the parallelepiped determined by $\vec{u}$, $\vec{v}$, and $\vec{w}$.

9.5.7 Exercises

1. If $\mathbf{a} = \mathbf{i} + \mathbf{j} + 4\mathbf{k}$ and $\mathbf{b} = \mathbf{i} + \mathbf{j} + 4\mathbf{k}$
Compute the cross product $\mathbf{a} \times \mathbf{b}$.
$\mathbf{a} \times \mathbf{b}$ =_____ $\mathbf{i}$ + _____ $\mathbf{j}$ + _____ $\mathbf{k}$

2. Suppose $\vec{v} \cdot \vec{w} = 7$ and $||\vec{v} \times \vec{w}|| = 3$, and the angle between $\vec{v}$ and $\vec{w}$ is θ. Find
(a) $\tan\theta$ = _____
(b) θ = _____

3. You are looking down at a map. A vector $\mathbf{u}$ with $|\mathbf{u}| = 8$ points north and a vector $\mathbf{v}$ with $|\mathbf{v}| = 1$ points northeast. The crossproduct $\mathbf{u} \times \mathbf{v}$ points:
A) south
B) northwest
C) up
D) down
Please enter the letter of the correct answer: _____
The magnitude $|\mathbf{u} \times \mathbf{v}|$ = _____

4. If $\mathbf{a} = \mathbf{i} + 9\mathbf{j} + \mathbf{k}$ and $\mathbf{b} = \mathbf{i} + 16\mathbf{j} + \mathbf{k}$, find a unit vector with positive first coordinate orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.
_____ $\mathbf{i}$ + _____ $\mathbf{j}$ + _____ $\mathbf{k}$

5. Sketch the triangle with vertices $O, P = (0, 7, 4)$ and $Q = (3, 0, 3)$ and compute its area using cross products.
Area=_____

6. Let $A = (-5, 5, 2)$, $B = (5, -3, -3)$, and $P = (k, k, k)$. The vector from A to B is perpendicular to the vector from A to P when k = _____.

7. Find two unit vectors orthogonal to $\mathbf{a} = \langle -3, 2, -4\rangle$ and $\mathbf{b} = \langle 2, -4, 3\rangle$
Enter your answer so that the first non-zero coordinate of the first vector is positive.
First Vector: $\langle$_____, _____, _____$\rangle$
Second Vector: $\langle$_____, _____, _____$\rangle$

8. Use the geometric definition of the cross product and the properties of the cross product to make the following calculations.
(a) $((\vec{i} + \vec{j}) \times \vec{i}) \times \vec{j}$ = _____
(b) $(\vec{j} + \vec{k}) \times (\vec{j} \times \vec{k})$ = _____
(c) $3\vec{i} \times (\vec{i} + \vec{j})$ = _____
(d) $(\vec{k} + \vec{j}) \times (\vec{k} - \vec{j})$ = _____

9. Are the following statements true or false?

(a) The value of $\vec{v} \cdot (\vec{v} \times \vec{w})$ is always zero.

(b) If $\vec{v}$ and $\vec{w}$ are any two vectors, then $||\vec{v} + \vec{w}|| = ||\vec{v}|| + ||\vec{w}||$.

(c) For any scalar c and any vector $\vec{v}$, we have $||c\vec{v}|| = c||\vec{v}||$.

(d) $(\vec{i} \times \vec{j}) \cdot \vec{k} = \vec{i} \cdot (\vec{j} \times \vec{k})$.

10. A bicycle pedal is pushed straight downwards by a foot with a 31 Newton force. The shaft of the pedal is 20 cm long. If the shaft is $\pi/5$ radians past horizontal, what is the magnitude of the torque about the point where the shaft is attached to the bicycle? _____ Nm

11. Find the cross product $\mathbf{c} \times \mathbf{d}$ where $\mathbf{c} = 2\boldsymbol{i} - 3\boldsymbol{j} + 3\boldsymbol{k}$ and $\mathbf{d} = \boldsymbol{i} - 4\boldsymbol{j} + 2\boldsymbol{k}$.
$\mathbf{c} \times \mathbf{d} =$ _____

12. Find the cross product $(2\mathbf{i} - \mathbf{j} + \mathbf{k}) \times (3\mathbf{i} - \mathbf{j} - \mathbf{k})$.
Answer: _____

13. Parallelogram $ABCD$ has vertices $A(4, 1, 1)$, $B(5, -4, 2)$, and $C(4, 2, 2)$.
Determine the coordinates of point D: _____
Determine the area of parallelogram $ABCD$
Area = _____

14. Find the area of the parallelogram with vertices $A(-1, 4, 3)$, $B(0, 6, 5)$, $C(1, 3, 5)$, and $D(2, 5, 7)$.
Area: _____

15. You are looking down at a map. A vector $\mathbf{u}$ with $|\mathbf{u}| = 1$ points north and a vector $\mathbf{v}$ with $|\mathbf{v}| = 9$ points northeast. The crossproduct $\mathbf{u} \times \mathbf{v}$ points:
A) south
B) northwest
C) up
D) down
Please enter the letter of the correct answer: _____
The magnitude $|\mathbf{u} \times \mathbf{v}| =$ _____

16. Find volume of the parallelepiped determined by the vectors $\mathbf{a} = \langle 5, 1, 0 \rangle$, $\mathbf{b} = \langle 1, 5, 1 \rangle$, and $\mathbf{c} = \langle 1, 1, 10 \rangle$.
Volume: _____

17. Calculate the 3x3 determinant:
$$\begin{vmatrix} -3 & -6 & -4 \\ 7 & -8 & 3 \\ -4 & -4 & 7 \end{vmatrix} =$$ _____

18. Let $\vec{u} = 2\hat{\imath} + \hat{\jmath}$ and $\vec{v} = \hat{\imath} + 2\hat{\jmath}$ be vectors in $\mathbb{R}^3$.

(a) Without doing any computations, find a unit vector that is orthogonal to both $\vec{u}$ and $\vec{v}$. What does this tell you about the formula for $\vec{u} \times \vec{v}$?

(b) Using the properties of the cross product and what you know about cross products involving the fundamental vectors $\hat{\imath}$ and $\hat{\jmath}$, compute $\vec{u} \times \vec{v}$.

(c) Next, use the determinant version of Equation (9.5.1) to compute $\vec{u} \times \vec{v}$. Write one sentence that compares your results in (a), (b), and (c).

(d) Find the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$.

19. Let $\vec{x} = \langle 1, 1, 1 \rangle$ and $\vec{y} = \langle 0, 3, -2 \rangle$.

(a) Are $\vec{x}$ and $\vec{y}$ orthogonal? Are $\vec{x}$ and $\vec{y}$ parallel? Clearly explain how you know, using appropriate vector products.

(b) Find a unit vector that is orthogonal to both $\vec{x}$ and $\vec{y}$.

(c) Express $\vec{y}$ as the sum of two vectors: one parallel to $\vec{x}$, the other orthogonal to $\vec{x}$.

(d) Determine the area of the parallelogram formed by $\vec{x}$ and $\vec{y}$.

20. Consider the triangle in $\mathbb{R}^3$ formed by $P = (3, 2, -1)$, $Q = (1, -2, 4)$, and $R = (4, 4, 0)$.

(a) Find $\overrightarrow{PQ}$ and $\overrightarrow{PR}$.

(b) Observe that the area of $\triangle PQR$ is half of the area of the parallelogram formed by $\overrightarrow{PQ}$ and $\overrightarrow{PR}$. Hence find the area of $\triangle PQR$.

(c) Find a unit vector that is orthogonal to the plane that contains points P, Q, and R.

(d) Determine the measure of $\angle QPR$.

21. One of the properties of the cross product is that $(\vec{u} + \vec{v}) \times \vec{w} = (\vec{u} \times \vec{w}) + (\vec{v} \times \vec{w})$. That is, the cross product distributes over vector addition *on the right*. Here we investigate whether the cross product distributes over vector addition *on the left*.

(a) Let $\vec{u} = \langle 1, 2, -1 \rangle$, $\vec{v} = \langle 4, -3, 6 \rangle$, and $\vec{w} = \langle 4, 7, 2 \rangle$. Calculate

$$\vec{u} \times (\vec{v} + \vec{w}) \quad \text{and} \quad (\vec{u} \times \vec{v}) + (\vec{u} \times \vec{w}).$$

What do you notice?

(b) Use the properties of the cross product to show that in general

$$\vec{x} \times (\vec{y} + \vec{z}) = (\vec{x} \times \vec{y}) + (\vec{x} \times \vec{z})$$

for any vectors $\vec{x}$, $\vec{y}$, and $\vec{z}$ in $\mathbb{R}^3$.

22. Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$, $\vec{v} = \langle v_1, v_2, v_3 \rangle$, and $\vec{w} = \langle w_1, w_2, w_3 \rangle$ be vectors in $\mathbb{R}^3$. In this exercise we investigate properties of the triple scalar product $(\vec{u} \times \vec{v}) \cdot \vec{w}$.

(a) Show that $(\vec{u} \times \vec{v}) \cdot \vec{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$.

(b) Show that $\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = - \begin{vmatrix} v_1 & v_2 & v_3 \\ u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$. Conclude that interchanging the first two rows in a 3×3 matrix changes the sign of the determinant. In general (although we won't show it here), interchanging any two rows in a 3×3 matrix changes the sign of the determinant.

(c) Use the results of parts (a) and (b) to argue that

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = (\vec{w} \times \vec{u}) \cdot \vec{v} = (\vec{v} \times \vec{w}) \cdot \vec{u}.$$

(d) Now suppose that $\vec{u}$, $\vec{v}$, and $\vec{w}$ do not lie in a plane when they emanate from a common initial point.

i. Given that the parallepiped determined by $\vec{u}$, $\vec{v}$, and $\vec{w}$ must have positive volume, what can we say about $(\vec{u} \times \vec{v}) \cdot \vec{w}$?

ii. Now suppose that $\vec{u}$, $\vec{v}$, and $\vec{w}$ all lie in the same plane. What value must $(\vec{u} \times \vec{v}) \cdot \vec{w}$ have? Why?

iii. Explain how (i.) and (ii.) show that if $\vec{u}$, $\vec{v}$, and $\vec{w}$ all eminate from the same initial point, then $\vec{u}$, $\vec{v}$, and $\vec{w}$ lie in the same plane if and only if $(\vec{u} \times \vec{v}) \cdot \vec{w} = 0$. This provides a straightforward computational method for determining when three vectors are co-planar.

23. In Subsection 9.5.3, we stated "Because the cross product works well with linear combinations, $\vec{u} \times \vec{v}$ is perpendicular to any linear combination of $\vec{u}$ and $\vec{v}$." We want to verify this algebraically here. Specifically, show that $\vec{u} \times \vec{v}$ will be orthogonal to any vector of the form $a\vec{u} + b\vec{v}$.

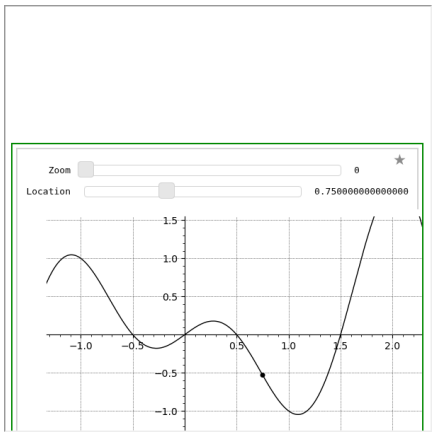
9.6 Lines in Space

Motivating Questions

- How are lines in $\mathbb{R}^3$ similar to and different from lines in $\mathbb{R}^2$?
- How do vectors allow us to describe lines in $\mathbb{R}^3$?

9.6.1 Introduction

In single variable calculus, we learn that a differentiable function is *locally linear*. In other words, if we zoom in on the graph of a differentiable function around a point, the graph will look like the tangent line to the function at that point. In Figure 9.6.1 you can look at the graph of a curve and zoom in around a point on that curve to see that eventually the graph of the function will look linear. You can use the sliders at the top of change how zoomed in the plot is displayed or change the location on the curve that is being highlighted. You should visually verify at a several points that this function is differentiable because at any point shown, the graph is locally linear.



Standalone
In Context
Embed

Figure 9.6.1 A plot with the ability to zoom in on a point of $y = f(x)$, where f is a nonlinear, differentiable function

Recall that linear functions have the form $f(x) = ax + b$. These functions play important roles in single variable calculus and are useful in approximating differentiable functions, in approximating roots of functions (Newton's

Method), and approximating solutions to first-order differential equations (Euler's Method). In multivariable calculus, we will study curves in space and we will see that connection between local linearity and differentiability exists for curves in three (or more) dimensions. As we extend our study of functions to include those of two (or more!) variables in Chapter 11, we will see that a function is locally linear at a point if the surface defined by the function looks like a plane (the tangent plane) as we zoom in on the graph.

Because of their important role in understanding curves and surfaces in $\mathbb{R}^3$, it is important for us to understand both lines and planes in space, as can both be thought of as generalizations of lines in two dimensions to three-dimensional space. In your study of one-variable functions, you probably worked with slope, direction, and measuring change using lines several times before you generalized those ideas for a curve in general. Similarly, we will want to understand how we talk about direction and measuring change along flat objects in $\mathbb{R}^3$ (lines and planes) before we start generalizing to curved objects. In this section, we discuss how we can use vectors to describe lines in space. The next section will be dedicated to understanding planes.

Preview Activity 9.6.1 We will start our work on lines by considering some familiar ideas in $\mathbb{R}^2$ but from a new perspective. You are probably familiar with equations of lines in the xy-plane in the form $y = mx + b$, where m is the slope of the line and $(0, b)$ is the y-intercept. In this activity, we explore a more flexible way of representing lines that is useful in the xy-plane and higher dimensions. To begin, consider the line through the point $(2, -1)$ with slope $\frac{2}{3}$ as shown in Figure 9.6.2.

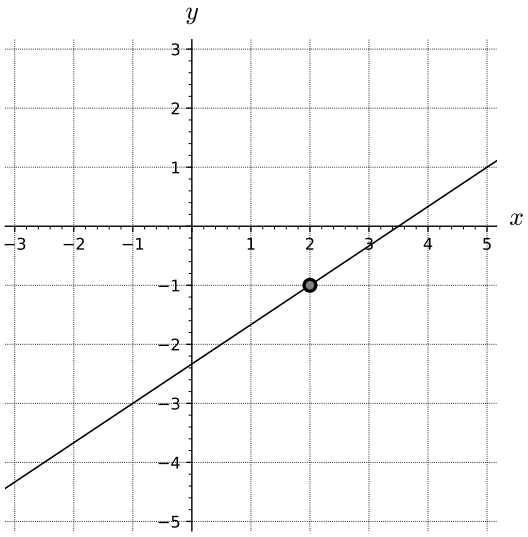


Figure 9.6.2 The line through $(2, -1)$ with slope $\frac{2}{3}$

(a) Suppose we increase x by 1 from the point $(2, -1)$. How does the y-value of a point on the line change? What is the point on the line with x-coordinate 3?

(b) Suppose we decrease x by 3.25 from the point $(2, -1)$. How does the y-value change? What is the point on the line with x-coordinate -1.25?

(c) Now, suppose we increase x by some arbitrary value $3t$ from the point $(2, -1)$. How does the y-value change? What is the point on the line with x-coordinate $2 + 3t$?

(d) Remember that the horizontal component of a vector describes the "run" and the vertical component measures the "rise" of the vector. So the slope of the line is related to *any* vector whose y-component divided by the x-component is the slope of the line. For the line in this activity, we might use the vector $\langle 3, 2\rangle$ to describe the direction of the line. We will look at the following vector-valued function of one variable:

$$\vec{r}(t) = \langle 2, -1\rangle + \langle 3, 2\rangle t,$$

For each of the values of t below, find $\vec{r}(t)$ and write your result as a single vector (of the form $\langle a, b\rangle$)

(a) $t = 0$

(b) $t = 1$

(c) $t = 2$

(d) $t = -2$

(e) $t = -\frac{1}{2}$

(f) $t = \frac{7}{3}$

(e) Draw the six vectors from the previous step in standard position (with initial point at the origin) on the plot below.

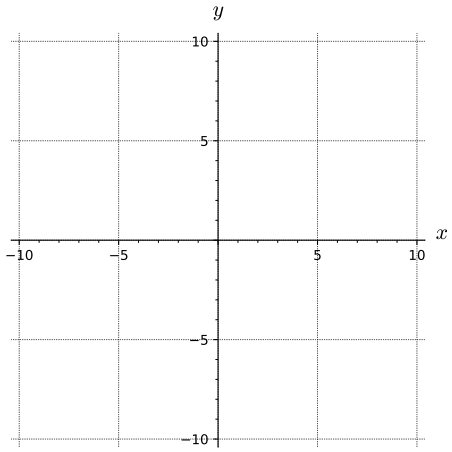


Figure 9.6.3 A blank 2D set of axes

(f) Write a few sentences that compare the endpoints of your vectors from the previous task to the plot of the line through the point $(2, -1)$ with a slope of $\frac{2}{3}$ (Figure 9.6.2). In particular you should address what aspects of the line are related to $\langle 2, -1\rangle$ or $\langle 3, 2\rangle$.

9.6.2 Lines in Space

The way most people draw a line is more connected to the ideas in Preview Activity 9.6.1 than slope or algebraic forms. Take a minute and draw a plot of the line described by $x + y - 1 = 0$. You should not continue reading until you have made a plot (by hand) of $x + y - 1 = 0$ that you would be proud to share with classmates.

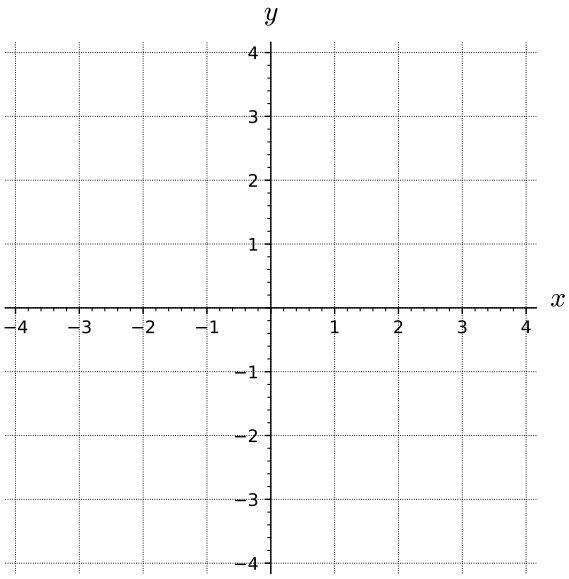


Figure 9.6.4 A blank 2D set of axes

You may have done some algebra or other calculation to get to slope-intercept form or another familiar algebraic structure, but when you went to draw the line you likely did the following steps:

1. Found a point to start your plot (maybe the y-intercept).
2. Found a second point on your plot (perhaps using the slope to find the second point).
3. Drew a line segment from the first point to the second, then extended the line past the second point to the edge of the plot.
4. Returned to the first point and extended the line in the direction opposite of the second point.

In other words, you made a starting point, extended indefinitely far in a particular direction, then added the line in the opposite direction. This is the essence of the general description for a line we will give below.

In two-dimensional space, a non-vertical line is defined to be the set of points satisfying the equation

$$y = mx + b,$$

for some constants m and b. The value of m (the slope) tells us how the dependent variable changes for every one unit increase in the independent variable, while the point $(0, b)$ is the y-intercept and anchors the line to a location on the y-axis. Alternatively, we can think of the slope as being related to the vector $\langle 1, m \rangle$, which tells us the direction of the line, as shown on the left in Figure 9.6.5. Thus, we can identify a line in space by fixing a point P and a direction $\vec{v}$, as shown on the right. Since we also have vectors in $\mathbb{R}^n$ to provide direction, this same idea of a point and a direction determining a line works in $\mathbb{R}^n$ for any n.

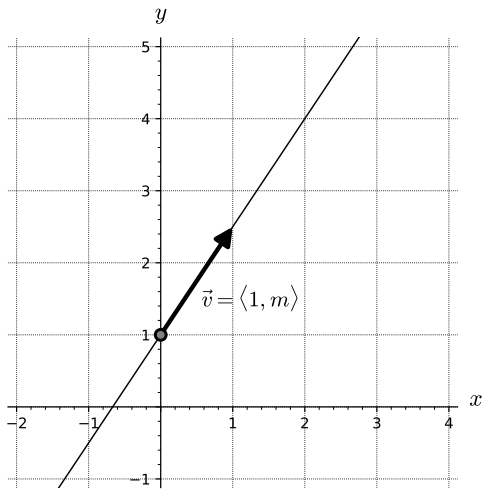


Figure 9.6.5 The line through $(0, 1)$ with slope m

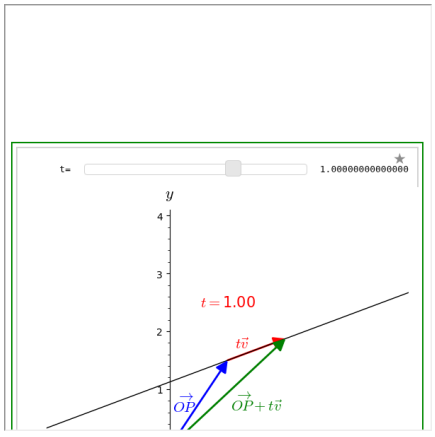
Definition 9.6.6 A **line** in space is the set of terminal points of vectors starting at a given point P that are parallel to a fixed vector $\vec{v}$. ◆

The vector $\vec{v}$ in Definition 9.6.6 is called a **direction vector** for the line. As we saw in Preview Activity 9.6.1, any vector parallel to $\vec{v}$ will have the form $t\vec{v}$ for some scalar t. As a result, any vector emanating from the point P in a direction parallel to the vector $\vec{v}$ will be of the form

$$\overrightarrow{OP} + \vec{v}t \tag{9.6.1}$$

for some scalar t. (Here we use O to denote the origin.)

Figure 9.6.7 shows the plot of a line in two-space. The vector $\overrightarrow{OP}$ is illustrated in blue and points from the origin to our starting point P. The direction vector $t\vec{v}$ as in Equation (9.6.1) is shown in red with its initial point at point P. The line is traced out by the terminal points of the green vector $\overrightarrow{OP} + t\vec{v}$. In other words, the tips (terminal points) of the green vectors trace out the line as t changes. You can use the slider to change the value of t for this example line. You should pay attention to how $\overrightarrow{OP} + \vec{v}t$ changes as you vary t. Be sure to see what happens with $t = 0$ as well as for some negative values of t.



Standalone
In Context
Embed

Figure 9.6.7 A line in 2-space with vector definition

To summarize, the terminal points of the vectors of the form in (9.6.1) define a linear function $\vec{r}$ in space of the following form, which is valid in any dimension and gives some geometric insight into lines in higher dimensions.

The vector form of a line.

The **vector form** of a line through the point P in the direction of the vector $\vec{v}$ is

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}, \tag{9.6.2}$$

where $\vec{r}_0$ is the position vector $\overrightarrow{OP}$ from the origin to the point P.

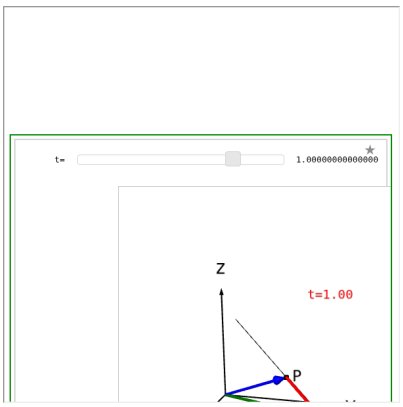
Example 9.6.8

(a) Lines in the xy-plane are commonly described using the slope-intercept form $y = mx + b$. The vector form of the line, as described above, is an alternative way to represent lines that has the two advantages. First, in two dimensions, we are able to represent vertical lines, whose slope m is not defined, using a vertical direction vector. For example, the vertical line described by $x = 2$ could be described in vector form with a direction vector such as $\vec{v} = \langle 0, 1\rangle$, giving $\vec{r}(t) = \langle 2, 0\rangle + t\langle 0, 1\rangle$. The second advantage of the vector form of a line is that this description of lines works in *any* dimension whereas the concept of slope of a line does not generalize to three (or more) dimensions.

(b) We will now find a vector form of the line that is the y-axis in $\mathbb{R}^3$. We need to pick a point on our line and a direction vector for our line. A convenient point would be the origin. We can use $\hat{j} = \langle 0, 1, 0\rangle$ as our direction vector. This gives a vector form of $\vec{r}_1(t) = \langle 0, 0, 0\rangle + t\langle 0, 1, 0\rangle$ which can be combined to a single vector form of $\vec{r}_1(t) = \langle 0, t, 0\rangle$. Since the y-axis consists of points with x- and z-coordinates of zero while the y-coordinate can be any real number, this form should be consistent with your intuition.

We had many other choices we could have made for the initial point and the direction vector used to create the vector form of the y-axis. We could have selected the point $(0, \pi, 0)$ and the direction vector given by $\langle 0, -\frac{11}{7}, 0\rangle$. While these are likely not choices you would have made, this point and direction vector still describe the necessary information to give the vector form of the y-axis. The associated vector form of the y-axis would be $\vec{r}_2(t) = \langle 0, \pi, 0\rangle + t\langle 0, -\frac{11}{7}, 0\rangle = \langle 0, \pi - \frac{11}{7}t, 0\rangle$. It is unlikely that anyone else has ever described the y-axis with this equation but the terminal points of the output vectors for $\vec{r}_2(t)$ do trace out the entire y-axis. In the next chapter, we will explore how the context in which we wish to arrive at the vector form of a line may influence the choice of initial point and direction vector.

(c) To look at a line that is less familiar, we now consider the line $\mathcal{L}_1$ in $\mathbb{R}^3$ passing through $(1, 2, 1)$ and in the direction $\langle 1, 1, -\frac{1}{2}\rangle$. The corresponding vector form is $\vec{r}(t) = \langle 1, 2, 1\rangle + \langle 1, 1, -\frac{1}{2}\rangle t$. Figure 9.6.9 shows a plot of this line with a slider that you can use to change the value for t and examine how the different aspects of this vector form look. Notice that when $t = 1$, we have $\vec{r}(1) = \langle 2, 3, 1/2\rangle$, so the point $P = (2, 3, 1/2)$ lies on $\mathcal{L}_1$.



Standalone
In Context
Embed

Figure 9.6.9 A line in 2-space with vector definition

(d) You can verify visually using Figure 9.6.9 that the line $\mathcal{L}_1$ from the previous part does not intersect the y-axis. In two dimensions, two lines were parallel if they did not intersect. However, this idea does not translate to higher dimensions because it does not seem appropriate to say that $\mathcal{L}_1$ is parallel to the y-axis. In general, two lines in $\mathbb{R}^n$ are **parallel** if they can be described using parallel direction vectors. (Recall that two vectors are parallel if they are non-zero scalar multiples of each other.) This definition is based in the perspective that parallel lines should move in the same direction, but that direction can be described by *any* parallel vector.

▲

We formalize the discussion in the last part of the previous activity by giving names to the two nonintersecting configurations that lines in $\mathbb{R}^n$ for $n \geq 3$ can have.

Definition 9.6.10 Two lines in $\mathbb{R}^n$ are said to be **parallel** provided that their direction vectors are parallel. Two lines that do not intersect and whose direction vectors are not parallel are said to be **skew**. ◆

Activity 9.6.2 Let $P_1 = (1, 2, -1)$ and $P_2 = (-2, 1, -2)$ and let $\mathcal{L}$ be the line in $\mathbb{R}^3$ through P_1 and P_2. Note that Figure 9.6.9 shows a similar example of a line in 3D defined by two points.

(a) Find a direction vector for the line $\mathcal{L}$.

(b) Find a vector equation of $\mathcal{L}$ in the form $\vec{r}(t) = \vec{r}_0 + t\vec{v}$.

(c) Consider the vector equation $\vec{s}(t) = \langle -5, 0, -3\rangle + t\langle 6, 2, 2\rangle$. What is the direction of the line given by $\vec{s}(t)$? Is this new line parallel to line $\mathcal{L}$?

(d) Do $\vec{r}(t)$ and $\vec{s}(t)$ represent the same line, $\mathcal{L}$? Write a couple of sentences to justify why you think $\vec{r}(t)$ and $\vec{s}(t)$ do or do not describe the same set of points.

9.6.3 The Parametric Equations of a Line

The vector form of a line, $\vec{r}(t) = \vec{r}_0 + t\vec{v}$ in Equation (9.6.2), describes a line as the set of terminal points of the vectors $\vec{r}(t)$. If $\vec{r}_0 = \langle x_0, y_0, z_0\rangle$ and $\vec{v} = \langle a, b, c\rangle$, then we can view this vector form in terms of components and get the following:

$$\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$$

$$= \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle$$
$$= \langle x_0 + at, y_0 + bt, z_0 + ct \rangle$$

Splitting this vector form into each component equation gives

$$x(t) = x_0 + at \qquad y(t) = y_0 + bt \qquad z(t) = z_0 + ct$$

These equations describe the coordinates of the points on the line separately where each is coordinate is a function of t. The variable t represents an arbitrary scalar and is called a **parameter**. In particular, we use the following language.

The parametric equations of a line.

The **parametric equations** for a line through the point $P = (x_0, y_0, z_0)$ in the direction of the vector $\vec{v} = \langle a, b, c \rangle$ are

$$x(t) = x_0 + at \qquad y(t) = y_0 + bt \qquad z(t) = z_0 + ct$$

Notice that there are *many* different parametric equations for the same line. For example, choosing another point P on the line or another (parallel) direction vector $\vec{v}$ produces another set of parametric equations. In many physical applications, it is useful to think of t as a time parameter and the parametric equations as telling us where we are on the line at each time. In this way, the parametric equations describe a particular way to walk along the line. There are, of course, many possible ways to walk along the same line, including in opposite directions and at different speeds. In mathematical terms, we say that the parameterization of a line is not unique.

Activity 9.6.3 Let $P_1 = (1, 2, -1)$ and $P_2 = (-2, 1, -2)$, and let $\mathcal{L}$ be the line in $\mathbb{R}^3$ through P_1 and P_2. (This is the same line as in Activity 9.6.2.)

(a) Find parametric equations of the line $\mathcal{L}$.

(b) Does the point $(1, 2, 1)$ lie on $\mathcal{L}$? If so, what value of t results in this point?

(c) Consider another line, $\mathcal{K}$, whose parametric equations are

$$x(s) = 11 + 4s, \quad y(s) = 1 - 3s, \quad z(s) = 3 + 2s.$$

What is the direction of the line $\mathcal{K}$?

(d) Do the lines $\mathcal{L}$ and $\mathcal{K}$ intersect? If so, provide the point of intersection and the t and s values, respectively, that result in the point. If not, explain why. To find a point of intersection, you can set the coordinate equations of each line equal to each other try to solve for t and s.

Hint. Remember that the two lines need to go through the same (x, y, z) point but do not need to have the same parameter value at that point (which is why we used *different* variable names for the parameters t and s).

Before we move on to our discussion of planes in the next section, we will talk about some terminology that may helpful in describing features and measurements later. A line is a one-dimensional, flat graph in three (or n) dimensions. We say the line is a one-dimensional graph because there is only one direction (forward/backward) to go while staying on the graph. We say

that a line is flat because the direction you are allowed to go doesn't change as you move along the graph. Note that flat does *not* mean horizontal. A circle would be another example of a one-dimensional graph that can be drawn in two or more dimensional space (because there is only one direction to move along the circle). However, a circle is not flat since the direction you go to stay on the figure changes as you move along the graph. As we will see in the next section, a plane in three dimensional space is an example of a flat, two-dimensional graph. If you were an ant on a plane, you would have two directions (forward/backward or left/right) to move in while staying on the graph. However, the directions that the graph extends do not change when you change locations on the plane.

Summary

- While lines in $\mathbb{R}^3$ do not have a slope, like lines in $\mathbb{R}^2$ they can be characterized by a point and a direction vector. Indeed, we define a line in space to be the set of terminal points of vectors emanating from a given point that are parallel to a fixed vector.
- Vectors play a critical role in representing the equation of a line. In particular, the terminal points of the vector $\vec{r}(t) = \vec{r}_0 + t\vec{v}$ define a linear function $\vec{r}$ in space through the terminal point of the vector $\vec{r}_0$ in the direction of the vector $\vec{v}$, tracing out a line in space.

9.6.4 Exercises

1. Rewrite the vector equation $\mathbf{r}(t) = (-1+t)\mathbf{i} + (3-5t)\mathbf{j} + (2-3t)\mathbf{k}$ as the corresponding parametric equations for the line.

$x(t) =$ _____
$y(t) =$ _____
$z(t) =$ _____

2. Find the vector and parametric equations for the line through the point P(0, 0, -4) and parallel to the vector $-2\mathbf{i} - 5\mathbf{j} + 2\mathbf{k}$.

Vector Form: $\mathbf{r} =$ ⟨_____, _____, -4 ⟩+t⟨ _____, _____, 2 ⟩

Parametric form (parameter t, and passing through P when t = 0):

$x = x(t) =$_____
$y = y(t) =$_____
$z = z(t) =$_____

3. Consider the line which passes through the point P(2, -4, -5), and which is parallel to the line $x = 1 + 1t, y = 2 + 2t, z = 3 + 4t$

Find the point of intersection of this new line with each of the coordinate planes:

xy-plane: (_____,_____, _____)
xz-plane: (_____,_____, _____)
yz-plane: (_____,_____, _____)

4. Find the point at which the line $\langle 3, 3, 5\rangle + t\langle 3, 5, 2\rangle$ intersects the plane $x + 5y - 4z = 178$.

(_____, _____, _____)

5. Point A is located at the coordinates $A = 0\hat{i} + 4\hat{j} - 1\hat{k}$. Point B is located at $B = -2\hat{i} + 0\hat{j} + 0\hat{k}$. A rope is stretched between points A and B with a tension force of $99\,N$. What are the directional components of the tension force as seen from point A? *Go to three significant digits*

$\mathbf{F}$ =_____$\hat{i}$+_____$\hat{j}$+_____$\hat{k}$ N

6. Find the vector and parametric equations for the line through the point $P = (5, -5, 3)$ and parallel to the vector $\langle -3, 5, 2 \rangle$.

Vector Form: $\mathbf{r} = \langle$_____, _____, 3 $\rangle + t\langle$ _____, _____, 2 $\rangle$

Parametric form (parameter t, and passing through P when $t = 0$):

$x = x(t) =$ _____

$y = y(t) =$ _____

$z = z(t) =$ _____

7. Find the vector and parametric equations for the line through the point $P = (-5, 3, -3)$ and the point $Q = (0, 8, -6)$.

Vector Form: $\mathbf{r} = \langle$_____, _____, -3 $\rangle + t\langle$ _____, _____, -3 $\rangle$

Parametric form (parameter t, and passing through P when $t = 0$):

$x = x(t) =$ _____

$y = y(t) =$ _____

$z = z(t) =$ _____

8. Find the simplest vector parametric expression $\vec{r}(t)$ for the line that passes through the points $P = (-3, -1, 1)$ at time t = 4 and $Q = (-1, 4, 3)$ at time t = 9.

$\vec{r}(t) =$ _____

9. Parameterize the line through $P = (-1, -5)$ and $Q = (0, -1)$ so that the points P and Q correspond to the parameter values $t = 15$ and 16.

$\vec{r}(t) =$ _____

10. ***(a)*** Find a parametric equation for the line through the point $(3, 5, 3)$ and in the direction of $a\vec{i} + b\vec{j} + c\vec{k}$.

$\vec{r}(t) =$ _____

(b) Find conditions on a, b, c so that the line you found in part (a) goes through the origin. (Be sure you can give a reason for your answer.) Then use your work to give two distinct triples a, b, c that result in the line passing through the origin.

Choice 1: $a =$ _____, $b =$ _____, $c =$ _____;

Choice 2: $a =$ _____, $b =$ _____, $c =$ _____.

11. Given a the vector equation $\mathbf{r}(t) = (0+4t)\mathbf{i}+(-2+4t)\mathbf{j}+(4-4t)\mathbf{k}$, rewrite this in terms of the parametric equations for the line.

$x(t) =$ _____

$y(t) =$ _____

$z(t) =$ _____

12. Determine whether the following lines represented by the vector equations below *intersect*, are *parallel*, are *skew*, or are *identical*.

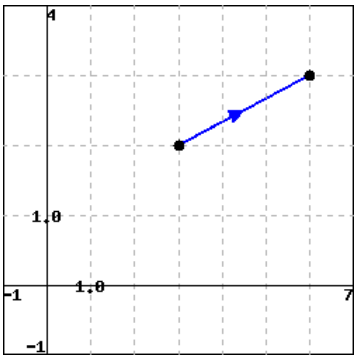
(☐ choose one ☐ Intersect ☐ Parallel ☐ Skew ☐ Identical) ***1.*** $\mathbf{r}(t) = \langle 1 - t, 3 + 2t, -3t \rangle$ and $\mathbf{s}(t) = \langle 3t, 2 - 6t, 3 + 9t \rangle$.

(☐ choose one ☐ Intersect ☐ Parallel ☐ Skew ☐ Identical) ***2.*** $\mathbf{r}(t) = \langle 2 - 3t, 5, 7 - 5t \rangle$ and $\mathbf{s}(t) = \langle 1 + 4t, 7, 2 - t \rangle$.

(☐ choose one ☐ Intersect ☐ Parallel ☐ Skew ☐ Identical) ***3.*** $\mathbf{r}(t) = \langle 4 - 2t, 4 + 5t, 1 + 2t \rangle$ and $\mathbf{s}(t) = \langle t, 5 - 3t, 23 \rangle$.

Note: *You only have two attempts at this problem.*

13. Find a parameterization for the curve shown in the figure below.



$\vec{r}(t) =$ _____,
where _____ $\leq t \leq$ _____
(Note: you need to enter answers in all three answer blanks to receive feedback.)

14. Two particles are traveling through space. At time t the first particle is at the point

$$(-2+3t, -4+t, -5+3t)$$

and the second particle is at

$$(-18-t, -20-3t, -21-t)$$

Describe the two paths in words. Then determine if the two particles collide, cross, or neither cross nor collide to fill in the following.

We see that (☐ the paths cross, and the particles collide there ☐ the paths of the particles cross, but with no collision ☐ neither of these is true) .

If they collide or their paths cross, indicate where (if neither is true, enter ***none***).

A collision or crossing occurs at $(x, y, z) =$ _____.

15. Determine whether the points lie on a straight line. Answer *Yes* or *No.*

(☐ Yes ☐ No) ***1.*** $A(0,0,0)$, $B(9,-4,3)$, and $C(-27,12,-8)$.
(☐ Yes ☐ No) ***2.*** $D(9,-4,3)$, $E(10,-2,6)$, and $F(14,6,18)$.
(☐ Yes ☐ No) ***3.*** $G(-1,0,1)$, $H(3,9,10)$, and $I(8,27,28)$.

Note: *You only have two attempts at this problem.*

16. Find the intersection of the lines $\mathbf{r}(t) = \langle -2-2t, 17+6t, 22+8t\rangle$ and $\mathbf{R}(s) = \langle 17+7s, -8-5s, 4+s\rangle$. ***Write your answer as a point*** (a,b,c) ***where*** a***,*** b***, and*** c ***are numbers.***

Answer: _____

Note: *If there is no intersection, write "none".*

17. Determine whether the following lines in $\mathbb{R}^2$ represented by the vector equations below *intersect*, are *parallel*, or are *identical.*

(☐ choose one ☐ Intersect ☐ Parallel ☐ Identical) ***1.*** $\mathbf{r}(t) = \langle -3-t, -3+2t\rangle$ and $\mathbf{s}(t) = \langle -3-3t, -3+t\rangle$.

(☐ choose one ☐ Intersect ☐ Parallel ☐ Identical) ***2.*** $\mathbf{r}(t) = \langle 2-3t, -3+2t\rangle$ and $\mathbf{s}(t) = \langle 2+6t, 7-4t\rangle$.

(☐ choose one ☐ Intersect ☐ Parallel ☐ Identical) ***3.*** $\mathbf{r}(t) = \langle 1-2t, 2+4t\rangle$ and $\mathbf{s}(t) = \langle -1-4t, 6+8t\rangle$.

Note: *You only have two attempts at this problem.*

18. Given a the vector equation $\mathbf{r}(t) = (-2+4t)\mathbf{i}+(4+3t)\mathbf{j}+(3+2t)\mathbf{k}$, rewrite this in terms of the parametric equations for the line.

$x(t) =$ _____
$y(t) =$ _____
$z(t) =$ _____

19. Find a vector parametric equation $\vec{r}(t)$ for the line through the points $P = (-1, 1)$ and $Q = (-4, 2)$.
$\vec{r}(t) =$ _____
Find a vector parametric equation $\vec{r}(t)$ for the line through the points $P = (3, -4, 5)$ and $Q = (-2, -2, 7)$.
$\vec{r}(t) =$ _____
Find a vector parametric equation $\vec{r}(t)$ for the line through the points $P = (-3, -2, 4)$ and $Q = (-7, -4, -1)$.
$\vec{r}(t) =$ _____

20. Determine whether the lines

$$L_1 : x = 9 + 3t, \quad y = 15 + 6t, \quad z = 16 + 5t$$

and

$$L_2 : x = -8 + 4t \quad y = -19 + 8t \quad z = -15 + 8t$$

intersect, are skew, or are parallel. If they intersect, determine the point of intersection; if not leave the remaining answer blanks empty.
Do/are the lines: (☐ parallel ☐ intersect ☐ skew)
Point of intersection: (_____, _____,_____)

21. Determine whether the lines

$$L_1 : \frac{x-4}{1} = \frac{y-22}{6} = \frac{z-1}{2}$$

and

$$L_2 : \frac{x-2}{-3} = \frac{y-6}{-18} = \frac{z+2}{-6}$$

intersect, are skew, or are parallel. If they intersect, determine the point of intersection; if not leave the remaining answer blanks empty.
Do/are the lines (intersect/skew/parallel):_____
Point of intersection: (_____, _____,_____)

22. Consider the two lines
$L_1 : \ x = -2t, \ y = 1 + 2t, \ z = 3t$
and
$L_2 : \ x = -8 + 4s, \ y = 2 + 3s, \ z = 3 + 3s$
Find the point of intersection of the two lines.
$P =$ (_____, _____, _____)

23.

(a) Give the parametric form of a line that goes through the points $(1, 2, 3)$ and $(4, 5, 6)$.

(b) Do the lines given by $\overrightarrow{r}_1(t) = \langle 0, 1, 1\rangle + t\langle 1, 2, 2\rangle$ and $\overrightarrow{r}_2(t) = \langle 2, 0, 3\rangle + t\langle 1, 4, 4\rangle$ intersect? If so, state at what point(s) they intersect.

(c) Give a parametric form of the y-axis.

(d) Give a parametric form of a line that does not intersect the xz-plane.

(e) Give a parametric form for the line that goes through points given by $x = 2y = 3z$.

Hint. Find two points that would satisfy these equations.

24. The vector and parametric forms of a line allow us to easily describe line segments in space.

Let $P_1 = (1, 2, -1)$ and $P_2 = (-2, 1, -2)$, and let $\mathcal{L}$ be the line in $\mathbb{R}^3$ through P_1 and P_2 as in Activity 9.6.2.

(a) What value of the parameter t makes $(x(t), y(t), z(t)) = P_1$? What value of t makes $(x(t), y(t), z(t)) = P_2$?

(b) What t values describe the line segment between the points P_1 and P_2?

(c) What about the line segment (along the same line) from $(7, 4, 1)$ to $(-8, -1, -4)$?

(d) Now, consider a segment that lies on a different line: parameterize the segment that connects point $R = (4, -2, 7)$ to $Q = (-11, 4, 27)$ in such a way that $t = 0$ corresponds to point Q, while $t = 2$ corresponds to R.

25. This exercise explores key relationships between a pair of lines. Consider the following two lines: one with parametric equations $x(s) = 4 - 2s$, $y(s) = -2 + s$, $z(s) = 1 + 3s$, and the other being the line through $(-4, 2, 17)$ in the direction $\vec{v} = \langle -2, 1, 5 \rangle$.

(a) Find a direction vector for the first line, which is given in parametric form.

(b) Find parametric equations for the second line, written in terms of the parameter t.

(c) Show that the two lines intersect at a single point by finding the values of s and t that result in the same point. Then find the point of intersection.

(d) Find the acute angle formed where the two lines intersect, noting that this angle will be given by the acute angle between their respective direction vectors.

9.7 Planes in Space

Motivating Questions

- How is a plane defined in terms of measurements of points and vectors?
- What different ways are there to determine a plane through geometric information?

9.7.1 Introduction

In Section 9.6, we saw how to describe a line in $\mathbb{R}^n$ by setting an initial point and allowing unrestricted movement in a direction given by a direction vector. In this section's Preview Activity, we will be looking at what happens when we allow only movement that is perpendicular to a given direction. Since we will specify the given direction by a vector, we will find the perpendicular directions using orthogonal vectors.

Preview Activity 9.7.1 We will consider what happens when we allow movement in $\mathbb{R}^3$ with the restriction that the movement must be orthogonal to $\vec{v} = \langle 1, 2, 3 \rangle$.

(a) Find values for a_0 and b_0 such that $\langle 0, a_0, b_0 \rangle$ is orthogonal to $\langle 1, 2, 3 \rangle$.

(b) Find values for c_0 and d_0 such that $\langle c_0, d_0, 0 \rangle$ is orthogonal to $\langle 1, 2, 3 \rangle$.

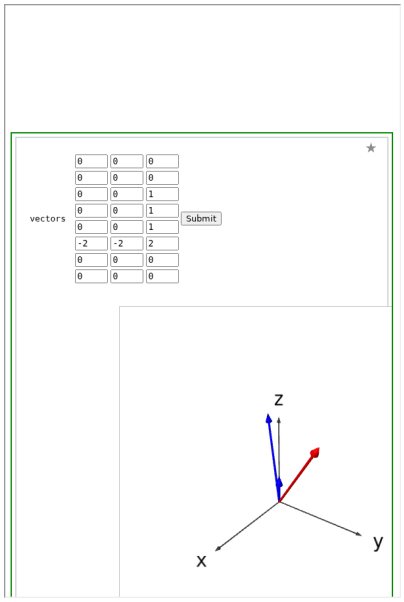
(c) Find values for c_1 and d_1 such that $\langle c_1, d_1, 1 \rangle$ is orthogonal to $\langle 1, 2, 3 \rangle$.

(d) Find two other values for each of c and d such that $\langle c, d, 1 \rangle$ is orthogonal to $\langle 1, 2, 3 \rangle$.

(e) Verify that each of the following vectors is also orthogonal to $\langle 1, 2, 3 \rangle$.

- $\langle -2, -2, 2 \rangle$
- $\langle c_0, a_0 + d_0, b_0 \rangle$ with your values from parts a and b
- $\langle -2c_1, a_0 - 2d_1, b_0 - 2 \rangle$ with your values from parts a and c

(f) Put all of the vectors you have computed into the interactive below to visually verify that each of them is orthogonal to $\langle 1, 2, 3 \rangle$. You should put the component values of each vector into this array with each vector corresponding to a row. If you do not have eight distinct vectors from the previous parts, multiply one of your repeated vectors by -1 and enter that instead of entering a vector multiple times.



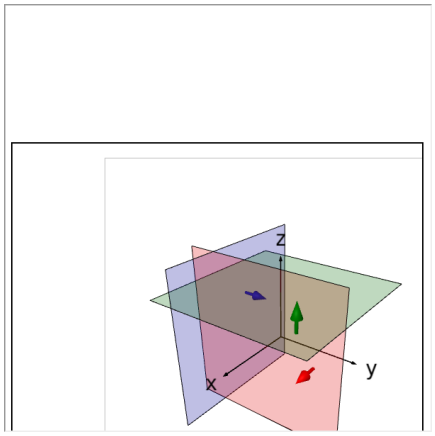
Standalone
In Context
Embed

Figure 9.7.1 A plot of your vectors that should be orthogonal to $\langle 1, 2, 3 \rangle$

(g) Describe what you think the plot of the set of vectors that are orthogonal to $\langle 1, 2, 3 \rangle$ will look like.

9.7.2 Planes in Space

Now that we have a way of describing lines, we would like to develop a means of describing planes in three dimensions. In Section 9.2, we studied the coordinate planes and planes parallel to them. In particular, $x = 1$, $y = -2$, and $z = \sqrt{3}$ are examples of fundamental planes. In general, the equations of fundamental planes have the form *coordinate* = *constant*.

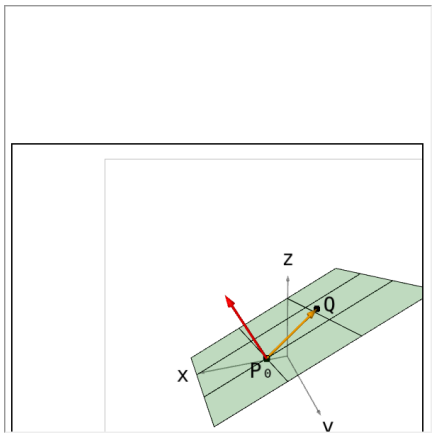


Standalone
In Context
Embed

Figure 9.7.2 A plot of $x = 1$ in red, $y = -2$ in blue, and $z = \sqrt{3}$ in green with a normal vector plotted for each fundamental plane

As shown in Figure 9.7.2, any vector in a plane with $x =$ constant will be orthogonal to the vector $\langle 1, 0, 0\rangle$, any vector in a plane with $y =$ constant will be orthogonal to the vector $\langle 0, 1, 0\rangle$, and any vector in a plane with $z =$ constant will be orthogonal to the vector $\langle 0, 0, 1\rangle$. We will use this idea to define a plane in general.

Definition 9.7.3 A **plane** p in space is the set of all terminal points of vectors emanating from a given point P_0 perpendicular to a fixed vector $\vec{n}$, as shown in Figure 9.7.4. We call $\vec{n}$ the **normal vector** of the plane. ◆



Standalone
In Context
Embed

Figure 9.7.4 A plot of a plane with normal vector $\vec{n}$ in red and vector on the plane $\overrightarrow{P_0Q}$ in orange

Like the definition of a line, the definition of a plane given above uses a starting point and a vector as the critical pieces of information. For a line, you begin at the starting point and move as much as you want parallel to the given vector (the direction vector). For a plane, you begin at the starting point and move as much as you want orthogonal to the given vector (the normal vector). For a line, you move only in the direction of the given vector whereas on a plane you cannot move at all in the direction of the given vector.

This description of a plane allows us to find the equation of a plane. Assume that $\vec{n} = \langle a, b, c\rangle$, $P_0 = (x_0, y_0, z_0)$, and that $Q = (x, y, z)$ is an arbitrary point on the plane. Since the vector $\overrightarrow{P_0Q}$ lies in the plane, it must be perpendicular to $\vec{n}$. This means that

$$0 = \vec{n} \cdot \overrightarrow{P_0Q}$$

$$\begin{aligned} &= \vec{n} \cdot \big(\langle x, y, z\rangle - \langle x_0, y_0, z_0\rangle\big) \\ &= \vec{n} \cdot \langle x - x_0, y - y_0, z - z_0\rangle \\ &= a(x - x_0) + b(y - y_0) + c(z - z_0). \end{aligned}$$

We may now summarize our new equation for a plane.

Equations of a plane.

- The **scalar equation** of the plane with normal vector $\vec{n} = \langle a, b, c\rangle$ containing the point $P_0 = (x_0, y_0, z_0)$ is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \tag{9.7.1}$$

- The **vector equation** of the plane with normal vector $\vec{n} = \langle a, b, c\rangle$ containing the points $P_0 = (x_0, y_0, z_0)$ and $Q = (x, y, z)$ is

$$\vec{n} \cdot \overrightarrow{P_0Q} = 0. \tag{9.7.2}$$

We may take the scalar equation of a plane a little further and note that since

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

it equivalently follows that

$$ax + by + cz = ax_0 + by_0 + cz_0.$$

That is, we may write an equation of a plane as $ax + by + cz = d$ where $d = \vec{n} \cdot \langle x_0, y_0, z_0\rangle$.

Before we look at some examples, we will mention a few ideas that relate planes (as a collection of points), equations for these planes, and normal vectors for these planes. As we saw with lines and will see for most other types of graphs, there are multiple different equations that can describe any given plane. This highlights the difference between an equation that describes a graph and the set of points that make up the graph or make the equation true. For instance, $x + y + z = 1$ and $2x + 2y + 2z = 2$ are different equations but describe the same graph. That is, the equations have the same set of (x, y, z) points that make the equation true. You may be tempted to say "Those are the same equations. They just differ by a scalar multiple!", but this difference is the same as the difference between parallel vectors.

The good news is that the orientation of a plane is encoded into the normal vector, thus we can use all the powerful tools we developed for vectors. We say that two planes are **parallel** if their normal vectors are parallel (as vectors). In fact, we can describe the angle between two planes using normal vectors. The angle between two planes is the acute angle between their respective normal vectors. While the angle between two vectors (as defined in Subsection 9.4.3) can be obtuse, the angle between planes will not be greater than $\frac{\pi}{2}$. When using normal vectors to describe the orientation of a plane, we need to consider all the vectors parallel to our choice of the normal vector because any vector parallel to our choice would give an equivalent equation for the plane. We will look at specifics of this in our next example.

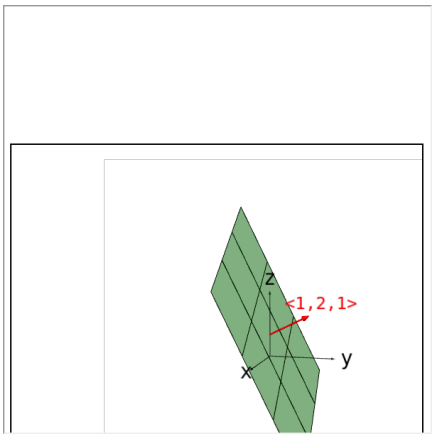
Example 9.7.5 If we would like to describe the plane passing through the point $P = (4, -2, 1)$ and perpendicular to the vector $\vec{n} = \langle 1, 2, 1\rangle$, we have

$$\langle 1, 2, 1\rangle \cdot \langle x, y, z\rangle = \langle 1, 2, 1\rangle \cdot \langle 4, -2, 1\rangle$$

or

$$x + 2y + z = 1.$$

Notice that the coefficients of x, y, and z in this description give a vector perpendicular to the plane.



Standalone
In Context
Embed

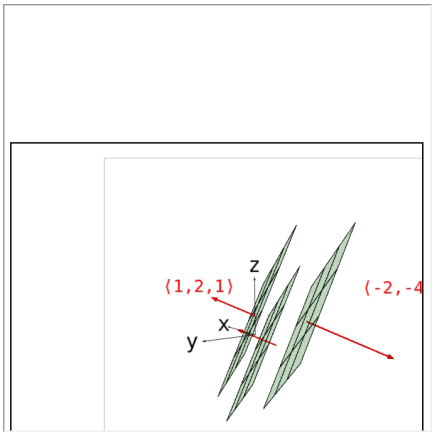
Figure 9.7.6 A plot of a plane with normal vector $\vec{n}$ in red

For instance, if we are presented with the plane

$$-2x + y - 3z = 4,$$

we know that $\vec{n} = \langle -2, 1, -3 \rangle$ is a vector perpendicular to the plane.

Using our definition of parallel planes, we can see that our plane given by $x+2y+z = 1$ will be parallel to planes given by $x+2y+z = -2$, $2x+4y+2z = 1$, and $-\frac{3}{2}x - 3y - \frac{3}{2}z = 12,345$ because the respective normal vectors ($\langle 1, 2, 1 \rangle$, $\langle 1, 2, 1 \rangle$, $\langle 2, 4, 2 \rangle$, and $\langle -\frac{3}{2}, -3, -\frac{3}{2} \rangle$) are all scalar multiples of each other.



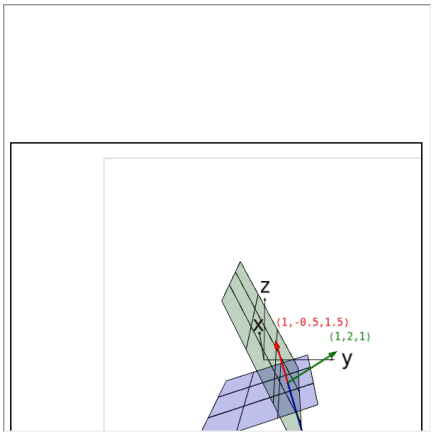
Standalone
In Context
Embed

Figure 9.7.7 A plot of three parallel planes with normal vectors

If we wanted to find the angle between $x+2y+z = 1$ and $-2x+y-3z = 4$, we would need to consider the angle between the normal vectors $\langle 1, 2, 1 \rangle$ and $\langle -2, 1, -3 \rangle$. The dot product of these normal vectors is $\langle 1, 2, 1 \rangle \cdot \langle -2, 1, -3 \rangle = -3$. This tells us that the angle between these vectors is obtuse. However, the (smallest) angle between the planes is *not* obtuse. The equation $x - \frac{1}{2}y + \frac{3}{2}z = -2$ will describe the same set of points as $-2x + y - 3z = 4$ but the normal vector for $x - \frac{1}{2}y + \frac{3}{2}z = -2$ will be $\langle 1, -\frac{1}{2}, \frac{3}{2} \rangle$. This vector makes an acute angle with $\langle 1, 2, 1 \rangle$. This allows us to conclude that the angle between $x + 2y + z = 1$

and $-2x + y - 3z = 4$ is

$$\arccos\left(\frac{\langle 1,2,1\rangle \cdot \langle 1, \frac{1}{2}, \frac{3}{2}\rangle}{\sqrt{7}\sqrt{\frac{7}{2}}}\right) \approx 1.2373 \approx 70.89°.$$



Standalone
In Context
Embed

Figure 9.7.8 A plot of planes $x + 2y + z = 1$ and $-2x + y - 3z = 4$ with normal vectors

Our next activity gives you an opportunity to practice with equations of planes and also to think about interactions between planes and lines. ▲

Activity 9.7.2

(a) Write a scalar equation of the plane p_1 passing through the point $(0, 2, 4)$ and perpendicular to the vector $\vec{n} = \langle 2, -1, 1\rangle$.

(b) Is the point $(2, 0, 2)$ on the plane p_1?

(c) Write a scalar equation of the plane p_2 that is parallel to p_1 and passes through the point $(3, 0, 4)$.

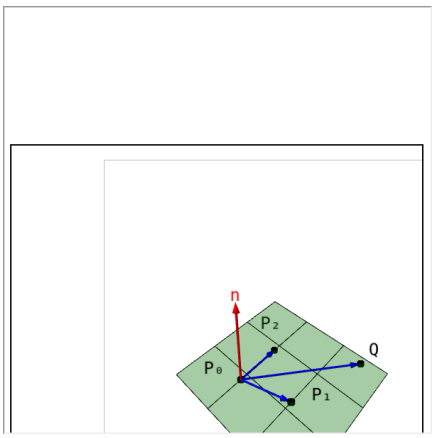
Hint. Compare normal vectors of the planes.

(d) Give parametric equations for the line $\mathcal{L}$ passing through the point $(2, 0, 2)$ and perpendicular to the plane p_3 described by the equation $x+2y-2z = 7$.

(e) Find the point at which $\mathcal{L}$ intersects p_3.

Just as two distinct points in space determine a line, three non-collinear points in space determine a plane. Our next example illustrates a method for finding an equation of the plane containing three non-collinear points in $\mathbb{R}^3$.

Example 9.7.9 Consider three points P_0, P_1, and P_2 in space, not all lying on the same line, as shown in Figure 9.7.10.



Standalone
In Context
Embed

Figure 9.7.10 A plane determined by three points P_0, P_1, and P_2

Notice that the vectors $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_0P_2}$ both lie in the plane p. If we form their cross-product

$$\vec{n} = \overrightarrow{P_0P_1} \times \overrightarrow{P_0P_2},$$

we obtain a normal vector to the plane p. Therefore, if Q is any other point on p, it then follows that $\overrightarrow{P_0Q}$ will be perpendicular to $\vec{n}$, and we have the equation:

$$\vec{n} \cdot \overrightarrow{P_0Q} = 0 \tag{9.7.3}$$

▲

The final activity of this section allows you to practice with finding an equation of a plane passing through three points.

Activity 9.7.3 Let $P_0 = (1, 2, -1)$, $P_1 = (1, 0, -1)$, and $P_2 = (0, 1, 3)$ and let p be the plane containing P_0, P_1, and P_2.

(a) Determine the components of the vectors $\overrightarrow{P_0P_1}$ and $\overrightarrow{P_0P_2}$.

(b) Find a normal vector $\vec{n}$ to p.

(c) Find a scalar equation of p.

(d) Consider a second plane q with scalar equation $-3(x-1) + 4(y+3) + 2(z-5) = 0$. Find two different points on q as well as a vector $\vec{m}$ that is normal to q.

(e) What is the angle between planes p and q?

Summary

- A plane in space is the set of all terminal points of vectors emanating from a given point perpendicular to a fixed vector.
- If P_1, P_2, and P_3 are non-collinear points in space, the vectors $\overrightarrow{P_1P_2}$ and and $\overrightarrow{P_1P_3}$ are vectors in the plane and the vector $\vec{n} = \overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}$ is a normal vector to the plane. So any point P in the plane satisfies the equation $\overrightarrow{PP_1} \cdot \vec{n} = 0$. If we let $P = (x, y, z)$, $\vec{n} = \langle a, b, c \rangle$ be the normal vector, and $P_1 = (x_0, y_0, z_0)$, we can also represent the plane with the equation

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

9.7.3 Exercises

1. Find the intercepts of the plane $6x + y + 7z = 42$. ***Write your answers as points*** (a, b, c) ***where*** a***,*** b***, and*** c ***are numbers.***

1. The x-axis intercept.
Answer: _____

2. The y-axis intercept.
Answer: _____

3. The z-axis intercept.
Answer: _____

Note: *If there is no intersection, write "none".*

2. Find the linear equation of the plane through the origin and the points $(6, 4, 2)$ and $(3, -1, 1)$.
Equation: _____

3. Find the linear equation of the plane through the origin and perpendicular to the vector $\langle -5, 2, 5\rangle$.
Equation: _____

4. Find the linear equation of the plane through the point $(3, -5, 2)$ and with normal vector $\mathbf{j} + 2\mathbf{k}$.
Equation: _____

5. Find an equation of a plane containing the three points (-2, 3, 0), (-3, 5, -4), (-3, 6, -2) in which the coefficient of x is 8.
_____ $= 0$.

6. Find a linear equation for the plane that passes through the point $(1, 6, 10)$ and is parallel to the plane $x + 2y + 5z + 4 = 0$.
Equation: _

7. Find a vector equation with parameter t for the line through the point $(7, 0, 1)$ and perpendicular to the plane $x + 4y + 6z = 5$.
Answer: $\mathbf{r}(t) =$ _____

8. Find a linear equation for the plane that passes through the point $(-2, 5, 5)$ and is perpendicular to the line represented by the vector equation $\mathbf{r}(t) = \langle 1 + 4t, -2 - t, 6 + 2t\rangle$.
Equation: _____

9. Find the point at which the line $\langle 3, -1, -3\rangle + t\langle -2, 2, -4\rangle$ intersects the plane $-5x + 4y - 2z = 65$.
(_____, _____, _____)

10. Find a linear equation for the plane that passes through the point $(1, 2, 3)$ and contains the line represented by the vector equation $\mathbf{r}(t) = \langle 3t, 6 - 2t, 1 - 4t\rangle$.
Equation: _____

11. An implicit equation for the plane passing through the points $(-3, 2, 5)$, $(2, 4, 0)$, and $(0, 3, 1)$ is _____.

12. ***(a)*** Find a vector $\vec{n}$ perpendicular to the plane

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

$\vec{n} =$ _____

(b) Find a vector $\vec{v}$ parallel to the plane.
$\vec{v} =$ _____

13. What do the following equations represent in R^3?

Match the two sets of letters:

a. a vertical plane

b. a horizontal plane

c. a plane which is neither vertical nor horizontal

___ A. $-7x + 10y = 8$

___ B. $x = -6$

___ C. $y = 2$

___ D. $z = -8$

14. Find a vector equation with parameter t for the line through the point $(6, 0, 2)$ and perpendicular to the plane $x + 3y + 6z = 5$.

Answer: $\mathbf{r}(t) =$ _____

15. Consider the following geometry problems in 3-space

Enter T or F depending on whether the statement is true or false. (You must enter T or F -- True and False will not work.)

(a) Two planes orthogonal to a line are parallel

(b) A plane and a line either intersect or are parallel

(c) Two lines orthogonal to a plane are parallel

(d) Two planes parallel to a third plane are parallel

(e) Two lines parallel to a third line are parallel

(f) Two lines either intersect or are parallel

(g) Two planes parallel to a line are parallel

(h) Two lines parallel to a plane are parallel

(i) Two planes orthogonal to a third plane are parallel

(j) Two lines orthogonal to a third line are parallel

(k) Two planes either intersect or are parallel

16. Find an equation of a plane through the point (0, 5, -4) which is orthogonal to the line

$$x = -2 + 1t, \quad y = 1 + 2t, \quad z = 1 - 2t$$

in which the coefficient of x is 1.

Equation: _____ = 0.

17. Find an equation of a plane through the point (1, 1, 0) which is parallel to the plane $-5x - 4y + 2z = -8$ in which the coefficient of x is -5.

_____ = 0.

18. Find an equation of a plane containing the line $\mathbf{r} = \langle 3, -3, 5\rangle + t\langle -1, -6, 3\rangle$ which is parallel to the plane $3x + 1y + 3z = 16$ in which the coefficient of x is 3.

_____ = 0.

19. Find an equation of a plane through the point (4, 5, -1) which is orthogonal to the line

$$x = -5 - 1t, \quad y = 4 + 2t, \quad z = 5 - 4t$$

in which the coefficient of x is -1.

Equation: _____ = 0.

20. Find the point at which the line $\langle 5,1,3\rangle + t\langle -3,1,5\rangle$ intersects the plane $-2x + 2y - 4z = -104$.

(_____, _____, _____)

21. Find the point P where the line x = 1 + t, y = 2t, z = -3t intersects the plane x + y - z = -5.

P = (_____, _____, _____)

22. (A) Find the parametric equations for the line through the point P = (-4, 4, 1) that is perpendicular to the plane $2x + 0y + 1z = 1$.

Use "t" as your variable, t = 0 should correspond to P, and the velocity vector of the line should be the same as the standard normal vector of the plane.

x = _____

y = _____

z = _____

(B) At what point Q does this line intersect the yz-plane?

Q = (_____, _____, _____)

23. Find the angle of intersection of the plane $-2x - 5y - 1z = 1$ with the plane $-5x - 3y + 1z = 0$.

Answer in radians:

and in degrees:

24. Find the angle of intersection of the plane $-1x - 4y - 4z = 5$ with the plane $-1x + 1y - 2z = 4$.

Answer in radians:

and in degrees:

25. Compare the the planes below to the plane $4x + 4y + 4z = -5$. Match the letter corresponding to the words parallel, orthogonal, or "neither" which describes the relation of the two planes.

(a) $-4x - 4y - 4z = -1$

(b) $-4x + 4y = -1$

(c) $4x + 5y + 4z = -3$

A. orthogonal

B. parallel

C. neither

26. Find the vector equation for the line of intersection of the planes $4x - 2y + z = 1$ and $4x + 3z = 0$

$\mathbf{r} = \langle$_____, _____,0 $\rangle + t\langle$-6, _____, _____ $\rangle$.

27. The plane that passes through the point $(0, 0, -2)$ and is perpendicular to both $2x + y + 5z = 12$ and $2z - (2x + 4y) = 16$ has _____ as its implicit equation.

Hint. Hint:

28. Find an equation for the plane that contains both of the lines described in Exercise 9.6.4.25.

29. This exercise explores key relationships between a pair of planes. Consider the following two planes: one with scalar equation $4x - 5y + z = -2$, and the other which passes through the points $(1, 1, 1)$, $(0, 1, -1)$, and

$(4, 2, -1)$.

(a) Find a vector normal to the first plane.

(b) Find a scalar equation for the second plane.

(c) Find the angle between the planes, where the angle between them is defined by the angle between their respective normal vectors.

(d) Find a point that lies on both planes.

(e) Since these two planes do not have parallel normal vectors, the planes must intersect, and thus must intersect in a line. Observe that the line of intersection lies in both planes, and thus the direction vector of the line must be perpendicular to each of the respective normal vectors of the two planes. Find a direction vector for the line of intersection for the two planes.

(f) Determine parametric equations for the line of intersection of the two planes.

30.

(a) Find the plane that goes through the points $(1, 2, 3)$, $(4, 5, 6)$, and $(7, 8, 9)$ or explain why there is not one plane that satisfies this condition.

(b) Find *all* planes in $\mathbb{R}^3$ whose line of intersection with the xz-plane is given by $\pi x + ez = \sqrt{5}$.

(c) Give two vectors that are in the plane given by $3x + 4y = 6z - 20$ and are orthogonal to each other.

Hint. What geometric information do you have about vectors in the given plane?

(d) Give the equations of *two* planes that make a $90°$ angle with $3x + 4y = 6z - 20$ and have an angle of $90°$ between them as well.

(e) Give the parametric form of the line of intersection for the planes given by $x + y + z = 10$ and $-2x - 4y + 2z = 7$.

31. Determining a plane that contains two lines. Find an equation for the plane containing the line in the xy-plane where $x = 6$, and the line in the yz-plane where $z = 12$.

equation: _____

32. ***(a)*** Find a vector parallel to the line of intersection of the planes $5x - y + 3z = -1$ and $-5x - y - 3z = 3$.

$\vec{v} =$ _____

(b) Show that the point $(-1, -1, 1)$ lies on both planes. Then find a vector parametric equation for the line of intersection.

$\vec{r}(t) =$ _____

33. Find the vector equation for the line of intersection of the planes $x - 5y + 5z = -1$ and $x + 3z = 1$

$\mathbf{r} = \langle$_____, _____,0 $\rangle + t\langle$-15, _____, _____ $\rangle$.

34. Find a vector equation with parameter t for the line of intersection of the planes $x + y + z = 3$ and $x + z = 0$.

Answer: $\mathbf{r}(t) =$ _____

35. Determine whether the following planes below are *parallel*, *perpendicular*, or *neither*.

(☐ choose one ☐ Parallel ☐ Perpendicular ☐ Neither) ***1.*** $x + 2y - 6z = 0$ and $-3x - 6y + 18z = -3$.

(☐ choose one ☐ Parallel ☐ Perpendicular ☐ Neither) ***2.*** $x - 3y + z = 0$ and $-x - 4y + z = 5$.

(☐ choose one ☐ Parallel ☐ Perpendicular ☐ Neither) ***3.*** $x + 7z = 0$ and $7x - z = 3$.

Note: *You only have two attempts at this problem.*

36. Find the cosine of the angle between the planes $x + y + z = 0$ and $3x + 2y + z = 1$.

Answer: $\cos\theta =$ _____

37. Find the angle in radians between the planes $1x + z = 1$ and $-1y + z = 1$.

38. Suppose that the line ℓ is represented by $\mathbf{r}(t) = \langle 9 + 2t, 12 + 4t, 5 + 2t\rangle$ and the plane P is represented by $3x - 2y + 5z = 8$.

1. Find the intersection of the line ℓ and the plane P. ***Write your answer as a point*** (a, b, c) ***where*** a***,*** b***, and*** c ***are numbers.***

Answer: _____

2. Find the cosine of the angle θ between the line ℓ and the normal vector of the plane P.

Answer: $\cos\theta =$ _____

Note: *If there is no intersection, write "none" for both answers.*

39. In this problem, we explore how we can use what we know about vectors and projections to find the distance from a point to a plane.

Let p be the plane with equation $z = -4x + 3y + 4$, and let $Q = (4, -1, 8)$.

(a) Show that Q does not lie in the plane p.

(b) Find a normal vector $\vec{n}$ to the plane p.

(c) Find the coordinates of a point P in p.

(d) Find the components of $\overrightarrow{PQ}$. Draw a picture to illustrate the objects found so far.

(e) Explain why $|\text{comp}_{\vec{n}}\overrightarrow{PQ}|$ gives the distance from the point Q to the plane p. Find this distance.

9.8 Common Graphs in Three Dimensions

Motivating Questions

- If an equation involving only two variables is plotted in three-dimensional space, what sort of surface is produced?
- What surfaces arise as graphs of an equation that contains all three of the variables x, y, and z and the equation is quadratic in at least one variable?

9.8.1 Introduction

In this section, we will introduce some examples of graphs in three dimensions that have nice algebraic properties and a variety of interesting geometric

features. In Preview Activity 9.8.1, we ask you to practice identifying conic sections such as circles, parabolas, ellipses, and hyperbolas from their equations. The section will then help you see how to use your knowledge of conic sections to identify key properties of graphs in *three* dimensions. For a complete discussion of conic sections, see Section A.1.

Preview Activity 9.8.1 For each equation below, identify its graph in Figure 9.8.1. Note that there are more graphs than equations, so some graphs will not be selected. To help you with identification, you might consider looking for x-intercepts, y-intercepts, and values of a variable for which the graph contains no points.

i. $\frac{x^2}{9} + \frac{y^2}{25} = 1$

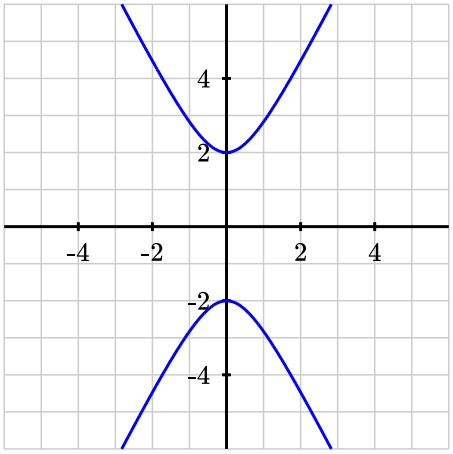
ii. $\frac{x}{3} + \frac{y}{5} = 1$

iii. $\frac{y^2}{4} - \frac{x^2}{1} = 1$

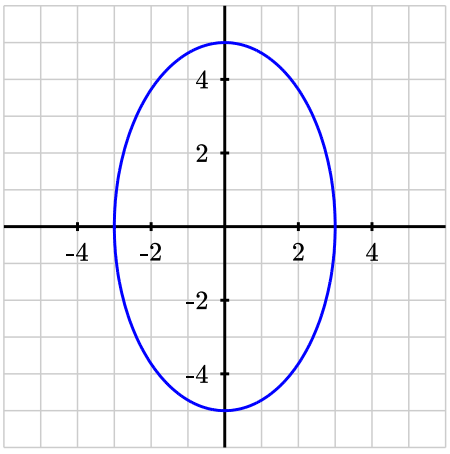
iv. $x = 2y^2$

v. $\frac{y^2}{9} - \frac{x^2}{25} = 0$

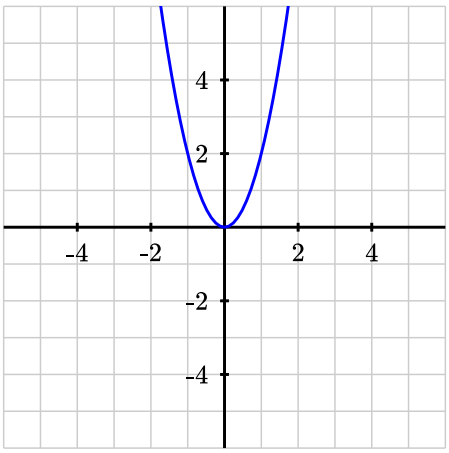
vi. $\frac{x^2}{1} - \frac{y^2}{4} = 1$


4
2
-4
-2
2
4
-2
-4

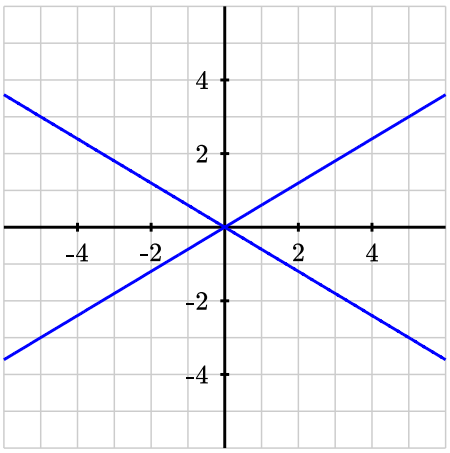
(a)


4
2
-4
-2
2
4
-2
-4

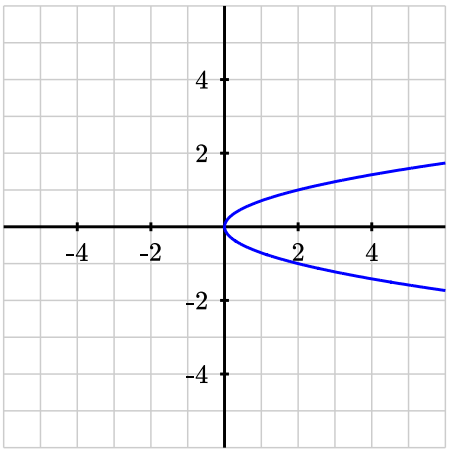
(b)


4
2
-4
-2
2
4
-2
-4

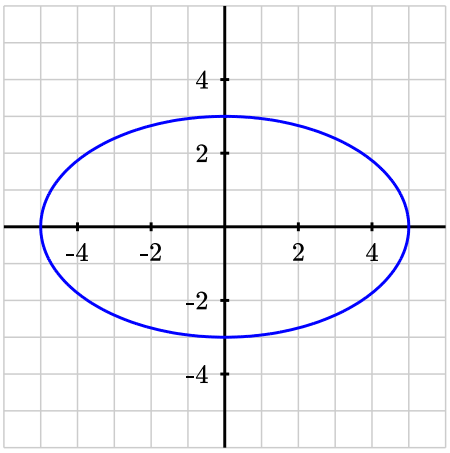
(c)


4
2
-4
-2
2
4
-2
-4

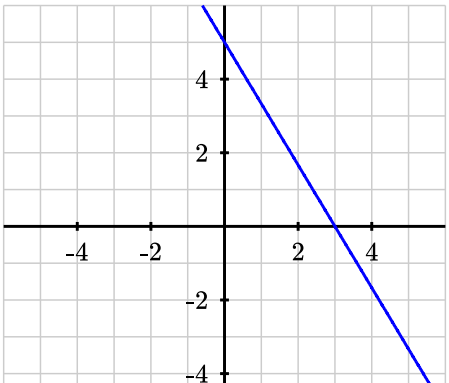
(d)

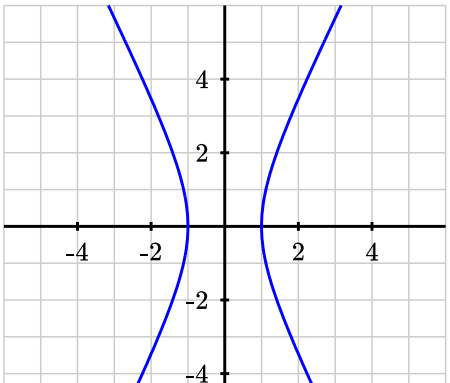

4
2
-4
-2
2
4
-2
-4

(e)


4
2
-4
-2
2
4
-2
-4

(f)


4
2
-4
-2
2
4
-2
-4


4
2
-4
-2
2
4
-2
-4

9.8.2 Cylinder Surfaces

Preview Activity 9.8.1 gave you an opportunity to recall that there are a number of interesting curves in the xy-plane that have nice algebraic forms (equations). What happens if we consider these same equations in three dimensions?

Example 9.8.2 We start by considering $(x-1) = \frac{(y+2)^2}{2}$. The graph of $(x-1) = \frac{(y+2)^2}{2}$ in two dimensions is a parabola centered at $(1, -2)$.

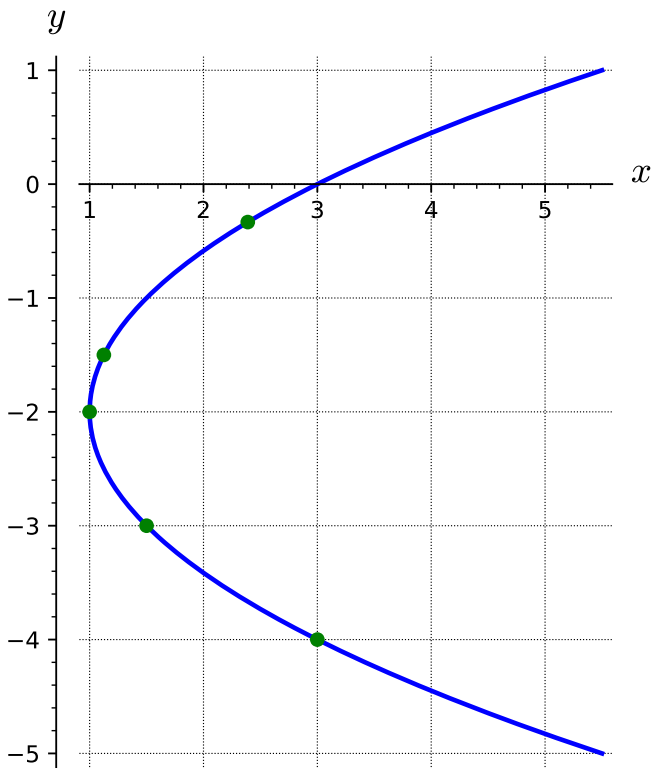
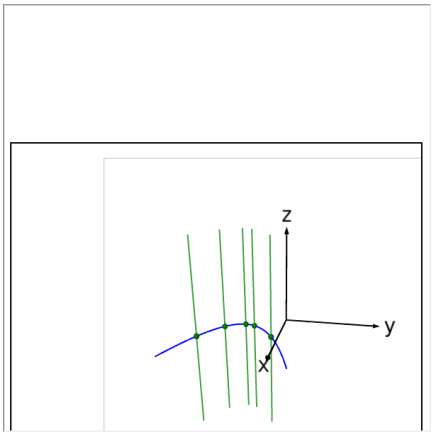


Figure 9.8.3 A plot of $(x-1) = \frac{(y+2)^2}{2}$ in the xy-plane with five points highlighted

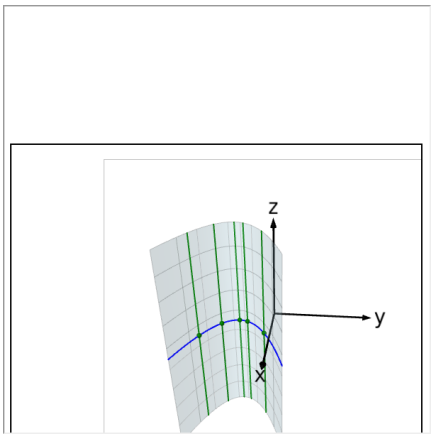
If we want to consider the graph of $(x-1) = \frac{(y+2)^2}{2}$ in three dimensions, then the graph must include all of the (x, y, z) points that will satisfy this equation. For any x and y values we pick that satisfy $(x-1) = \frac{(y+2)^2}{2}$, we can make *any* choice of z to get a point that also satisfies the given equation. Hence, for each of the highlighted points on Figure 9.8.3, we can extend the graph of $(x-1) = \frac{(y+2)^2}{2}$ parallel to the z-axis, forming a vertical line. Figure 9.8.4 shows how the points on the parabola (in blue) can be extended to include any z-coordinate.



Standalone
In Context
Embed

Figure 9.8.4 A plot of $(x-1) = \frac{(y+2)^2}{2}$ generating curve and several rulings

Extending all points from the parabola $(x-1) = \frac{(y+2)^2}{2}$ in the xy-plane parallel to the z-axis will give a surface. This kind of surface is called a **cylinder surface**. We call the two-dimensional curve used to make the surface the **generating curve**, and the lines that extend in the direction of the missing variable are called **rulings**. In Figure 9.8.5, the surface is plotted in blue, the generating curve in black, and the rulings in green. This surface is called a **parabolic cylinder surface** because the generating curve is a parabola.



Standalone
In Context
Embed

Figure 9.8.5 A plot of the cylinder surface $(x-1) = \frac{(y+2)^2}{2}$ with generating curve and several rulings

▲

The next activity prompts you to look at the most important cylinder surfaces. The conic sections that were reviewed in Preview Activity 9.8.1 will be helpful as you do this.

Activity 9.8.2 Each equation below is a cylinder surface in $\mathbb{R}^3$. To sketch the cylinder surfaces in xyz-space, you should first draw the generating curve in the xy-plane, xz-plane, or yz-plane (depending on which two variables appear in the equation) and then sketch a three-dimensional cylinder surface by thinking about how the rulings will run.

(a) $2x - y + 1 = 0$ is called a **linear cylinder surface**.

(b) $(x-1)^2 + (y+2)^2 = 4$ is called a **right-circular cylinder surface**.

(c) $\frac{x^2}{9} + \frac{z^2}{4} = 1$ is called an **elliptic cylinder surface**.

(d) $x^2 - y^2 = 1$ is called a **hyperbolic cylinder surface**.

9.8.3 Quadric Surfaces

The defining characteristic of the equations of cylinder surfaces is that one variable is completely omitted from the equation. Now we will examine surfaces with algebraic equations that are *quadratic* in x, y, and z. This will give us a category of example surfaces that are simple algebraically but exhibit a variety of interesting and important characteristics. To understand these **quadric surfaces**, we focus on the intercepts and two-dimensional graphs formed by the intersection of the quadric surface with fundamental planes. Next activity leads you through this process to help you learn to recognize and sketch quadric surfaces. After this, we will summarize the key ideas for recognizing and sketching quadric surfaces and then you will have the opportunity to apply those key ideas in additional activities.

Activity 9.8.3 For this activity, we will be looking at a variety properties that will help us draw a graph of the surface described by $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$.

(a) Find all x-, y-, and z-intercepts of $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$.

Hint. The x-intercepts are where $y = 0$ and $z = 0$.

(b) Find an equation for the curve given by the intersection of $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$ with the xy-plane, the yz-plane, and the xz-plane. Draw a two-dimensional plot of each intersection.

(c) Find equations for the curve given by the intersection of $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$ with the each of the following fundamental planes. You should state the shape and any other characteristics (like center or direction) for each of these intersections.

(i) $z = \sqrt{3}$

(ii) $z = -2$

(iii) $x = 3$

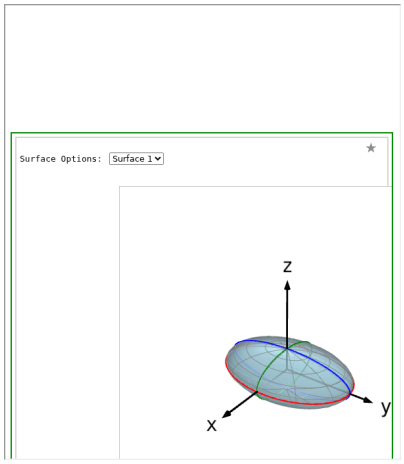
(iv) $x = -1$

(v) $y = 2$

(vi) $y = -4$

(d) Sketch each of these intersections on the proper fundamental planes in three dimensions.

(e) Which of the following surface plots will correspond to $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$? You can determine this by comparing the features on your previous part to these options.



Standalone
In Context
Embed

Figure 9.8.6 Choose the surface with equation $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$

In designing Activity 9.8.3, we carefully chose the fundamental planes for which we asked you to find the intersection with the surface. By examining the equation that defines a surface, you can strategically select fundamental planes to investigate. Doing so will allow you to identify key characteristics of a quadric surface without needing to test an exceedingly large number of fundamental planes. In Key Idea 9.8.7, we suggest some steps to follow. Afterward, you will have an activity that gives you a chance to practice implementing these steps.

Key Idea 9.8.7 Recognizing Quadric Surfaces.

1. *Identify the intercepts of the surface with the x-axis, y-axis, and z-axis, including noting if the surface does not intersect one or more of these coordinate axes. To find the intercepts, set two of the variables equal to 0 and solve for the third.*

2. *Identify the intersection curve of the surface with the xy-plane, yz-plane, and xz-plane, including noting if the surface does not intersect one or more of these coordinate planes. To find these intersections, set one variable equal to 0 and manipulate the equation so that you can recognize the shape of the curve. In most cases, the curve will be a conic section, and you should pay particular attention to the intercepts of the conic section.*

3. *Identify the intersection curve of the surface with at least two fundamental planes parallel to each of the three coordinate planes. To find these intersections, set a variable equal to a nonzero constant and manipulate the equation so that you can recognize the shape of the curve. In most cases, the curve will be a conic section, and you should pay particular attention to the intercepts of the conic section.*

   Practice will help you develop a strategy for choosing these constants, but the following suggestions will be helpful as you are practicing:

   - *Use what you learned in the two prior steps. For example, if a surface doesn't intersect the xy-plane ($z = 0$), find values of z that do create intersections with fundamental planes.*
   - *Identify if there are (other) values for a variable where the surface no longer intersects with the fundamental planes with that variable fixed.*

- *Use a mixture of positive and negative values for the constants determining your fundamental planes.*
- *Determine if the shape of the intersection varies depending on the value you choose for a fixed coordinate. For example, if all intersections with planes of the form $x = c$ are hyperbolas, do some values of c cause the hyperbolas to have intercepts on one axis while other values of c cause the hyperbolas to have intercepts on another axis? If there are variations of this form, identify them as best as you can.*

To illustrate parts of Key Idea 9.8.7, we can return to Activity 9.8.3 and discuss some of our strategy in selecting the fundamental planes we asked you to consider:

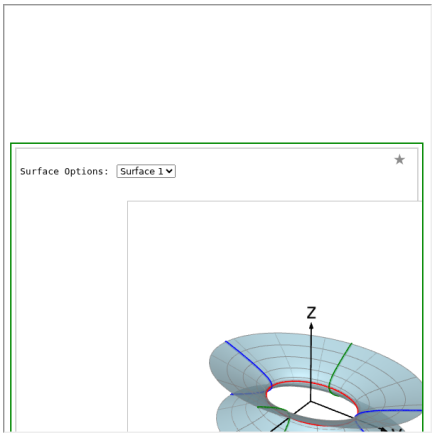
- For planes of the form $z = c$, the intersection equation is always of the form $\frac{x^2}{4} + \frac{y^2}{9} = 1 + c^2$, which gives an ellipse for all values of c.
- For planes of the form $x = c$, the intersection equation is $\frac{y^2}{9} - z^2 = 1 - \frac{c^2}{4}$. When $c^2/4 < 1$,, this is a hyperbola with positive coefficient on the y^2 term. This hyperbola has intercepts on the y-axis. When $c^2/4 > 1$, the right-hand side of the equation is negative, so a more standard form for the equation would be $z^2 - \frac{y^2}{9} = \frac{c^2}{4} - 1$, which is a hyperbola that has intercepts on the z-axis.
- For planes of the form $y = c$, the intersection equation is $\frac{x^2}{4} - z^2 = 1 - \frac{c^2}{9}$. As with planes of the form $x = c$, these are hyperbolas where the location of the intercepts depends on the sign of $1 - \frac{c^2}{9}$.

Activity 9.8.4 For each equation below, use the process in Key Idea 9.8.7 to determine the shape of the quadric surface defined by the equation and identify the graph of the quadric surface in Figure 9.8.8.

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

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

(c) $\frac{x^2}{4} - \frac{y^2}{9} - \frac{z}{1} = 0$



Standalone
In Context
Embed

Figure 9.8.8 A plot of surfaces to select from

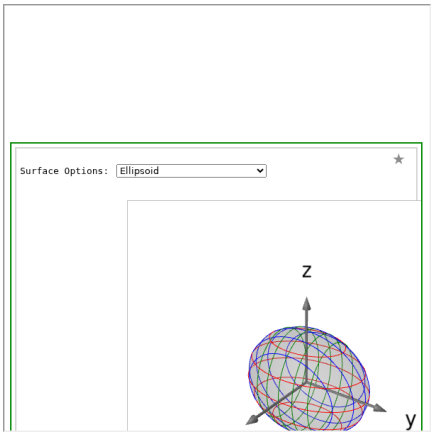
As mentioned before, the surfaces given by $\frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{1} = 1$, $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$, $\frac{x^2}{4} - \frac{y^2}{9} - \frac{z^2}{1} = 1$, and $\frac{x^2}{4} - \frac{y^2}{9} - \frac{z}{1} = 1$ are examples of **quadric surfaces**.

Quadric surfaces are the surfaces generated by polynomials that are quadratic in at least one of the three coordinate variables: x, y, and z. There are six main shapes of quadric surfaces. The box below gives the forms of the equations of these. Figure 9.8.9 shows plots of each of the main shapes with the algebraic form used to describe each shape. Note that all of these examples are oriented along the z-axis and centered at the origin. However, as with conic sections, quadric surfaces can be oriented along other coordinate directions and/or centered away from the origin.

Quadric Surfaces.

The following is a list of the six types of quadric surfaces and their general algebraic form centered at the origin with stretches of a, b, and c in the respective coordinate directions. Each of these is oriented along the z-axes (when possible).

Ellipsoid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$
Hyperboloid of 1-Sheet	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$
Hyperboloid of 2-Sheet	$-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$
Elliptic Paraboloid	$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z$
Hyperbolic Paraboloid (Saddle)	$\frac{x^2}{a^2}-\frac{y^2}{b^2}=z$
Cone	$\frac{x^2}{a^2}-\frac{y^2}{b^2}=\frac{z^2}{c^2}$



Standalone
In Context
Embed

Figure 9.8.9 Quadric Surfaces

In Chapter 11, we will talk about surfaces where one coordinate can be expressed as a function of the others. This is an extension of the idea that graphs that pass the vertical line test can be expressed with y as a function of x. For instance, in the plot of the saddle surface shown in Figure 9.8.9, any line parallel to the z-axis will intersect the saddle surface at only one place. This means that the saddle surface can be expressed with z as a function of x and y. This is not surprising since the saddle surface is given by $\frac{x^2}{a^2}-\frac{y^2}{b^2}=z$.

Further discussion of writing one variable as a function of the others can be found in Exercise 9.2.5.9 and Section 11.2.

Compare this to the graph of a hyperboloid of one sheet in Figure 9.8.9, where a line that is parallel to the z-axis will intersect the surface at two places. Lines parallel to the x- or y-axes will also intersect the hyperboloid of one sheet in more than one place. This means that we *cannot* express the graph of a hyperboloid of one sheet with one of the coordinates as a function of the other two. Algebraically, this corresponds to there being more than one solution when you try to solve the equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ for one variable. In particular, the multiple solutions come from needing to consider the positive and negative square roots. We will return to this idea in Chapter 11.

Summary

- Translated Conic Sections have the following forms:
  - Ellipse: $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ with center (h, k), vertices $(h \pm a, k)$ and covertices $(h, k \pm b)$
  - Hyperbola: $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ with center (h, k), vertices $(h \pm a, k)$, and asymptotes $(y - k) = \frac{b}{a}(x - h)$
  - Parabola: $(x - h) = A(y - k)^2$ with vertex/center (h, k)

  Other versions of these forms that are oriented vertically will have the roles of x and y switched.
- Cylinder surfaces are described algebraically by an equation involving only two coordinate variables. Geometrically, a cylinder surface is generated by a curve/graph in the coordinate plane involving the two coordinate variables in the equation and stretching this generating curve parallel to the missing coordinate variable's axis.
- Quadric surfaces are a category of surfaces created by quadratic equations in x, y, and z. Quadric surfaces have six typical shapes: ellipsoid, hyperboloid of one sheet, hyperboloid of two sheets, elliptic paraboloid, hyperbolic paraboloid (saddle surface), and cone. The same coordinate transformations that generalize conic sections can be applied to quadric surfaces.

9.8.4 Exercises

1. Identify the type of quadric surface defined by the equation

$$4y^2 - 36x^2 + z^2 = 1,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with

x-intercepts when $x =$ _____,
y-intercepts when $y =$ _____, and
z-intercepts when $z =$ _____.

Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

2. State the type of the quadratic surface:
$\left(\frac{x}{7}\right)^2 - \left(\frac{y}{6}\right)^2 + \left(\frac{z}{8}\right)^2 = 1$
1. Ellipsoid
2. Hyperboloid of one sheet
3. Hyperboloid of two sheets
4. None of these

Describe the trace obtained by intersecting with the plane $z = 1$:
1. Ellipse
2. Hyperbola
3. Circle
4. Empty set

3. Identify the following surfaces as an elliptical paraboloid, hyperbolic paraboloid, a hyperboloid of one sheet, a hyperboloid of two sheets, a cone, a circular cylinder, an elliptical cylinder, or a parabolic cylinder, and identify the axis of symmetry as the x-axis, the y-axis, or the z-axis.

a. (☐ circular cylinder ☐ elliptical paraboloid ☐ cone ☐ hyperboloid of two sheets ☐ parabolic cylinder ☐ hyperboloid of one sheet ☐ hyperbolic paraboloid ☐ elliptical cylinder) (☐ along the x-axis ☐ along the z-axis ☐ along the y-axis) $4x^2 + 6y^2 - z = 0$

b. (☐ elliptical cylinder ☐ hyperbolic paraboloid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ parabolic cylinder ☐ elliptical paraboloid ☐ cone ☐ circular cylinder) (☐ along the y-axis ☐ along the z-axis ☐ along the x-axis) $4y^2 - 4x^2 - 9z^2 + 1 = 0$

c. (☐ circular cylinder ☐ elliptical paraboloid ☐ cone ☐ hyperboloid of two sheets ☐ parabolic cylinder ☐ hyperboloid of one sheet ☐ hyperbolic paraboloid ☐ elliptical cylinder) (☐ along the y-axis ☐ along the x-axis ☐ along the z-axis) $-9x^2 - 4y^2 + z^2 = 1$

d. (☐ elliptical cylinder ☐ hyperbolic paraboloid ☐ elliptical paraboloid ☐ cone ☐ circular cylinder ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ parabolic cylinder) (☐ along the z-axis ☐ along the x-axis ☐ along the y-axis) $-4x^2 + 4y^2 - z^2 = 0$

e. (☐ elliptical cylinder ☐ hyperbolic paraboloid ☐ hyperboloid of one sheet ☐ parabolic cylinder ☐ hyperboloid of two sheets ☐ cone ☐ elliptical paraboloid ☐ circular cylinder) (☐ along the y-axis ☐ along the z-axis ☐ along the x-axis) $6y^2 + 6z^2 = 1$

4. Match the surfaces with the appropriate descriptions.

(a) $z = x^2$

(b) $z = 2x^2 + 3y^2$

(c) $z = 4$

(d) $x^2 + 2y^2 + 3z^2 = 1$

(e) $z = y^2 - 2x^2$

(f) $z = 2x + 3y$

(g) $x^2 + y^2 = 5$

A. horizontal plane

B. elliptic paraboloid

C. parabolic cylinder

D. nonhorizontal plane

E. hyperbolic paraboloid

F. ellipsoid

G. circular cylinder

5. By completing the square, put the equation

$$3x^2 + y^2 + 7z^2 = 13 - 6x - 12y + 14z$$

into the standard form $a(x - x_0)^2 + b(y - y_0)^2 + c(z - z_0)^2 = d$. Then, use your answer to sketch a graph of this quadric surface on paper.

_____ (_____)^2+ _____ (_____)^2+ _____ (_____)^2 = _____

6. Identify the type of quadric surface defined by the equation

$$\frac{y^2}{25} - \frac{x^2}{9} + z = 1,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with

x-intercepts when x = _____,
y-intercepts when y = _____, and
z-intercepts when z = _____.

Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

7. Identify the type of quadric surface defined by the equation

$$y^2 + 5z^2 = 25,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with

x-intercepts when x = _____,
y-intercepts when y = _____, and
z-intercepts when z = _____.

Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

8. Identify the type of quadric surface defined by the equation

$$5x^2 + 4y^2 + z = 20,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with

x-intercepts when $x =$ _____,
y-intercepts when $y =$ _____, and
z-intercepts when $z =$ _____.
Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

9. Identify the type of quadric surface defined by the equation

$$5x^2 + 5y^2 - z^2 = -25,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with
x-intercepts when $x =$ _____,
y-intercepts when $y =$ _____, and
z-intercepts when $z =$ _____.
Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

10. Identify the type of quadric surface defined by the equation

$$5x^2 + 6y^2 + 3z^2 = 90,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with
x-intercepts when $x =$ _____,
y-intercepts when $y =$ _____, and
z-intercepts when $z =$ _____.
Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

11. Identify the type of quadric surface defined by the equation

$$x^2 - \frac{y^2}{8} + \frac{z^2}{64} = 0,$$

and find all x-, y-, and z-intercepts of the resulting graph. Sketch the graph of this quadric surface on paper.

The quadric surface is a / an (☐ Choose ☐ elliptic cylinder ☐ hyperbolic cylinder ☐ parabolic cylinder ☐ elliptic paraboloid ☐ hyperbolic paraboloid ☐ ellipsoid ☐ hyperboloid of one sheet ☐ hyperboloid of two sheets ☐ elliptic cone) with
x-intercepts when $x =$ _____,
y-intercepts when $y =$ _____, and
z-intercepts when $z =$ _____.
Enter your answers as comma separated lists, or enter *NONE* if there are no intercepts of a particular type.

12. Match the surfaces with the appropriate descriptions.

(a) $z = x^2$

(b) $z = 2x^2 + 3y^2$

(c) $x^2 + 2y^2 + 3z^2 = 1$

(d) $z = y^2 - 2x^2$

(e) $z = 4$

(f) $x^2 + y^2 = 5$

(g) $z = 2x + 3y$

A. hyperbolic paraboloid

B. nonhorizontal plane

C. parabolic cylinder

D. ellipsoid

E. circular cylinder

F. elliptic paraboloid

G. horizontal plane

13. Match the surfaces with the appropriate descriptions.

(a) $x^2 + 2y^2 + 3z^2 = 1$

(b) $z = 2x + 3y$

(c) $z = 4$

(d) $x^2 + y^2 = 5$

(e) $z = x^2$

(f) $z = 2x^2 + 3y^2$

(g) $z = y^2 - 2x^2$

A. nonhorizontal plane

B. elliptic paraboloid

C. ellipsoid

D. horizontal plane

E. circular cylinder

F. parabolic cylinder

G. hyperbolic paraboloid

14. State whether the equation

$\left(\frac{x}{3}\right)^2 - \left(\frac{y}{5}\right)^2 - \left(\frac{z}{8}\right)^2 = 1$

defines (enter number of statement):

1. A hyperboloid of two sheets
2. A hyperboloid of one sheet
3. An ellipsoid
4. None of these

15. State whether the equation

$9x^2 + 4y^2 - 16z^2 = 1$

defines (enter number of statement):
1. A hyperboloid of two sheets
2. A hyperboloid of one sheet
3. An ellipsoid
4. None of these

16. State whether the equation
$z = \left(\frac{x}{4}\right)^2 + \left(\frac{y}{7}\right)^2$
defines: (☐ An elliptic cone ☐ An elliptic paraboloid ☐ None of these ☐ A hyperbolic paraboloid)

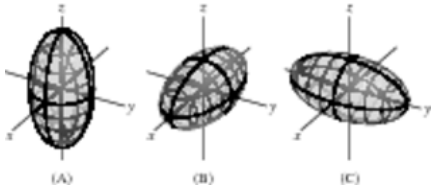
17. State the type of the quadratic surface:
$x^2 - \left(\frac{y}{5}\right)^2 + z^2 = 1$
1. Hyperboloid of two sheets
2. Hyperboloid of one sheet
3. Ellipsoid
4. None of these

Describe the trace obtained by intersecting with the plane $y = 0$:
1. Ellipse
2. Hyperbola
3. Circle
4. Empty set

18. State the type of the quadratic surface:
$\left(\frac{x}{6}\right)^2 + \left(\frac{y}{2}\right)^2 + z^2 = 1$
1. Hyperboloid of two sheets
2. Hyperboloid of one sheet
3. Ellipsoid
4. None of these

Describe the trace obtained by intersecting with the plane $z = 0.25$:
1. Ellipse
2. Hyperbola
3. Circle
4. Empty set

19.



Match the ellipsoids shown in the figure above with the equations:
_ 1) $x^2 + 16y^2 + 16z^2 = 256$
_ 2) $16x^2 + y^2 + 16z^2 = 256$
_ 3) $16x^2 + 16y^2 + z^2 = 256$

20. By completing the square, put the equation

$$x^2 - 25y^2 + z^2 + 16x - 10z = -89$$

into the standard form $a(x - x_0)^2 + b(y - y_0)^2 + c(z - z_0)^2 = d$. Then, use your answer to sketch a graph of this quadric surface on paper.
_____ (_____)^2+ _____ (_____)^2+ _____ (_____)^2 = _____

21. (a) Describe the set of points whose distance from the z-axis equals the distance from the xy-plane.

◯ A cylinder opening along the z-axis

◯ A cone opening along the y-axis

◯ A cylinder opening along the x-axis

◯ A cylinder opening along the y-axis

◯ A cone opening along the x-axis

◯ A cone opening along the z-axis

(b) Find the equation for the set of points whose distance from the z-axis equals the distance from the xy-plane.

◯ $z^2 = x^2 + y^2$

◯ $y^2 = x^2 + z^2$

◯ $y^2 + z^2 = r^2$

◯ $x^2 + y^2 = r^2$

◯ $x^2 = y^2 + z^2$

◯ $x^2 + z^2 = r^2$

22. Translating Center of Figure. In this exercise, we will look at a few different surfaces and the algebraic forms used to translate the center or key feature of the surface. In particular, we will look at spheres, cones, and a parabolic cylinder surface with translations away from the origin.

(a) For this part, we will be looking at transformations of the sphere given by $x^2 + y^2 + z^2 = 1$.

(i) We need to start with a description of the key features of our surface. For a sphere, you need to state the center and radius since the surface is going to be complete described by these two features. Write a sentence that gives the center and radius for the sphere given by $x^2 + y^2 + z^2 = 1$.

(ii) Write a sentence that gives the center and radius for the sphere given by $(x-2)^2 + y^2 + z^2 = 1$.

(iii) Give the equation for a sphere with center $(-4, 2, -3)$ and radius 1.

(iv) Generalize your work to give the equation of a sphere of radius 1 with center (x_0, y_0, z_0).

(b) For this part, we will be looking at transformations of the cone given by $x^2 + y^2 = z^2$.

(i) We need to start with a description of the key features of our surface. For a cone, you need to state the center (where the vertex of the cone occurs) and the axis of symmetry that the cone is centered on, since the cone is going to be complete described by these two features. Write a sentence that gives the center and axis of symmetry for the cone given by $x^2 + y^2 = z^2$.

(ii) Write a sentence that gives the center and axis of symmetry for the cone given by $(x)^2 + (y-2)^2 = (z+1)^2$.

(iii) Give the equation for a cone with axis of symmetry along the x-axis and center at the origin.

Hint. Switch the roles of x and z.

(iv) Give the equation for a cone with axis of symmetry along the y-axis and center at $(1, -3, 2)$.

(v) Generalize your work to give the equation of cone with axis of symmetry along the x-axis and center at (x_0, y_0, z_0).

(c) For this part, we will be looking at transformations of the parabolic cylinder given by $x^2 = y$.

(i) We need to start with a description of the key features of our surface. For a parabolic cylinder surface, you need to state the plane with the generating curve and the features of the generating curve, since the surface is going to be complete described by these two features. Write a couple of sentences to describe this cylinder surface given by $y = x^2$. Be sure to include which plane contains the generating curve and the important features of the generating curve. You should also state which axis the rulings will be parallel to.

(ii) Write a couple of sentences to describe the parabolic cylinder given by $-(y + 1) = (x - 2)^2$.

(iii) Write a couple of sentences to describe the parabolic cylinder given by $(x + 3) = (z - 1)^2$.

(iv) Give the equation for a parabolic cylinder surface with generating curve in the yz-plane, with vertex of the generating parabola at $(0, -2, 4)$, and parabola extending in the negative z direction.

23. In this exercise, we will look at how stretches of each coordinate direction appear algebraically as a transformation of our quadric and cylinder surfaces.

(a) In this part, we will look at coordinate stretching/shrinking being applied to a sphere.

(i) For the sphere given by $x^2 + y^2 + z^2 = 1$, describe how the surface is stretched (relative to the center) in each of the coordinate directions.

(ii) For the sphere given by $x^2 + y^2 + z^2 = 16$, describe how the surface is stretched (relative to the center) in each of the coordinate directions.

(iii) What kind of surface is given by $\frac{x^2}{9} + \frac{y^2}{4} + z^2 = 1$. Give the shape of the surface and how it is stretched relative to the basic sphere in the previous task.

(iv) Give the equation of an ellipsoid with center $(1, 2, 3)$ that has been stretched by a by a factor of 6 parallel to the x-axis, by a factor of $2/3$ parallel to the y-direction, and by a factor of π parallel to the z-direction.

(b) In this part of the exercise, we will look at coordinate stretching/ shrinking being applied to a hyperboloid of one-sheet.

(i) For the hyperboloid of one-sheet given by $x^2 + y^2 - z^2 = 1$, describe how the surface is stretched (relative to the center)

in each of the coordinate directions. You should also give the coordinate axis that describes the axis of symmetry/rotation for this hyperboloid.

(ii) For the hyperboloid of one-sheet given by $\frac{x^2}{9} + \frac{y^2}{4} - z^2 = 1$, describe how the surface is stretched (relative to the center) in each of the coordinate directions. You should also give the coordinate axis that describes the axis of symmetry/rotation for this hyperboloid.

(iii) Give the equation of a hyperboloid of one sheet centered on the y-axis that has been stretched by a factor of five parallel to the x-axis and by a factor of $\frac{6}{7}$ parallel to the z-axis.

(iv) How would the description of the surface given by $\frac{(x-2)^2}{25} + \frac{(z-\pi)^2}{\frac{36}{49}} - \frac{y^2}{4} = 1$ be different than the surface from the previous task

Hint. Make sure you look at where the axis of symmetry/ rotation will be.

24. Using your newly found understanding of cylinder and quadric surfaces, for each equation given below, write a couple sentences to describe the shape and key features of each surface. You should try to draw a good plot of each equation by hand and label your graphs.

(a) $25x^2 + 4y^2 + 100z^2 = 100$

(b) $25x^2 + 4y^2 - 100z^2 = 100$

(c) $-25x^2 - 4y^2 + 100z^2 = 100$

(d) $25x^2 - 4y^2 = 0$

(e) $25x^2 + 4y^2 + 100z^2 = 0$

(f) $25x^2 + 4y^2 + 100z = 0$

(g) $25x^2 - 4y^2 + 100z = 0$

(h) $-4y^2 + 100z = 0$

(i) $25x + 4y + 100z = 100$

(j) $-25x^2 - 4y^2 - 100z^2 = 100$

(k) $25x^2 - 4y^2 + 100z^2 = -100$

9.9 Polar, Cylindrical, and Spherical Coordinates

Motivating Questions

- What are the polar coordinates of a point, and how are they related to rectangular coordinates?
- What are the cylindrical coordinates of a point, and how are they related to rectangular coordinates?
- What are the spherical coordinates of a point, and how are they related to rectangular coordinates?

- How can we convert coordinates of points or equations between coordinate systems?

9.9.1 Introduction

In this preview activity, we will investigate some measurements that will be useful when defining new ways to specify locations. Many of these measurements involve measuring the length of specific line segments and angles between line segments in three dimensions.

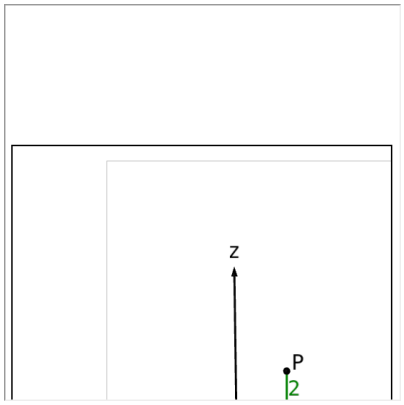
Preview Activity 9.9.1 An angle in the xy-plane is in standard position if the initial side of the angle is on the positive x-axis and is measured with the positive direction going counterclockwise.

(a) If an angle θ is in standard position, for which quadrants will $\sin(\theta)$ be positive? For which quadrants will $\sin(\theta)$ be negative? Which angles correspond to $\sin(\theta) = 0$?

(b) If an angle θ is in standard position, for which quadrants will $\cos(\theta)$ be positive? For which quadrants will $\cos(\theta)$ be negative? Which angles correspond to $\cos(\theta) = 0$?

(c) If an angle θ is in standard position, for which quadrants will $\tan(\theta)$ be positive? For which quadrants will $\tan(\theta)$ be negative? Which angles correspond to $\tan(\theta) = 0$? Which angles correspond to $\tan(\theta)$ being undefined?

(d) Consider the plot of point P in Figure 9.9.1.



Standalone
In Context
Embed

Figure 9.9.1 A plot of point P with y, z, and r labeled

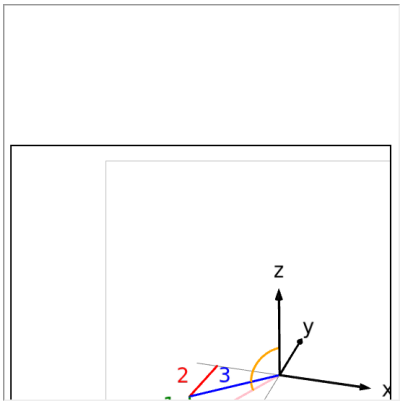
Use trigonometry and the distance formula to find the following:

(i) The x-coordinate of P.

(ii) The angle in the xy-plane between the x-axis and the line segment (in blue) labeled 5 that is from the origin to a point directly below P.

(iii) The distance between the origin and P.

(e) Consider the plot of point P in Figure 9.9.2.



Standalone
In Context
Embed

Figure 9.9.2 A plot of point P with y, z, and r labeled

Use trigonometry and the distance formula to find the following:

(i) The x-coordinate of P.

(ii) The angle between the z-axis and the pink line segment from the origin to P.

(iii) The distance between the origin and P.

In this section we will define and work with new ways to specify the location of a point. These new measurements will be useful as coordinate systems and as a way of thinking about how to describe different shapes relative to rotational or other symmetries.

9.9.2 Polar Coordinates

We previously defined the rectangular coordinates of a point to be signed distances from the point to the axes as shown below. Remember that the sign on the coordinate tells you whether to go above or below the other axis.

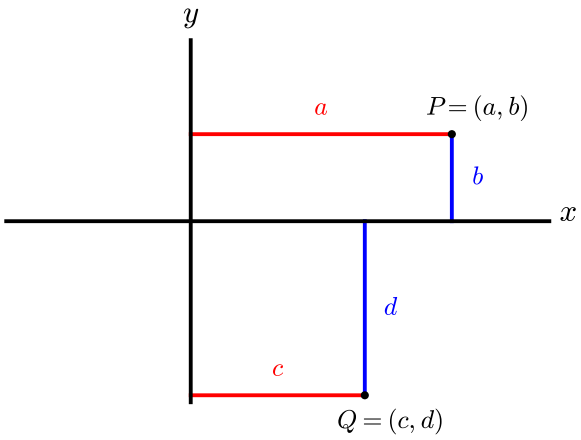


Figure 9.9.3 A plot with the 2D coordinate measurements labeled

We define the **polar coordinates** of a point P in two dimensions to be (r, θ), where r is the signed distance from the origin to P and θ is the counterclockwise angle from the positive horizontal axis to the line segment connecting the origin and P. We use the sign on the r to indicate if the point is locked ahead or behind when facing the θ direction at the origin.

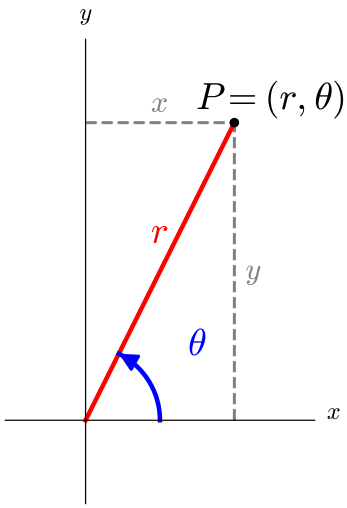


Figure 9.9.4 A plot with polar coordinate measurements labeled

A helpful way to visualize a location based on polar coordinates is to stand at the origin, facing the positive horizontal axis; turn an angle of θ counterclockwise; and move r in the direction you are facing.

Example 9.9.5 Our first example will look at the location given by $P : (r, \theta) = (2, \frac{\pi}{2})$. To understand where this point is, we do the following steps: from the origin, face the positive x-axis; turn $\frac{\pi}{2}$ radians counterclockwise (to the positive y axis); and move two units in the direction you are now facing.

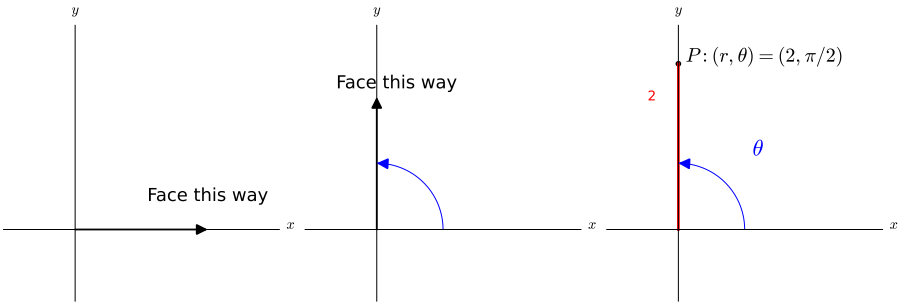


Figure 9.9.6 Starting orientation of polar coordinates

Figure 9.9.7 Rotate by $\frac{\pi}{2}$ counterclockwise

Figure 9.9.8 Move two units in this direction

Geometrically, we can see that this location would have rectangular coordinates of $(x, y) = (0, 2)$. While geometry is useful for understanding meaning, it is rarely wise to compute things like coordinates using only a graph. Later in this section, we will talk about algebraic tools that will allow us to convert between rectangular and polar coordinates.

If we considered the point with polar coordinates $Q : (r, \theta) = (-2, -\frac{\pi}{2})$, we would do the following geometric instructions:

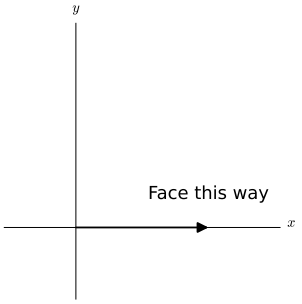


Figure 9.9.9 Starting orientation of polar coordinates

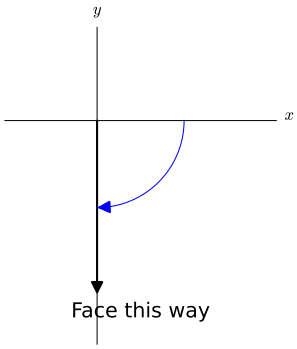


Figure 9.9.10 Rotate by $\frac{\pi}{2}$ clockwise

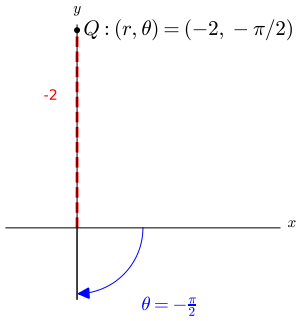


Figure 9.9.11 Move two units backwards while facing this direction

Notice that both $P : (r, \theta) = (2, \frac{\pi}{2})$ and $Q : (r, \theta) = (-2, -\frac{\pi}{2})$ correspond to the location with rectangular coordinates $(0, 2)$. This example highlights how polar coordinates are measured and how there is ***not*** a unique set of polar coordinates for a location. ▲

If you look at Figure 9.9.4, you can see a right triangle in the first quadrant with base x and height y. The angle the hypotenuse makes with the base is θ. From this, you can deduce the four relationships shown below, which are used to convert between rectangular and polar coordinates. It may be insightful to draw triangles of this type in each of the quadrants to confirm for yourself that these conversion formulas are valid in all four cases.

Converting between rectangular and polar coordinates.

From polar to rectangular Given polar coordinates (r, θ) of a point P, the rectangular coordinates (x, y) of P satisfy

$$x = r\cos(\theta) \qquad y = r\sin(\theta).$$

From rectangular to polar Given rectangular coordinates (x, y) of a point P, the polar coordinates (r, θ) of P satisfy

$$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x},$$

assuming $x \neq 0$.

It is important to notice that the equations used in the conversion from polar to rectangular work to convert coordinates of a point for *any* values of r and θ, including negative r-values and values of θ outside the interval $[0, 2\pi]$. The conversion equations to go from rectangular to polar require you to interpret the results of your calculation by choosing from multiple suitable values for r and θ. This is because we are not able to solve uniquely for r and θ in terms of x and y. This is because there is not a unique set of polar coordinates for a location. A point can have many different polar coordinates that refer to the same location.

For example, the point with rectangular coordinates $(-1, -1)$ can be given the polar coordinates $(r, \theta) = (\sqrt{2}, \frac{5\pi}{4})$ or $(r, \theta) = (-\sqrt{2}, \frac{-\pi}{4})$ or $(r, \theta) = (\sqrt{2}, \frac{13\pi}{4})$ or $(r, \theta) = (\sqrt{2}, -\frac{3\pi}{4})$ or $(r, \theta) = (-\sqrt{2}, -\frac{7\pi}{4})$ (amongst many others). The good news is that when you need to choose appropriate r and θ, you can make proper choices based on your experience with quadrants and where

the point is (in terms of x and y).

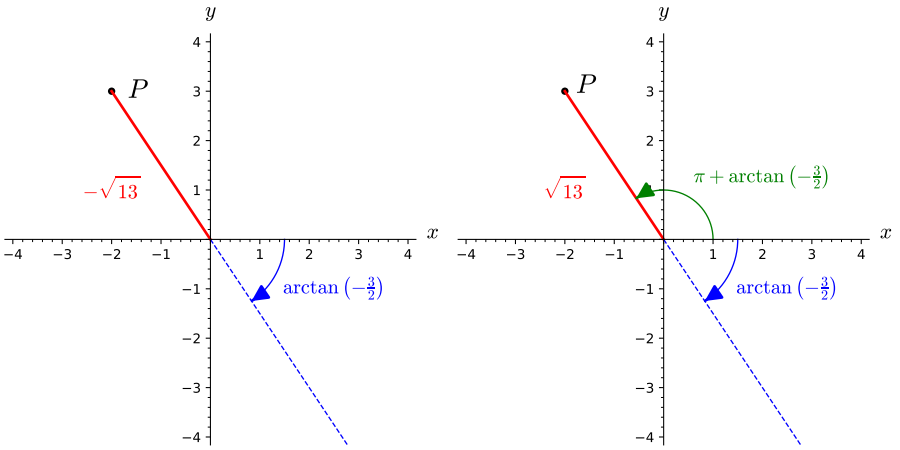
Example 9.9.12 In this example, we will look at how to convert the points with rectangular coordinates $P = (-2, 3)$ and $Q = (-3, -2)$ into polar coordinates. Note here that P is in the second quadrant and Q is in the third quadrant.

Using the conversion equations for $P = (-2, 3)$, we see that

$$r^2 = (-2)^2 + (3)^2 \qquad \tan(\theta) = -\frac{3}{2}$$

which means we need to pick r to be $\sqrt{13}$ or $-\sqrt{13}$. We will need to pick a appropriate angles for these possible r-values that satisfy $\tan(\theta) = -\frac{3}{2}$. There are angles in the fourth and second quadrants that will satisfy $\tan(\theta) = -\frac{3}{2}$, namely $\arctan(-\frac{3}{2})$ and $\pi + \arctan(-\frac{3}{2})$. You could also choose any angle coterminal to either of these angles. This is not a situation where we can arbitrarily combine one of the choices of r will work with any choice of θ. We must interpret these choices *together* to get a description for the correct location.

If we choose $r = \sqrt{13}$, then we need to select the angle that corresponds to the second quadrant, $\pi + \arctan(-\frac{3}{2})$, to get polar coordinates $P = (r, \theta) = (\sqrt{13}, \pi + \arctan(-\frac{3}{2}))$. If we want to not have to use the complementary angle to $\arctan(-\frac{3}{2})$, then we need to use the negative value of r to get $P = (r, \theta) = (-\sqrt{13}, \arctan(-\frac{3}{2}))$. Figure 9.9.13 illustrates these two different polar coordinates that identify the point P

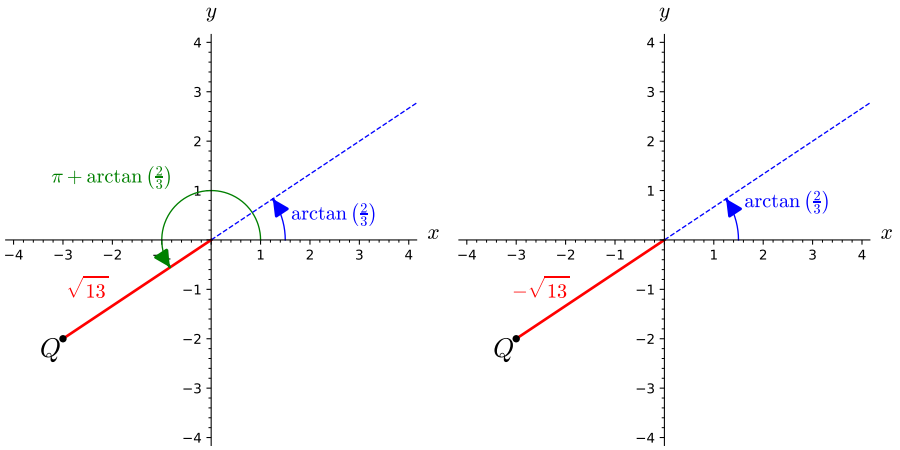


(a) A visualization of P with polar coordinates $(-\sqrt{13}, \arctan(-\frac{3}{2}))$

(b) A visualization of P with polar coordinates $(\sqrt{13}, \pi + \arctan(-\frac{3}{2}))$

Figure 9.9.13

Now we will look at how to convert $Q = (-3, -2)$ into polar coordinates. Similar to our arguments above, we find that $Q = (-3, -2)$ implies that $r^2 = 9 + 4 = 13$ and $\tan(\theta) = \frac{-2}{-3}$. If we want to consider a positive value for r and a value of θ between 0 and 2π, then we get $r = \sqrt{13}$ and $\theta = \pi + \arctan(\frac{2}{3})$. We could also consider a negative value for r and a value of θ between 0 and 2π, which would yield $r = -\sqrt{13}$ and $\theta = \arctan(\frac{2}{3})$. Figure 9.9.14 illustrates these two perspectives on finding polar coordinates for Q



(a) A visualization of Q with polar coordinates $(\sqrt{13}, \pi + \arctan(\frac{2}{3}))$

(b) A visualization of Q with polar coordinates $(-\sqrt{13}, \arctan(\frac{2}{3}))$

Figure 9.9.14

Remember that inverse trig functions have limited domains and ranges, the r-coordinate can be positive or negative, and polar coordinates do not have unique values for a given point. ▲

The following activity gives you an opportunity to practice converting between rectangular and polar coordinates.

Activity 9.9.2

(a) For each of the points listed below, you should:

- Graph the point on a set of axes and label how each of the rectangular coordinates is measured.
- Draw and label how the polar coordinates are measured for each point.
- Compute the exact value of r and θ for each point. Exact values include things like $\sqrt{3}$, arcsin(3/4), etc. Simplify trigonometric function values of any common angles you encounter, like $\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$.

(i) $(x, y) = (-3, -7)$

(ii) $(x, y) = (7, -3)$

(iii) $(x, y) = (\sqrt{5}, -2)$

(iv) $(x, y) = \left(\frac{3}{2}, -\frac{3}{2}\right)$

(v) $(x, y) = \left(\frac{3}{2}, \frac{3}{2}\right)$

(b) For each of the points listed below, you should:

- Graph the point on a set of axes and label how each of the polar coordinates is measured.
- Draw and label how the rectangular coordinates are measured for each point.
- Compute the exact value of x and y for each point. Exact values include things like $\sqrt{3}$, arcsin(3/4), etc. Simplify trigonometric function values of any common angles you encounter, like $\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$.

(i) $(r, \theta) = \left(1, -\frac{\pi}{6}\right)$

(ii) $(r, \theta) = \left(\sqrt{3}, -\arctan\left(\frac{\sqrt{2}}{2}\right)\right)$

(iii) $(r, \theta) = \left(-\sqrt{5}, 0\right)$

(iv) $(r, \theta) = \left(7, \frac{3\pi}{4}\right)$

(v) $(r, \theta) = \left(3, \pi - \frac{\pi}{6}\right)$

We now want to look at converting an equation between rectangular and polar coordinates. Remember that the graph of an equation is the set of all points that satisfy the equation. When we say that we want to convert an equation such as $x = y^2$ to polar coordinates, we mean that we want to find an equation in r and θ whose graph will be exactly the same set of points as $x = y^2$. This will be fairly easy because we have conversion equations that are solved explicitly and uniquely for x and y: $x = r\cos(\theta)$ and $y = r\sin(\theta)$.

Example 9.9.15 We can convert the equation $x = y^2$ to polar coordinates by substituting $x = r\cos(\theta)$ and $y = r\sin(\theta)$ to get

$$\begin{aligned} x &= y^2 \\ r\cos(\theta) &= (r\sin(\theta))^2 \\ r\cos(\theta) &= r^2\sin(\theta)^2 \\ \frac{\cos(\theta)}{\sin(\theta)^2} &= r \\ r &= \cot(\theta)\csc(\theta) \end{aligned}$$

This tells us that the (r, θ) points satisfying $r = \cot(\theta)\csc(\theta)$ are the same locations as the (x, y) points that satisfy $x = y^2$. ***Except***, we divided by $\sin(\theta)$, which will not give an equivalent equation if $\sin(\theta) = 0$. Notice that if $\theta = 0$ (or any integer multiple of π), then $\cot(\theta)\csc(\theta)$ does not exist. In a technical sense, we would either need to stop our algebra at $\cos(\theta) = r\sin(\theta)^2$ or use a piecewise defined expression such as

$$r = \begin{cases} 0 & \theta = k\pi \text{ for } k \in \mathbb{Z} \\ \cot(\theta)\csc(\theta) & \theta \neq k\pi \text{ for } k \in \mathbb{Z} \end{cases}$$

This example highlights the tricky part of converting equations: without careful algebraic steps, you can miss points that need to be included in the graphs/ equations or possibly get nonsensical statements. Often we will use geometric intuition and information to help convert equations between coordinate systems and make sense of the resulting expressions. ▲

The previous example showed that converting an equation from rectangular coordinates to polar coordinates is mostly straightforward, although we must pay attention to issues where expressions may become undefined. Our next example focuses on the conversion from polar to rectangular coordinates, which has additional subtleties.

Example 9.9.16 In this example, we will convert the equation $r = 3\cos(\theta)$ into rectangular coordinates, identify the shape of its graph, and plot its graph. Remember that we do not have conversion equations for r and θ, but we do have some equations relating the variables, such as $r^2 = x^2 + y^2$. This will allow us to convert to an expression in x and y that does not require interpretation.

We start by squaring both sides of our polar equation, since this will give us an expression on the left-hand side of the equation that will be easy to convert

to rectangular coordinates. This yields

$$r = 3\cos(\theta) \Rightarrow r^2 = 9(\cos(\theta))^2 \Rightarrow x^2 + y^2 = 9(\cos(\theta))^2$$

Our last equation is a mix of rectangular and polar coordinates so we are not done with our conversion to rectangular coordinates. Unfortunately the easiest way to convert $(\cos(\theta))^2$ to rectangular is to use

$$x = r\cos(\theta) \Rightarrow (\cos(\theta))^2 = \frac{x^2}{r^2} \Rightarrow (\cos(\theta))^2 = \frac{x^2}{x^2 + y^2}$$

which will give us a rational expression for our conversion. This would require some extra algebra and interpretation, so let's try a slightly different approach.

Instead of squaring both sides of our polar equation, we can start by multiplying both sides by r. You may have seen this as a better approach from the start or gained this insight after our work above.

After we multiply both sides of our original equation by r, we have

$$r = 3\cos(\theta) \Rightarrow r^2 = 3r\cos(\theta)$$

This equation is great because *both* the left and right hand sides have explicit conversions to rectangular coordinates that do not require separate interpretation of quadrants. Thus, we see that the (r, θ) points that satisfy $r = 3\cos(\theta)$ are the same locations as the (x, y) points that satisfy $x^2 + y^2 = 3x$.

We have now converted our polar equation to a rectangular equation. What shape will the graph of $x^2 + y^2 = 3x$ (or $r = 3\cos(\theta)$) be? We can use the algebraic technique of completing the square to transform our equation into the the standard form of a circle: $(x - h)^2 + (y - k)^2 = R^2$.

$$\begin{aligned} x^2 \qquad\quad + y^2 &= 3x \\ x^2 - 3x \qquad + y^2 &= 0 \\ x^2 - 3x + \frac{9}{4} + y^2 &= 0 + \frac{9}{4} \\ \left(x - \frac{3}{2}\right)^2 + (y - 0)^2 &= \left(\frac{3}{2}\right)^2 \end{aligned}$$

We can see now that this circle will have center $(\frac{3}{2}, 0)$ and radius $\frac{3}{2}$. The graph of $r = 3\cos(\theta)$ is the same as $x^2 + y^2 = 3x$ or $(x - \frac{3}{2})^2 + (y - 0)^2 = \left(\frac{3}{2}\right)^2$, which is plotted below.

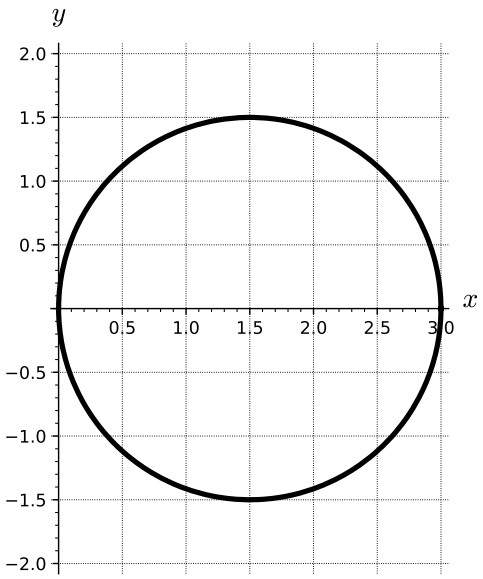


Figure 9.9.17 A plot of $r = 3\cos(\theta)$

While many people will find the rectangular equation more comfortable to graph and work with, this conversion from polar to rectangular is not always needed. You can find many resources on plotting points of a polar equation without converting to rectangular coordinates, especially when one coordinate (typically r) is solved explicitly in terms of the other coordinate. ▲

You can see that both of the rectangular and polar coordinates of the points in part a.iv, part 9.9.2.a.v, and part 9.9.2.b.ii will satisfy the equations of Example 9.9.16. The plot below shows how you can visualize part 9.9.2.b.ii.

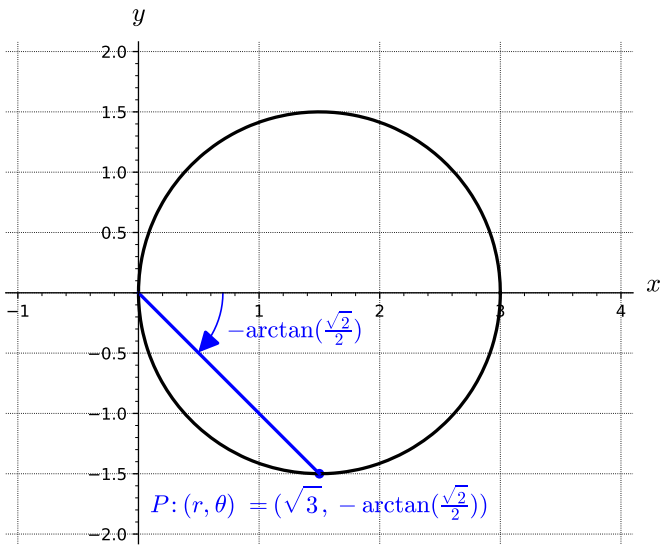


Figure 9.9.18 A plot of $r = 3\cos(\theta)$ with the polar coordinate measurements of part 9.9.2.b.ii

Activity 9.9.3 This activity gives you an opportunity to graph some equations in polar coordinates where r is expressed as a function of θ.

(a) The following activity will take you through how to create a graph of an

equation in r and θ.

(i) Draw a graph of $y = \cos(x)$ for at least $0 \leq x \leq 2\pi$.

(ii) Consider the polar equation $r = \cos(\theta)$. Complete the table below by computing the value of r for each value of θ for this equation.

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π	$\frac{5\pi}{4}$	$\frac{3\pi}{2}$	$\frac{7\pi}{4}$	2π
r											

(iii) Plot each of the eleven points you found in the previous part on the polar plane and connect them to make a plot of the graph of $r = \cos(\theta)$.

(b) Graph the equation $r = 1 + \cos(\theta)$ in the polar plane.

(c) Graph the equation $r = 1 - \cos(\theta)$ in the polar plane.

(d) What are the points of intersections for the graphs of $r = 1 + \cos(\theta)$ and $r = 1 - \cos(\theta)$

Now that you have some experience with plotting polar equations, the next activity will allow you to practice converting between polar and rectangular forms and seeing how graphs can help you with understanding these equations.

Activity 9.9.4 Below are several equations, some in polar form and some in rectangular form. For each equation, complete the following steps:

i. Convert the equation to the other coordinate system (either rectangular to polar or polar to rectangular).

ii. State the shape of the graph for the equation and any other information needed to graph

iii. Graph the given equation and write a sentence explaining how the graph helps understand the converted equation you found in the first step.

(a) $r = 2$

(b) $\theta = \frac{\pi}{3}$

(c) $x = -1$

(d) $y = 2$

(e) $y = -x$

(f) $r = \dfrac{3}{1 - 2\cos(\theta)}$

Hint. You may find it useful to have an intermediate step with an equation involving both r and x and completing the square as in Example 9.9.16 will likely be useful as well.

In Figure 9.2.1, you saw how a rectangular grid makes it easy to determine the rectangular coordinates of a point. A polar grid gives the same tools by plotting an array of constant coordinate graphs. An example of a polar grid is given in Figure 9.9.19, and thinking about this polar grid should help you answer the questions in Activity 9.9.5

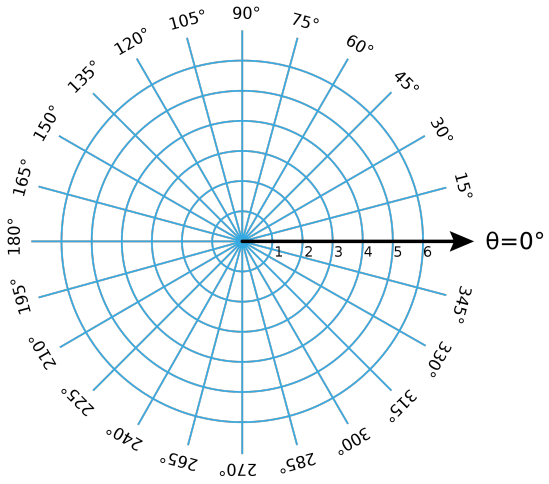


Figure 9.9.19 A polar grid (from WikiMedia)

Activity 9.9.5

(a) Write a couple of sentences describing the graph of $r = c$, where c is a constant. You should consider what values of c will correspond to different shapes.

(b) Write a couple of sentences describing the graph of $\theta = d$, where d is a constant. You should consider what values of d will correspond to different shapes.

The last idea of conversion we will look at between rectangular and polar coordinates is how to convert regions or inequalities between coordinate systems. Just as a graph of a an equation is a plot of the set of points that satisfies the equation, the graph of an inequality (or set of inequalities) is the region of points that satisfy the inequality.

Example 9.9.20 The graph of the inequality $-1 \leq x \leq 1$ is the vertical strip of points with horizontal coordinate between -1 and 1 (inclusive) and has graph shown in Figure 9.9.21.

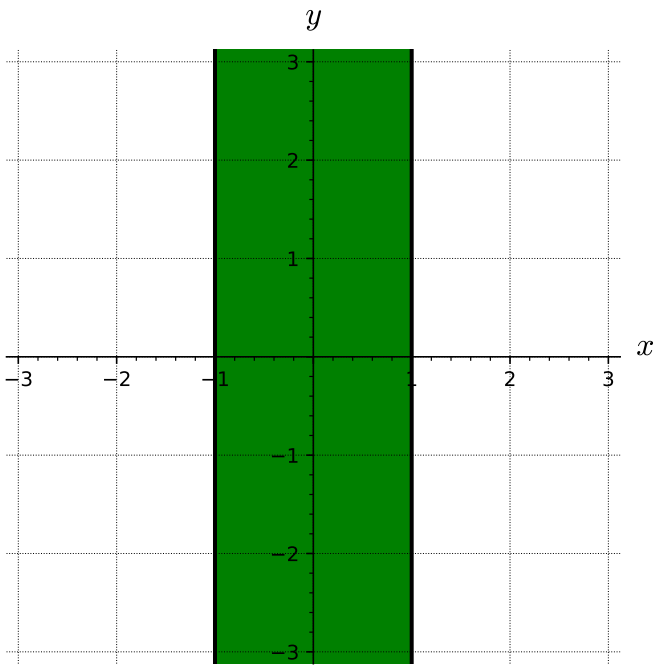
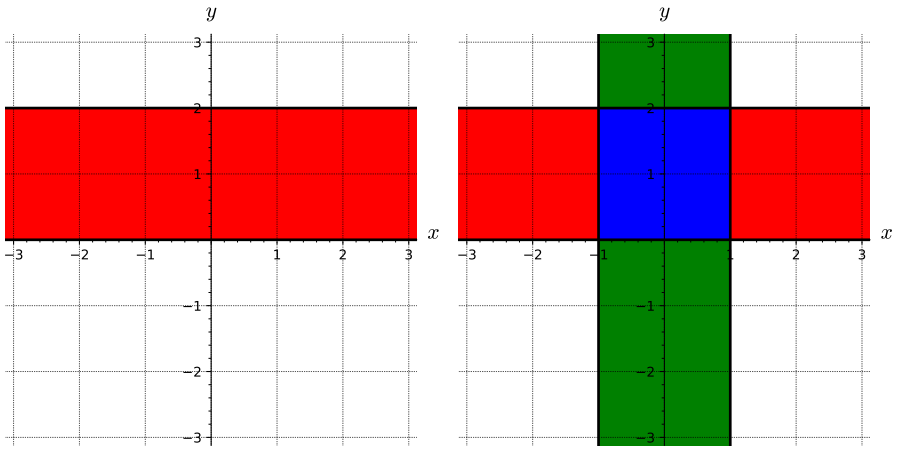


Figure 9.9.21 A plot of $-1 \leq x \leq 1$

If we consider the set of inequalities given by $-1 \leq x \leq 1$ and $0 \leq y \leq 2$, then we are looking for the region of points that satisfies both inequalities. The second inequality is shown in red in Figure 9.9.22(a) and the intersection of the two plots corresponds to the region described by $-1 \leq x \leq 1$ and $0 \leq y \leq 2$, which is plotted in blue Figure 9.9.22(b).



(a) A plot of $0 \leq y \leq 2$ **(b)** A plot of $-1 \leq x \leq 1$ and $0 \leq y \leq 2$

Figure 9.9.22

▲

Our final two activities on polar coordinates ask you to plot regions that satisfy inequalities in rectangular and polar coordinates.

Activity 9.9.6

(a) Graph the points that satisfy $y = x + 1$.

(b) On the same axes as in the previous part, graph the points that satisfy

$y \geq x + 1$.

Hint. Will the points above or below the line you drew in the first part of this activity satisfy this inequality?

(c) Graph the points that satisfy the inequality $x^2 + y^2 \leq 4$.

(d) Graph the points that satisfy both $x^2 + y^2 \leq 4$ and $y \geq x + 1$.

Activity 9.9.7

(a) Plot the region that corresponds to the polar inequality $1 \leq r \leq 3$. Write a sentence or two to describe this region.

(b) Plot the region that corresponds to the polar inequality $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$. Write a sentence or two to describe this region.

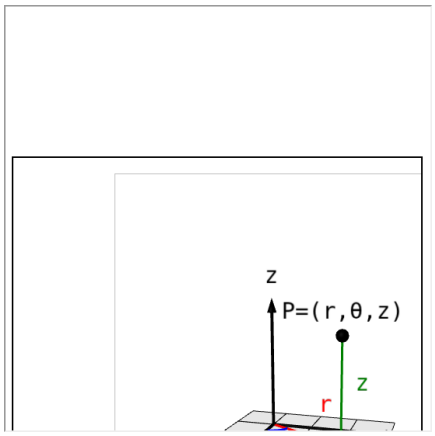
(c) Plot the region that corresponds to the polar inequalities $1 \leq r \leq 3$ and $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$. Write a sentence or two to describe this region.

(d) Plot the region that corresponds to the polar inequality $1 \leq r \leq 3\cos(\theta)$. Write a sentence or two to describe this region.

9.9.3 Cylindrical Coordinates

So far in this chapter, we have encountered two different coordinate systems in $\mathbb{R}^2$: rectangular and polar coordinates. Earlier in this section, we saw how polar coordinates form a convenient alternative in certain situations. In a similar way, there are two more coordinate systems in $\mathbb{R}^3$ that come from different ideas of rotational measurement. This subsection introduces cylindrical coordinates, which we can view as a vertical extension of polar coordinates to three dimensions. The next subsection will introduce spherical coordinates, which are useful for situations with significant rotational symmetry with respect to the origin.

Cylindrical coordinates are a coordinate system for $\mathbb{R}^3$ that consists of using polar coordinates, r and θ, in place of x and y coordinates. Cylindrical coordinates are given in the order of (r, θ, z) and can be described as "polar plus z". The z-coordinate is measured the same way as in rectangular coordinates (signed distance above or below the xy-plane) and polar coordinates are measured as a projection of the point in three dimensions onto the xy-plane. To convert between rectangular and cylindrical coordinates, we use the same conversion equations as between rectangular and polar coordinates (in two dimensions).



Standalone
In Context
Embed

Figure 9.9.23 A plot of point P with y, z, and r labeled

Converting between rectangular and cylindrical coordinates.

From cylindrical to rectangular

Given cylindrical coordinates (r, θ, z) of a point P, the rectangular coordinates (x, y, z) of P satisfy

$$x = r\cos(\theta) \qquad y = r\sin(\theta) \qquad z = z.$$

From rectangular to cylindrical

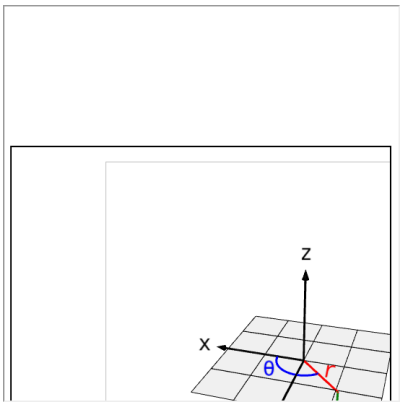
Given rectangular coordinates (x, y, z) of a point P, the cylindrical coordinates (r, θ, z) of P satisfy

$$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x} \qquad z = z$$

assuming $x \neq 0$.

Example 9.9.24

(a) For our first example, we will look at the point with cylindrical coordinates $(r, \theta, z) = \left(\sqrt{5}, \frac{3\pi}{4}, -2\right)$. This point is shown in Figure 9.9.25 with the measurements of the cylindrical coordinates.



Standalone
In Context
Embed

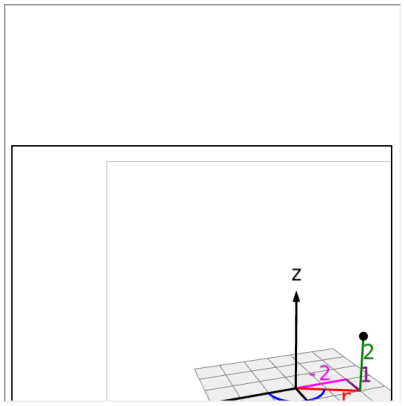
Figure 9.9.25 A plot of $(r, \theta, z) = (\sqrt{5}, \frac{3\pi}{4}, -2)$

To convert our cylindrical coordinates to rectangular, we can use $x = r\cos(\theta)$ and $y = r\sin(\theta)$. This gives rectangular coordinates of $(x, y, z) = \left(-\frac{\sqrt{10}}{2}, \frac{\sqrt{10}}{2}, -2\right)$.

(b) Next, we will find cylindrical coordinates for the point with rectangular coordinates $(x, y, z) = (-2, 1, 2)$. Just as the polar coordinates of a point in two dimensions are not unique, the cylindrical coordinates of a point in three dimensions will not be unique.

Using our conversion equations from rectangular to cylindrical, we get $r^2 = (-2)^2 + (1)^2 = 5$ and $\tan(\theta) = \frac{1}{-2}$. Remember that we must interpret the angle *and* our choice of r to make sure our coordinates describe a point in the second octant. In particular, we will choose a positive r value of $\sqrt{5}$ which means we will need to select $\theta = \arctan\left(-\frac{1}{2}\right) + \pi$.

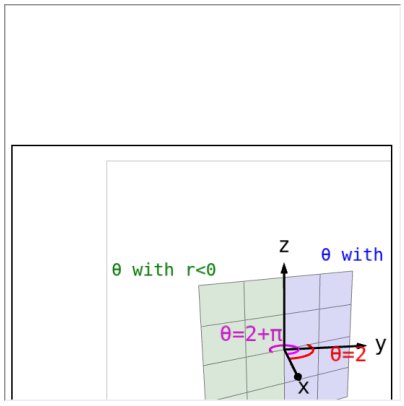
Alternatively, we can use $\cos(\theta) = \frac{x}{r}$ since arccos will give an output of an angle in the first or second quadrants. You can verify that $\theta = \arctan\left(-\frac{1}{2}\right) + \pi = \arccos\left(-\frac{2}{\sqrt{5}}\right)$. So rectangular coordinates $(x, y, z) = (-2, 1, 2)$ will correspond to cylindrical coordinates of $(r, \theta, z) = \left(\sqrt{5}, \arctan\left(-\frac{1}{2}\right) + \pi, 2\right)$ or $(r, \theta, z) = \left(\sqrt{5}, \arccos\left(-\frac{2}{\sqrt{5}}\right), 2\right)$.



Standalone
In Context
Embed

Figure 9.9.26 A plot of $(x, y, z) = (-2, 1, 2)$

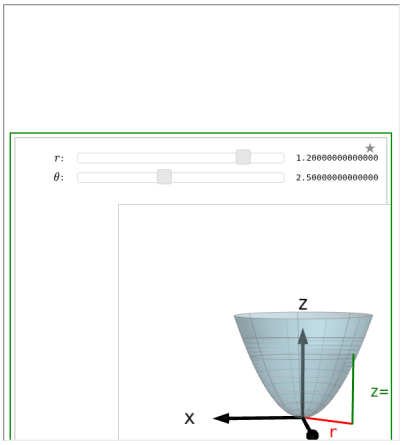
(c) When introducing polar coordinates, we saw how the graph of $\theta = 2$ is a line with slope $\tan(2)$. If we consider only positive r coordinates then we get a ray with an angle of 2 radians from the positive x-axis. If we wanted to consider the graph in three dimensions of $\theta = 2$ using *cylindrical coordinates*, then we have a plane containing the z-axis and extending in the $\theta = 2$ direction. This is a cylinder surface with the line given by $y = \tan(2)x$ as the generating curve and rulings parallel to the x-axis.



Standalone
In Context
Embed

Figure 9.9.27 A plot of $\theta = 2$ with positive r values in blue and negative r values in green

(d) In this example we will convert the elliptic paraboloid given by $z = x^2 + y^2$ to cylindrical coordinates. Algebraically, this is requires little work because we can substitute $r^2 = x^2 + y^2$ to obtain $z = r^2$. This should make sense for the graph of the elliptic paraboloid because the graph is rotationally symmetric around the z-axis because there is no explicit dependence on θ in the equation and the height of the surface above the xy-plane increases quadratically as the location moves away from the z-axis.



Standalone
In Context
Embed

Figure 9.9.28 A plot of $z = r^2$ with segments for r in red and $z = r^2$ in green

▲

The activities that close this subsection are meant to give you not only some algebraic experience with the cylindrical coordinate measurements but also some examples where you can make sense of your results from a geometric perspective.

Activity 9.9.8

(a) Convert $(x, y, z) = (1, 2, 3)$ to cylindrical coordinates.

(b) Use inequalities on cylindrical coordinates to describe the region given by the 6th octant.

(c) Convert the cone $z^2 = x^2 + y^2$ to cylindrical coordinates. Write a couple of sentences to make sense of how you can simplify your conversion and describe the shape of the graph in terms of z and r.

(d) Draw a plot of the surface in $\mathbb{R}^3$ with equation $r = 2$ in cylindrical coordinates. Write a couple of sentences about the shape and properties of this surface.

(e) Draw a plot of the region given by $0 \leq \theta \leq \pi, 0 \leq r \leq 2, 0 \leq z \leq r^2$. Write a couple of sentences about the shape and properties of this region.

Activity 9.9.9 In this activity, we graph some surfaces using cylindrical coordinates. To improve your intuition and test your understanding, you should first think about what each graph should look like before you plot it using appropriate technology.

(a) What familiar surface is described by the points in cylindrical coordinates with $r = 2$, $0 \leq \theta \leq 2\pi$, and $0 \leq z \leq 2$? How does this example suggest that we call these coordinates *cylindrical coordinates*? How does the

surface change if we restrict θ to $0 \leq \theta \leq \pi$?

(b) What familiar surface is described by the points in cylindrical coordinates with $\theta = 2$, $0 \leq r \leq 2$, and $0 \leq z \leq 2$?

(c) What familiar surface is described by the points in cylindrical coordinates with $z = 2$, $0 \leq \theta \leq 2\pi$, and $0 \leq r \leq 2$?

(d) Plot the graph of the cylindrical equation $z = r$, where $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 2$. What familiar surface results?

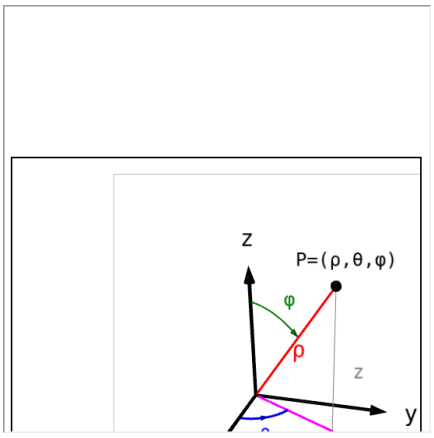
(e) Plot the graph of the cylindrical equation $z = \theta$ for $0 \leq \theta \leq 4\pi$. What does this surface look like?

9.9.4 Spherical Coordinates

Cylindrical coordinates used an angular measurement in the xy-plane as part of the description of location. Spherical coordinates use one linear measurement and two angular measurements to specify the location of a point. The three measurements used to define the spherical coordinates of a point in $\mathbb{R}^3$ are ρ (rho), θ, and ϕ (phi), where

- ρ is the distance from the point to the origin,
- θ has the same interpretation it does in polar coordinates, and
- ϕ is the (smallest) angle between the positive z-axis and the vector from the origin to the point.

We illustrate this in Figure 9.9.29, where the measures r and z from cylindrical coordinates are also depicted. You should convince yourself that any point in $\mathbb{R}^3$ can be represented in spherical coordinates with $\rho \geq 0$, $0 \leq \theta < 2\pi$, and $0 \leq \phi \leq \pi$.



Standalone
In Context
Embed

Figure 9.9.29 A plot of a point P with spherical coordinate measurements

Converting between rectangular and spherical coordinates.

From rectangular to spherical

Given rectangular coordinates (x, y, z) of a point P, the spherical coordinates (ρ, θ, ϕ) of P satisfy

$$\rho = \sqrt{x^2 + y^2 + z^2}$$
$$\tan(\theta) = \frac{y}{x}$$
$$\cos(\phi) = \frac{z}{\rho}$$

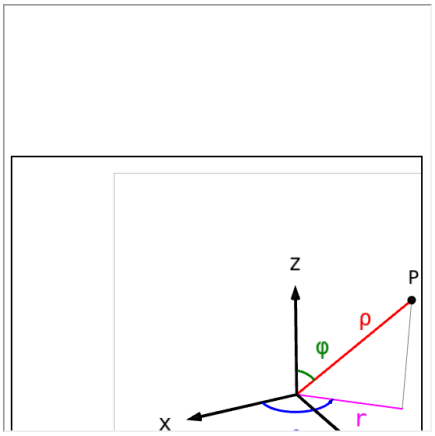
where in the second and third equations, we require $x \neq 0$ and $\rho \neq 0$.

From spherical to rectangular

Given spherical coordinates (ρ, θ, ϕ) of a point P, the rectangular coordinates (x, y, z) of P satisfy

$$x = \rho \sin(\phi) \cos(\theta)$$
$$y = \rho \sin(\phi) \sin(\theta)$$
$$z = \rho \cos(\phi).$$

Example 9.9.30 In this example, we investigate how to think about the spherical coordinates of a point and how to convert between spherical coordinates and rectangular coordinates. Consider the point P whose rectangular coordinates are $(-2, 2, \sqrt{8})$, which we illustrate in Figure 9.9.31



Standalone
In Context
Embed

Figure 9.9.31 A plot of a point $P = (-2, 2, \sqrt{8})$ with spherical coordinate measurements

What is the distance from P to the origin? The result is the value of ρ in the spherical coordinates of P, so

$$\rho = \sqrt{(-2)^2 + 2^2 + (\sqrt{8})^2} = 4$$

In order to find the θ-coordinate of P, we will need to determine the point that is the projection of P onto the xy-plane. We can use this projection to find the value of θ in the polar coordinates of the projection of P that lies in the plane. The projection of P onto the xy-plane is the point $(-2, 2, 0)$. As we saw earlier with polar and cylindrical coordinates, we can use $\tan(\theta) = \frac{y}{x}$ to find θ. For the point P, $\tan(\theta) = \frac{2}{-2} = -1$. Remember that you must interpret the

result of inverse trigonometric functions according to the particular geometric situation given in the problem. In particular, $\arctan(-1) = -\frac{\pi}{4}$ but we want will need to interpret θ as a an angle in the second quadrant. We use $\theta = \frac{3\pi}{4} = \pi + \arctan(-1)$.

Based on the illustration in Figure 9.9.31, we can use the right triangle with sides that correspond to the measurement of z, r, and ρ to find the angle ϕ. It may not be immediately clear where the angle ϕ is in this triangle, but Figure 9.9.32 shows how the alternate interior angles of the parallel lines associated with z will be equal. In other words, the upper right angle in our right triangle will be the same size as ϕ.

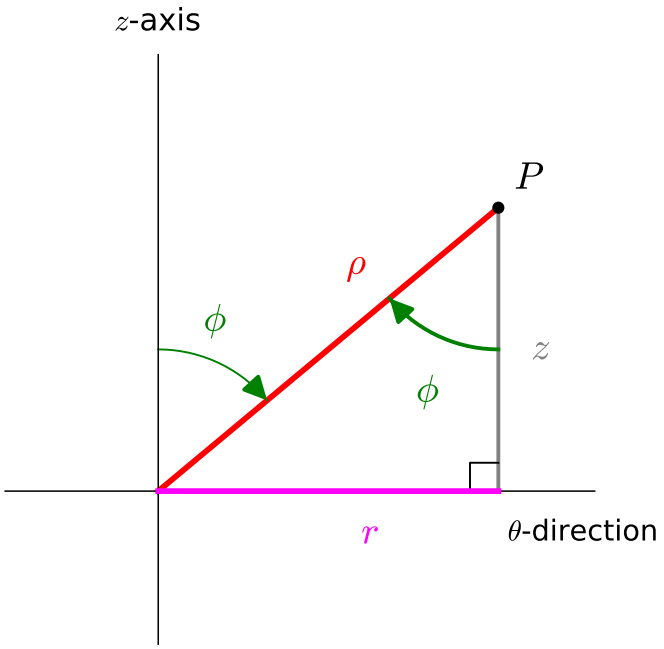


Figure 9.9.32 A 2D plot of a point P with spherical coordinate measurements in the plane of the measurement of θ

We can calculate the ϕ-coordinate of P using the property that $\cos(\phi) = \frac{z}{\rho} = \frac{\sqrt{8}}{4} = \frac{\sqrt{2}}{2}$, which means that $\phi = \frac{\pi}{4}$.

The point P with rectangular coordinates $(-2, 2, \sqrt{8})$ will be given in spherical coordinates as $(\rho, \theta, \phi) = \left(4, \frac{3\pi}{4}, \frac{\pi}{4}\right)$. ▲

Our first spherical coordinates activity focuses on understanding how spherical coordinates can be converted to rectangular coordinates and visualizing what the spherical coordinates measure.

Activity 9.9.10

(a) Use the conversion equations between rectangular and spherical coordinates to find the rectangular coordinates of the following points:

(i) $(\rho, \theta, \phi) = (3, \frac{\pi}{3}, \frac{\pi}{2})$
(ii) $(\rho, \theta, \phi) = (4, \frac{4\pi}{3}, \frac{3\pi}{4})$
(iii) $(\rho, \theta, \phi) = (\sqrt{7}, 20°, \pi)$

(b) Draw each of the points from part a and show how the spherical coordinates of each point is being measured. You should use your plots to make sense of the rectangular coordinate measurements that were your answer to part a.

Now that we have worked through how to think about *points* in spherical coordinates, we will do as we did with the other coordinate systems in this section and think about equations and surfaces.

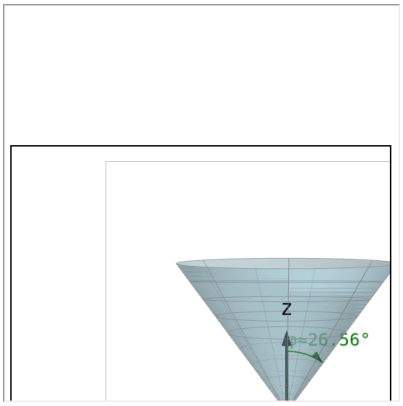
Example 9.9.33

(a) First, we will consider the surface defined by $\rho = 2$. Geometrically, we can think about $\rho = 2$ as the set of points that are two units away from the origin, which creates a sphere of radius 2. Algebraically, we can square both sides of our equation to get $\rho^2 = 4$ and since $\rho^2 = x^2 + y^2 + z^2$, the spherical coordinate equation $\rho = 2$ corresponds to the rectangular equation $x^2 + y^2 + z^2 = 4$. In fact, for all positive values of k, the surface $\rho = k$ will be a sphere of radius k.

(b) Next, we consider how to convert the equation of the half cone given by $z^2 = 2x^2 + 2y^2$ where $z \geq 0$ to spherical coordinates. Algebraically, we can use a couple of ideas from our conversion equations that include $z = \rho\cos(\phi)$ and $r^2 = x^2 + y^2 = \rho^2\sin(\phi)^2$. Specifically, this means that

$$\begin{aligned} z^2 &= 2x^2 + 2y^2 \\ z^2 &= 2\left(x^2 + y^2\right) \\ \rho^2\cos(\phi)^2 &= 2\left(\rho^2\sin(\phi)^2\right) \\ \tan(\phi)^2 &= \frac{1}{2} \end{aligned}$$

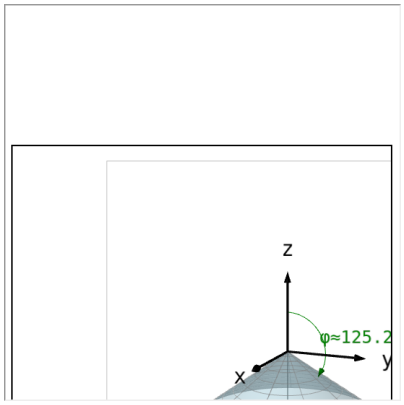
Since we are considering the part of the cone with $z \geq 0$, we can use $\phi = \arctan(\sqrt{\frac{1}{2}})$. This should not be that surprising that half cones centered on the z-axis correspond to a constant value of ϕ.



Standalone
In Context
Embed

Figure 9.9.34 A plot of the half cone $z^2 = 2x^2 + 2y^2$ with $z \geq 0$ shown in orange with measurement of ϕ

If we considered the bottom half of the cone given by $z^2 = 2x^2 + 2y^2$ where $z \leq 0$, we would need to find the angle ϕ between $\frac{\pi}{2}$ and π that satisfies $\tan(\phi) = \sqrt{\frac{1}{2}}$. Using complementary angles we see that the bottom half of the cone given by $z^2 = 2x^2 + 2y^2$ where $z \leq 0$ corresponds to the spherical equation $\phi = \pi - \arctan(\sqrt{\frac{1}{2}})$.



Standalone
In Context
Embed

Figure 9.9.35 A plot of the bottom half of the cone $z^2 = 2x^2 + 2y^2$ with $z \leq 0$ shown in orange with measurement of ϕ

(c) Finally, we consider how to convert the plane given by $z = x + y$ to spherical coordinates. We can use $x = \rho\sin(\phi)\cos(\theta)$, $y = \rho\sin(\phi)\sin(\theta)$, and $z = \rho\cos(\phi)$ to make the following algebraic simplifications:

$$\begin{aligned} z &= x + y \\ \rho\cos(\phi) &= \rho\cos(\theta)\sin(\phi) + \rho\sin(\theta)\sin(\phi) \\ \cos(\phi) &= \sin(\phi)\left(\cos(\theta) + \sin(\theta)\right) \\ \cot(\phi) &= \cos(\theta) + \sin(\theta) \end{aligned}$$

If the graph of $\cot(\phi) = \cos(\theta) + \sin(\theta)$ seems unfamiliar, that is normal. This is likely not an expression you have seen before and does not simplify or offer geometric insight. This example shows how easily converting equations between coordinate systems without some geometric reasoning for making the conversion can quickly turn into an exercise of algebra without much meaning.

▲

The next two activities help you develop your understanding of the angle ϕ in spherical coordinates and then ask you to think about some surfaces and how you can express them using spherical coordinates.

Activity 9.9.11 For many students, the measurement of ϕ is the most difficult and most unfamiliar measurement for spherical coordinates. In order to help understand the measurement of the ϕ-coordinate, we will look at the collection of points described by a constant value of ϕ. For each of the equations below, you should:

i. Draw a plot of the points that satisfy that equation. For some equations, your plot will be a path in space and others will correspond to a surface. You should draw how the ϕ-coordinate is measured and related to the graph.

ii. Write a sentence or two that describes the shape and features of your plot of the equation. You should mention any other descriptions that could be applied to your set of points, such as axes or coordinate planes.

(a) $\phi = 0$

(b) $\phi = 1$

(c) $\phi = \frac{\pi}{2}$

(d) $\phi = \frac{3\pi}{4}$

(e) $\phi = \pi$

Activity 9.9.12 For each of the surfaces described below, do the following:

i. Find a corresponding equation in spherical coordinates (involving only ρ, θ, and ϕ). You may want to consider algebraic or geometric approaches to these problems.

ii. Draw a plot of surface and label any important measurements in terms of the spherical coordinates.

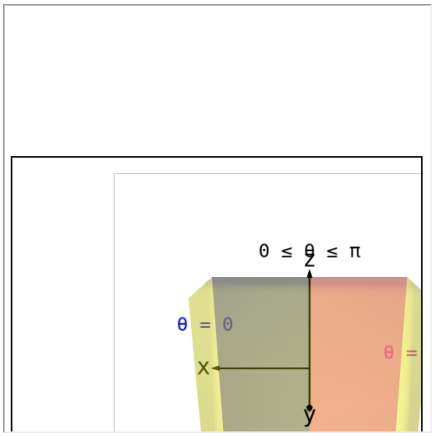
iii. Write a sentence or two about the corresponding equation in spherical coordinates and any assumptions made in your work.

(a) the xy-plane

(b) the xz-plane

(c) $z = x^2 + y^2$

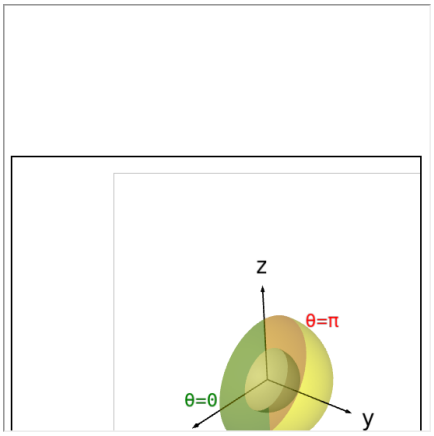
Example 9.9.36 In this example, we will develop a method for understanding the region given by the spherical coordinates $0 \leq \theta \leq \pi$, $1 \leq \rho \leq 2$, and $\frac{\pi}{4} \leq \phi \leq \pi$. We will look at these inequalities in sequence to figure out what the set of points that satisfies all three looks like. Our first inequality, $0 \leq \theta \leq \pi$, corresponds to all points in the first, second, fifth, and sixth octants. It may be tempting to draw this region with a spherical outer boundary. However, with no restrictions on the other coordinates, your plot should emphasize that the region continues throughout all of these octants (without an outer boundary).



Standalone
In Context
Embed

Figure 9.9.37 A plot of the region $0 \leq \theta \leq \pi$

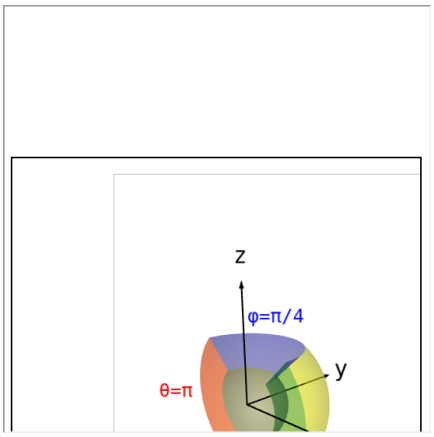
We now want to think about the set of points that satisfies $0 \leq \theta \leq \pi$ and $1 \leq \rho \leq 2$. The region given by $1 \leq \rho \leq 2$ is the set of points between the spheres of radius 1 and radius 2 centered at the origin. Hence, the region given by $0 \leq \theta \leq \pi$ and $1 \leq \rho \leq 2$ is the region between the spheres of radius 1 and radius 2 centered at the origin with non-negative y-coordinates.



Standalone
In Context
Embed

Figure 9.9.38 A plot of the region $0 \leq \theta \leq \pi$ and $1 \leq \rho \leq 2$

When we consider the additional restriction of $\frac{\pi}{4} \leq \phi \leq \pi$, we obtain the part of the region shown in Figure 9.9.38 that is *below* the cone given by $\phi = \frac{\pi}{4}$. This gives the region shown in Figure 9.9.39. Note that we have shown each of the boundary surfaces in different colors to illustrate how each of the upper or lower bounds on each of the coordinates gives rise to a different part of our region.



Standalone
In Context
Embed

Figure 9.9.39 A plot of the region $0 \leq \theta \leq \pi$, $1 \leq \rho \leq 2$, and $\frac{\pi}{4} \leq \phi \leq \pi$

▲

The next activity gives you a chance to apply some of the skills introduced in the previous example.

Activity 9.9.13

(a) Draw the region described by $\rho \leq 1$ and $0 \leq \phi \leq \frac{\pi}{2}$.

(b) Draw the region described by $x^2 + y^2 \leq 9$ with $x \leq y$ and $-1 \leq z \leq 3$

(c) Give a set of inequalities in spherical coordinates that describes the points in the seventh octant.

Example 9.9.40 For many people, the idea of latitude and longitude are the most familiar use of spherical coordinates. In particular, if you look at a globe as a map of the earth, then you will notice a grid of lines that run from the top to bottom or around the globe (horizontally). These lines measure latitude and longitude of different locations on the surface of the Earth. This latitude and longitude method of describing locations is actually spherical coordinates with

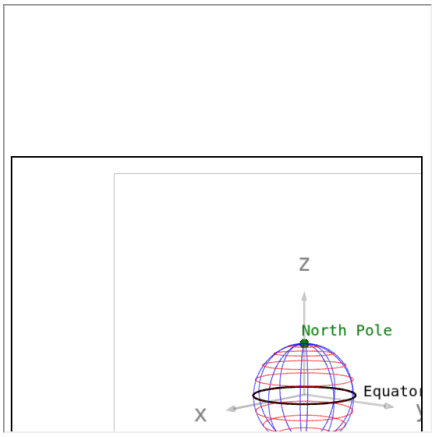
a (slightly incorrect) assumption that the surface of the Earth corresponds to a constant value of ρ.

If you look at the lines of longitude and latitude on a globe, you will see that these are angular measurements because each line in the grid is marked with a degree measurement. latitude describes how far a location is north or south of the equator and corresponds to a shifted measurement of the spherical coordinate ϕ. Specifically, a latitude of $0°$ corresponds to the equator, which corresponds to points on the surface of the sphere with $\phi = \frac{\pi}{2} = 90°$. The location identified with the "top" of the globe is often called the North Pole and corresponds to a latitude of $90°$ North and a spherical coordinate measurement of $\phi = 0$. Similarly, the "bottom" of the globe is called the South Pole and corresponds to a latitude of $90°$ South and a spherical coordinate measurement of $\phi = \pi = 180°$.

You can convert between latitude and ϕ measurements by

$$\phi = 90° + \text{latitude},$$

where north latitudes are considered negative and south latitudes are considered positive.



Standalone
In Context
Embed

Figure 9.9.41 A plot of lines of constant latitude (in red) and longitude (in blue), with labels for the North and South Poles as well as the equator

The blue curves in Figure 9.9.41 are lines with constant longitude. Longitude corresponds to an angular measurement of how far around (horizontally) on the surface of the sphere. Longitude corresponds to a shifted version of the θ coordinate. Lines of Longitude are called meridians and the Prime Meridian corresponds to the meridian that goes through the Greenwich Naval Observatory in Greenwich, England. If you were to flatten out the surface of the globe to make the latitude and longitude lines form a rectangular grid, then you get a Mercator projection map which you may recognize. While these flat maps are convienent there is a considerable amount of distortion in these maps as you get closer to the poles. As you can see in Figure 9.9.41, the lines with constant latitude are NOT all the same length, but need to be stretched out considerably near the poles to get a rectangular grid, as in Figure 9.9.42. Measuring the distortion of area (or volume) based on these coordinates will be an important idea in the later part of Chapter 12.

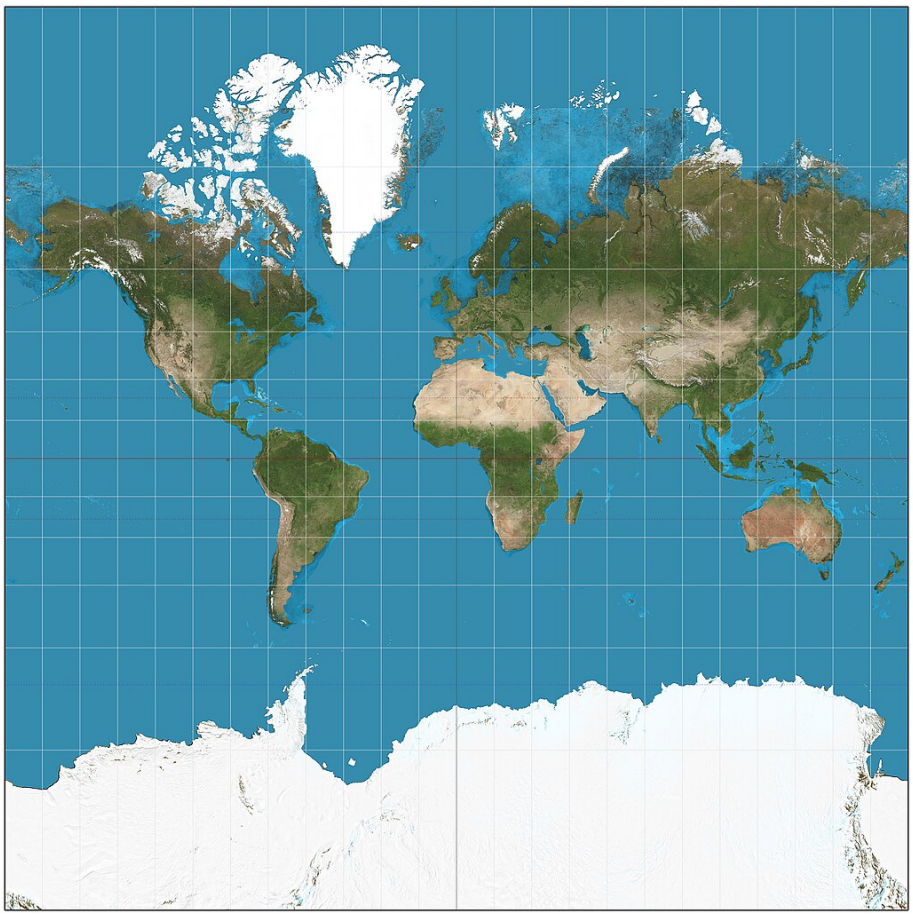
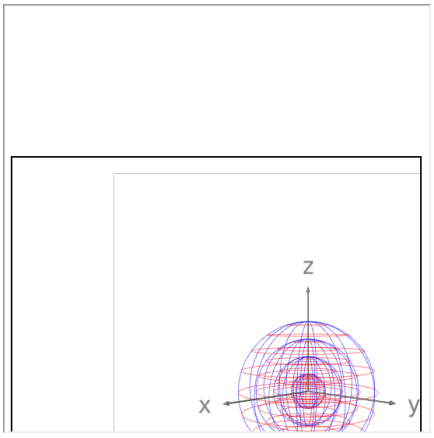


Figure 9.9.42 A Mercator projection of the Earth with latitude and longitude lines By Strebe - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=17700069

If we extended the grid of latitude and longitude to allow for multiple radii of spheres, we get a spherical coordinate grid as shown in Figure 9.9.43. You can see how visually cluttered this grid is and why you will likely not see this idea used in plotting again.



Standalone
In Context
Embed

Figure 9.9.43 A plot of a grid of spherical coordinates

▲

Summary

- The polar coordinates of a point P in $\mathbb{R}^2$ are (r, θ) where r is the distance from the origin to P and θ is the angle that the line segment from the origin to P makes with the positive x-axis. When P has rectangular coordinates (x, y), it follows that its polar coordinates are given by

$$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x}.$$

  When P has given polar coordinates (r, θ), its rectangular coordinates are

$$x = r\cos(\theta) \qquad y = r\sin(\theta).$$

- The cylindrical coordinates of a point P are (r, θ, z) where r is the distance from the origin to the projection of P onto the xy-plane, θ is the angle that the projection of P onto the xy-plane makes with the positive x-axis, and z is the vertical distance from P to the projection of P onto the xy-plane. When P has rectangular coordinates (x, y, z), it follows that its cylindrical coordinates are given by

$$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x} \qquad z = z.$$

  When P has given cylindrical coordinates (r, θ, z), its rectangular coordinates are

$$x = r\cos(\theta) \qquad y = r\sin(\theta) \qquad z = z.$$

- The spherical coordinates of a point P in 3-space are ρ (rho), θ, and ϕ (phi), where ρ is the distance from P to the origin, θ is the angle that the projection of P onto the xy-plane makes with the positive x-axis, and ϕ is the angle between the positive z axis and the vector from the origin to P. When P has rectangular coordinates (x, y, z), the spherical coordinates are given by

$$\rho^2 = x^2 + y^2 + z^2 \qquad \tan(\theta) = \frac{y}{x} \qquad \cos(\phi) = \frac{z}{\rho}.$$

  Given the point P in spherical coordinates (ρ, θ, ϕ), its rectangular coordinates are

$$x = \rho\sin(\phi)\cos(\theta) \qquad y = \rho\sin(\phi)\sin(\theta) \qquad z = \rho\cos(\phi).$$

- The conversion formulas above can be used to convert equations between coordinate systems.

9.9.5 Exercises

1. (a) The Cartesian coordinates of a point are $(-1, -\sqrt{3})$.

(i) Find polar coordinates (r, θ) of the point, where $r > 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____

$\theta =$ _____

(ii) Find polar coordinates (r, θ) of the point, where $r < 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____
$\theta =$ _____

(b) The Cartesian coordinates of a point are $(-2, 3)$.

(i) Find polar coordinates (r, θ) of the point, where $r > 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____
$\theta =$ _____

(ii) Find polar coordinates (r, θ) of the point, where $r < 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____
$\theta =$ _____

2. (a) You are given the point $(1, \pi/2)$ in polar coordinates.

(i) Find another pair of polar coordinates for this point such that $r > 0$ and $2\pi \leq \theta < 4\pi$.

$r =$ _____
$\theta =$ _____

(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____
$\theta =$ _____

(b) You are given the point $(-2, \pi/4)$ in polar coordinates.

(i) Find another pair of polar coordinates for this point such that $r > 0$ and $2\pi \leq \theta < 4\pi$.

$r =$ _____
$\theta =$ _____

(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $-2\pi \leq \theta < 0$.

$r =$ _____
$\theta =$ _____

(c) You are given the point $(3, 2)$ in polar coordinates.

(i) Find another pair of polar coordinates for this point such that $r > 0$ and $2\pi \leq \theta < 4\pi$.

$r =$ _____
$\theta =$ _____

(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $0 \leq \theta < 2\pi$.

$r =$ _____
$\theta =$ _____

3. Convert from rectangular to polar coordinates:

Note: Choose r and θ such that r is nonnegative and $0 \leq \theta < 2\pi$

(**a**) $(3, 0) \quad \Rightarrow_{(r,\theta)} \quad (_____, _____)$
(**b**) $(3, \frac{3}{\sqrt{3}}) \quad \Rightarrow_{(r,\theta)} \quad (_____, _____)$
(**c**) $(-2, 2) \quad \Rightarrow_{(r,\theta)} \quad (_____, _____)$
(**d**) $(-1, \sqrt{3}) \quad \Rightarrow_{(r,\theta)} \quad (_____, _____)$

4. Convert from polar to rectangular coordinates:

(**a**) $(5, \pi) \quad \Rightarrow_{(x,y)} \quad (_____, _____)$
(**b**) $(7, \frac{\pi}{6}) \quad \Rightarrow_{(x,y)} \quad (_____, _____)$
(**c**) $(5, \frac{\pi}{3}) \quad \Rightarrow_{(x,y)} \quad (_____, _____)$

5. Given the point $(2, -3)$ in x-y space, what is this point in polar coordinates? (use angles in degrees)

A $(56.3, 3.61)$

B $(3.61, 56.3)$

C $(3.61, -56.3)$

D $(3.61, 33.7)$

E $(3.61, -33.7)$

6. Decide if the points given in polar coordinates are the same. If they are the same, enter T . If they are different, enter F .

_ a.) $(6, \frac{\pi}{3}), (-6, \frac{-\pi}{3})$
_ b.) $(2, \frac{75\pi}{4}), (2, -\frac{75\pi}{4})$
_ c.) $(0, 6\pi), (0, \frac{9\pi}{4})$
_ d.) $(1, \frac{61\pi}{4}), (-1, \frac{\pi}{4})$
_ e.) $(17, \frac{62\pi}{3}), (-17, \frac{-\pi}{3})$
_ f.)$(6, 10\pi), (-6, 10\pi)$

7. Decide if the points given in polar coordinates are the same. If they are the same, enter T . If they are different, enter F .

_ a.) $(7, \frac{\pi}{3}), (-7, \frac{-\pi}{3})$
_ b.) $(2, \frac{51\pi}{4}), (2, -\frac{51\pi}{4})$
_ c.) $(0, 7\pi), (0, \frac{6\pi}{4})$
_ d.) $(1, \frac{117\pi}{4}), (-1, \frac{\pi}{4})$
_ e.) $(16, \frac{38\pi}{3}), (-16, \frac{-\pi}{3})$
_ f.)$(7, 6\pi), (-7, 6\pi)$

8. A curve with polar equation

$$r = \frac{8}{8\sin\theta + 25\cos\theta}$$

represents a line. Write this line in the given Cartesian form.

$y =$ _____

Note: Your answer should be a function of x .

9. Find a polar equation of the form $r = f(\theta)$ for the curve represented by the Cartesian equation $x = -y^2$.

Note: Since θ is not a symbol on your keyboard, use t in place of θ in your answer.

$r =$ _____

10. A Cartesian equation for the polar equation $r = 2$ can be written as:

$x^2 + y^2 =$ _____

11. Find the equation in polar coordinates of the line through the origin with slope $\frac{1}{3}$.

$\theta =$ _____

12. Which of the following two equations defines a horizontal line?

Write "T" if the equation defines a horizontal line, "F" otherwise.

(a) $r = 9\sec\theta$

(b) $r = 9\csc\theta$

13. Find an equation for the paraboloid $z = 2 - (x^2 + y^2)$ in cylindrical coordinates. (Type theta for θ in your answer.)

equation: _____

14. The following polar equation describes a circle in rectangular coordinates:

$$r = 10\sin\theta$$

Locate its center on the xy-plane, and find the circle's radius.
$(x_0, y_0) =$ (_____,_____)
$R =$_____

15. What are the rectangular coordinates of the point whose cylindrical coordinates are
$(r = 7,\ \theta = \frac{2\pi}{3},\ z = 7)$?
x= _____
y= _____
z= _____

16. Convert the following point from cylindrical to rectangular coordinates:

$$\left(5, \frac{4\pi}{3}, -0.5\right).$$

$(x, y, z) =$ _____
Usage:To enter a point, for example (x, y, z), type "(x, y, z)".

17. Convert the following point from rectangular to cylindrical coordinates:

$$(5, -5\sqrt{3}, 1).$$

$(r, \theta, z) =$ _____
Usage:To enter a point, for example (x, y, z), type "(x, y, z)".

18. What are the cylindrical coordinates of the point whose rectangular coordinates are $(x = 5,\ y = 4,\ z = -2)$?
$r =$ _____
$\theta =$ _____
$z =$ _____

19. What are the cylindrical coordinates of the point whose rectangular coordinates are $(x = -5,\ y = 3,\ z = 2)$?
$r =$ _____
$\theta =$ _____
$z =$ _____

20. Express the point given in Cartesian coordinates in cylindrical coordinates (r, θ, z).
A) $\left(3\left(\frac{\sqrt{3}}{2}\right),\ 3\left(\frac{1}{2}\right),\ -8\right)$ = (_____,_____,_____)
B) $\left(-3\left(\frac{\sqrt{3}}{2}\right), 3\left(\frac{1}{2}\right), -8\right)$ = (_____,_____,_____)
C) $\left(3\left(\frac{\sqrt{3}}{2}\right), -3\left(\frac{1}{2}\right), -8\right)$ = (_____,_____,_____)
D) $\left(-3\left(\frac{\sqrt{3}}{2}\right), -3\left(\frac{1}{2}\right), -8\right)$ = (_____,_____,_____)

21. Given the point $(1, 2, -2)$ in x-y-z space, what is this point in cylindrical coordinates? (use angles in degrees)

A $(3, 63.4, 63.4)$

B $(2.24, 63.4, 63.4)$

C $(2.24, 63.4, -63.4)$

D $(2.24, 63.4, -2)$

E $(3, 63.4, 2)$

22. What are the rectangular coordinates of the point whose cylindrical coordinates
are $(r = 4,\ \theta = 6,\ z = -2)$?
x = _____
y = _____
z = _____

23. What are the rectangular coordinates of the point whose cylindrical coordinates are
$(r = 6,\ \theta = \frac{3\pi}{9},\ z = -9)$?
x= _____
y= _____
z= _____

24. What are the rectangular coordinates of the point whose spherical coordinates are
$\left(4, \frac{1}{3}\pi, \frac{1}{2}\pi\right)$?
x = _____
y =_____
z =_____

25. What are the cylindrical coordinates of the point whose spherical coordinates are
$(4,\ -2,\ \frac{2\pi}{6})$?
r = _____
θ = _____
z= _____

26. Find an equation for the paraboloid $z = x^2 + y^2$ in spherical coordinates. (Enter rho, phi and theta for ρ, ϕ and θ, respectively.)
equation: _____

27. Convert the following point from spherical to rectangular coordinates:

$$\left(7, \frac{\pi}{3}, \frac{7\pi}{4}\right).$$

(x, y, z) = _____
Usage:To enter a point, for example (x, y, z), type "(x, y, z)".

28. Convert the following point from rectangular to spherical coordinates:

$$\left(\frac{5\sqrt{6}}{4}, \frac{5\sqrt{2}}{4}, \frac{5\sqrt{2}}{2}\right).$$

(ρ, θ, ϕ) = _____
Usage:To enter a point, for example (x, y, z), type "(x, y, z)".

29. What are the spherical coordinates of the point whose rectangular coordinates are
$(1,\ 5,\ 4)$?
ρ = _____
θ = _____
ϕ = _____

30. A point is given in spherical coordinates as $(1, 1, 1)$. What are the coordinates of the point in x-y-z coordinates? (use angles in radians)

A $(0.45, 0.71, 0.54)$

B $(0.02, 0.00, 1.00)$

C $(0.54, 0.45, 0.71)$

D $(0.29, 0.45, 0.84)$

E $(0.45, 0.29, 0.84)$

31. Express the point given in Cartesian coordinates in spherical coordinates (ρ, θ, ϕ). Note: you really only have to do the work for one, if you use a little geometry and your knowledge of the trig functions.

A) $(\frac{3}{4}\sqrt{2}, \frac{3}{4}\sqrt{2}, \frac{3}{2}\sqrt{3}) =$ _____

B)$(-\frac{3}{4}\sqrt{2}, \frac{3}{4}\sqrt{2}, \frac{3}{2}\sqrt{3}) =$ _____

C)$(\frac{3}{4}\sqrt{2}, -\frac{3}{4}\sqrt{2}, -\frac{3}{2}\sqrt{3}) =$ _____

D)$(-\frac{3}{4}\sqrt{2}, -\frac{3}{4}\sqrt{2}, -\frac{3}{2}\sqrt{3}) =$ _____

32. Suppose you are standing at a point on the equator of a sphere, parameterized by spherical coordinates θ, and ϕ. Let "north" point to the top of the sphere and "west" and "east" similarly be chosen as if the sphere were a globe.

What can you say about your initial ϕ coordinate?

$\phi_0 =$ _____

Let your initial θ coordinate be $\theta = a$. If you go east half of the way around the equator and halfway up toward the south pole along a longitude, what are your new θ and ϕ coordinates?

$\theta =$ _____

$\phi =$ _____

33. Match the given equation with the verbal description of the surface:

A. Circular Cylinder

B. Plane

C. Half plane

D. Elliptic or Circular Paraboloid

E. Cone

F. Sphere

(a) $\theta = \dfrac{\pi}{3}$

(b) $r = 2\cos(\theta)$

(c) $\rho = 4$

(d) $r = 4$

(e) $\rho\cos(\phi) = 4$

(f) $z = r^2$

(g) $r^2 + z^2 = 16$

(h) $\rho = 2\cos(\phi)$

(i) $\phi = \dfrac{\pi}{3}$

34. Consider a rectangular coordinate system with origin at the center of the earth, z-axis through the North Pole, and x-axis through the prime-meridian. Find the rectangular coordinates of London, England (51°32′N, 0°5′W). A minute is $1/60°$. Assume the earth is a sphere of radius $R = 6367$ km.

London has coordinates _____.

Usage:To enter a point, for example (x, y, z), type "(x, y, z)".

35. Match the given equation with the verbal description of the surface:

A. Circular Cylinder

B. Plane

C. Elliptic or Circular Paraboloid

D. Cone

E. Half plane

F. Sphere

(a) $r = 2\cos(\theta)$

(b) $\phi = \dfrac{\pi}{3}$

(c) $\rho\cos(\phi) = 4$

(d) $\rho = 2\cos(\phi)$

(e) $z = r^2$

(f) $\theta = \dfrac{\pi}{3}$

(g) $r = 4$

(h) $r^2 + z^2 = 16$

(i) $\rho = 4$

36. Consider the solid region S bounded above by the paraboloid $z = 16 - x^2 - y^2$ and below by the paraboloid $z = 3x^2 + 3y^2$.

(a) Describe parametrically the curve in $\mathbb{R}^3$ in which these two surfaces intersect.

(b) In terms of x and y, write an equation to describe the projection of the curve onto the xy-plane.

(c) What coordinate system do you think is most natural for setting up inequalities to describe the volume of the solid?

(d) Set up a set of inequalities in whichever coordinate system you choose that will describe the solid volume S.

(e) Try to draw a plot of the solid volume S by hand.

Chapter 10

Vector-Valued Functions of One Variable

10.1 Classic Calculus Approach

In this introduction, we will very briefly review some ideas and themes that ran through your work on single variable calculus because we will repeatedly use this same progression of ideas throughout our work in multivariable calculus.

The key ideas from your early calculus courses included limits, derivatives, and integrals, which were used these to measure change and accumulation of scalar functions. Each of these ideas and associated applications were likely introduced as a way to measure a specific feature where you approximate the feature of interest, then looked at how to improve this approximation. Limits offered a powerful tool to precisely describe how (and when) these approximations converge to the measurement of interest. We will call this process the **classic calculus approach**.

Key Idea 10.1.1 Classic Calculus Approach. *We highlight and reference the following steps as a fundamental approach to introducing and defining ideas in the calculus of both single variable and multivariable functions. We may abbreviate the Classic Calculus Approach as CCA .*

CCA Step 1: Approximate the measurement of interest

CCA Step 2: Quantify how the approximation you made changes under a finer scale

CCA Step 3: Use a limit to show the approximation converges to the measurement of interest

Before we start our work on the calculus of vector-valued functions, we will review the ideas limits, differentiation, and integration as well as look at how the Classic Calculus Approach was used in the definition of each of these concepts. This review is not meant to be comprehensive but is intended to ensure that all students are reminded of key concepts that will be vital for our development of calculus for new types of of functions. We will provide links to activities where you can review and explore the ideas from single-variable calculus.

Limits. The limit of a function measures what the output of the function approaches as an input gets closer to a particular value. Limits are examples of local measurements, since they describe a measurement that applies to a region around a particular value. Remember that a limit does *not* measure

what is happening at the particular input value; the limit describes behavior in a region *around* the particular input value.

In fact, limits are most useful and interesting when it does not make sense to evaluate the function at the particular input value. For instance,

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

is not a very interesting limit because the function we are taking the limit of will output the value of the limit. In other words, substituting in $x = 5$ into $x^2 - 2x+1$ will provide the same result as taking the limit as x approaches 5. These kinds of functions are called continuous functions. While continuous functions are important and nice, evaluating and understanding limits of continuous functions does not offer a different insight or understanding: what you see is what you get.

A more interesting limit is

$$\lim_{x\to 5} \frac{x^2 - 25}{x - 5}.$$

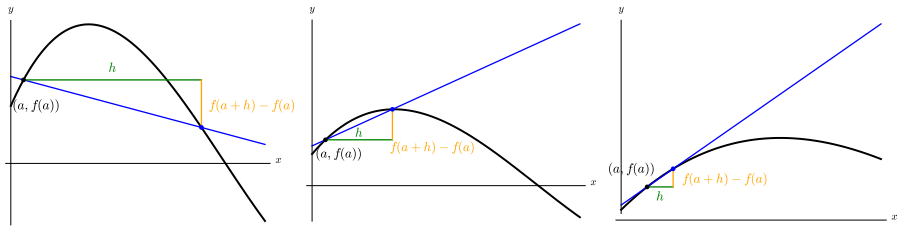
This limit is of an indeterminant form, which means that we will need to use other tools (rather than just attempting to substitute the particular input into the function). This limit and many of the limits you looked at in your definitions of calculus operations is of the form $\lim_{x\to a} f(x)\, g(x)$ where $\lim_{x\to a} f(x) = 0$ and $g(x)$ becomes unbounded as $x \to a$. Note that although this limit cannot be evaluated by substitution, we cannot immediately conclude that the limit does not exist. We will have to use other tools to find the limit value, if the limit value does exist.

Differentiation. The derivative of a function measures the rate of change of the value of a function at a particular input value. Geometrically, this is commonly visualized as the slope of a line tangent to the graph of the function at the point corresponding to the particular input value. The derivative is a local measurement because the rate of change or slope of the tangent line only models the behavior of the function for an arbitrarily small region around the input value.

The development of the definition of the derivative is an example of the Classic Calculus Approach. In order to understand how quickly the output of a function is changing at a particular input value, we approximate with the average rate of change over an interval, which corresponds to the slope of a secant line. This approximation is step 1 of the CCA.

As the points involved in the secant line become closer to the particular input value, the approximation becomes closer to the true rate of change at the particular input value. The slope of the secant line is the difference quotient: the change in the function value divided by the change in the inputs to the function. This quantification of how the slope of the secant line changes depending on the step size of the input is step 2 of the CCA.

By looking at the limit of this difference quotient as the change in the inputs approaches zero, we can compute the true rate of change for the function.



Algebraically, we can view this classic calculus approach as:

$$\text{rate of change of } f \approx \frac{f(a+h)-f(a)}{h} \qquad \text{(slope of secant line)}$$

$$\text{rate of change of } f = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \qquad \text{(slope of tangent line)}$$

Note that other derivative definitions correspond to different ways of thinking about how the two nearby points involved in the secant line approach the input of interest. The formula above approaches the point of interest from the right, but derivative formulas exist for the left difference or central difference methods. After evaluating the limit, all of these formulas calculate the same value for the derivative since there is only one instantaneous rate of change for the function at the particular input of interest.

You may not have noticed, but the limit formula for the derivative at a point is of the form of an "interesting" limit because

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

has

$$\lim_{h\to 0} f(a+h)-f(a) = 0 \quad \text{and} \quad \frac{1}{h} \to \infty \text{ as } h \to 0.$$

This means that we need other tools to evaluate the limit involved in derivatives.

Your first calculus course likely generalized the derivative at a point to $f'(x) = \frac{df}{dx}$, a function which outputs the rate of change of f for the input value x. This led to a collection of rules for efficiently calculating derivatives (e.g., power rule, sum/difference, product rule, chain rule, etc.) by thinking about the derivative $f'(x)$ as a function related to the function f. This perspective also leads to taking derivatives of derivatives, which is often used to measure features of graphs such as concavity.

We will take moment here to note another progression of ideas that was common in single variable calculus that will be very prevalent in our work with multivariable calculus. Most calculus topics follow the progression:

1. motivation given by the measuring a geometric feature
2. definition developed by the Classic Calculus Approach
3. stated definition of quantity involving limit form (which is difficult to calculate with, but explains what is being measured)
4. introduction of algebraic rules to aid in calculation or theorem about how to efficiently calculate quantity

The concept of derivatives in single variable calculus followed this progression of ideas and ended with learning many algebraic rules about how to calculate derivatives (sum/difference rule, coefficient rule, product rule, exponent rule, etc.)

Integration. The (definite) integral of a function on an interval measures the net signed area between the graph of the function and the horizontal axis. To approximate this area, we break the interval of interest into subintervals and use rectangles whose bases are the subintervals and whose heights are determined by the function's value at some point in the subinterval. Adding the area of these rectangles leads to an approximation of the exact area under the curve, which completes step 1 of the CCA.

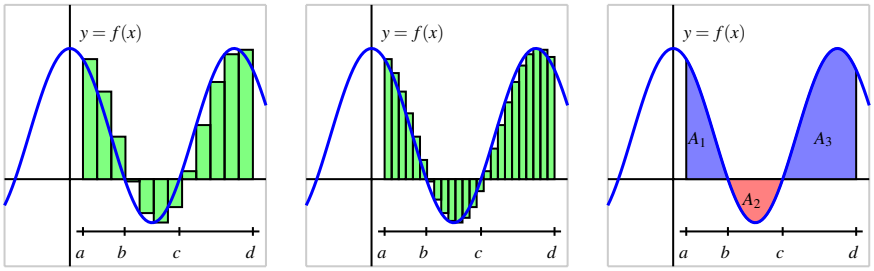
By using smaller and smaller subintervals, this rectangular approximation gets closer to the exact area under the curve. We can quantify this refinement and complete step 2 of the CCA with the following equation.

$$\text{area under the graph of } f \approx \sum_{i=0}^{n} f(x_i^*)\Delta x_i$$

This kind of sum is referred to as a Riemann sum for the function f over the interval from a to b.

Thus, taking the limit of the sum of the areas of these rectangles as the width of all of the subintervals goes to zero gives the exact area under the curve.

$$\text{area under the graph of } f = \lim_{\Delta x_i \to 0} \sum_{i=0}^{n} f(x_i^*)\Delta x_i$$



Your introduction to integrals likely investigated multiple rules for selecting which point you use to find the height of the rectangles in your approximation (e.g., right endpoint, left endpoint, midpoint, trapezoid etc.). However, a key idea is that the actual value of the integral does not depend on this choice.

If we use equally-sized subintervals in our approximation, then the Riemann sum argument above fits the form of our "interesting" limits. Specifically,

$$\lim_{\Delta x \to 0} \sum_{i=0}^{n} f(x_i^*)\,\Delta x$$

with

$$\lim_{\Delta x \to 0} \Delta x = \lim_{n \to \infty} \frac{b-a}{n} = 0 \quad \text{and} \quad \sum_{i=0}^{n} f(x_i^*) \to \infty \text{ as } n \to \infty$$

Note that increasing the number of subintervals corresponds exactly with decreasing the width of these equally-sized subintervals to zero.

You may have generalized the idea of the definite integral by looking at the accumulation of a function. Recovering Position from Velocity recalls how position can be thought of as the accumulation of change in velocity. The idea of accumulation of change (not limited to velocity leading to change in position) will be important in our future study.

Recovering Position from Velocity.

Suppose that an object moves along a number line (i.e., in one dimension) with velocity given as $v(t)$. The object's position is also a function of time, which we denote by $x(t)$ so that $x'(t) = v(t)$. We also

let $x_0 = x(0)$. In this framework, we can express the position at time $t = a$ as

$$x(a) = x_0 + \int_0^a v(t)\,dt.$$

Proof. In many physical problems, it is easier to monitor velocity rather than position because velocity is a local measurement, while position is *not* a local measurement. This may seem surprising at first, so we will consider our example of piloting a car. Tracking the position of a car for all times on a trip requires a coordinate system and constant measurement of distance between the car and reference objects such as axes or coordinate planes. This requires measurement beyond the car and an external view of the car. This is why systems like GPS satellites and receivers are necessary for accurate global navigation. In contrast, note that the car does *not* need something outside of itself to measure its speed because the speed is a measurement of change in position relative to a small step around the current location.

Because velocity is the change in position over time, velocity $= \frac{\Delta\text{position}}{\Delta\text{time}}$, we can measure a *change* in position by multiplying velocity by the change in time (Δposition = velocityΔtime) if velocity is constant. Over a small time interval, velocity changes very little, or is close to constant. Therefore, we can apply this relation and use a classic calculus approach to approximate, refine, and precisely state the position as a function of time.

We know the position of the object at time 0 because $x(0) = x_0$. Thus, we want to use the velocity to measure the change in position (on the number line) between $t = 0$ and time $t = a$. We approximate this change in position as Δposition = (velocity)(Δtime). To do this, we divide the interval $[0, a]$ into n equally-sized pieces of size $\Delta t = \frac{a}{n}$. The endpoints of these intervals are where $t_0 = 0, t_1 = \Delta t, \ldots, t_{n-1} = (n-1)\Delta t, t_n = a$. The position at $t = a$ is given by

$$x(a) = x_0 + \Delta x_1 + \ldots + \Delta x_n$$

where Δx_i is change in position from t_{i-1} to t_i.

On each interval $[t_{i-1}, t_i]$, we evaluate the velocity at some point in the interval, denoted $v(t_i^*)$, and then approximate the change in position by multiplying $v(t_i^*)$ by Δt. This gives us the estimate of the change in position on the interval as $\Delta x_i \approx v(t_i^*)\Delta t$. Our estimate of the position of our object at $t = a$ is then

$$x(a) \approx x_0 + \sum_{i=1}^{n} v(t_i^*)\Delta t \qquad \text{for some } t_i^* \in [t_{i-1}, t_i].$$

This approximation is step one of the classic calculus approach.

Step 2 of the classic calculus approach requires an understanding of how this approximation works on a finer scale. A "finer scale" in this case means a larger number n of subintervals. Increasing n means that Δt decreases. In other words, a finer scale means more terms in the sum but a smaller step size multiplier (Δt).

The estimate is now a Riemann sum where the function being evaluated in the Riemann sum is the velocity $v(t)$. Recognizing this leads to translating the Riemann sum into a definite integral as step three of the classic calculus approach. As $n \to \infty$ or $\Delta t \to 0$,

$$x_0 + \sum_{i=1}^{n} v(t_i^*)\Delta t \to x(a) = x_0 + \int_0^a v(t)\,dt.$$

In practice, this approximation is what GPS systems do at very fine level: use hyper-accurate timing and changes in relative velocities to satellites to give current position. ■

Accumulation will also apply to situations when working with density of a material as well. For instance, if considering a mining machine that harvests a mineral from the surface, the amount the machine harvests is the accumulation of the densities of the mineral along the path the machine followed. Mathematically this can be expressed as

$$\begin{aligned}\text{total harvested} &= \sum \text{amount in each section driven} \\ &\approx \sum (\text{density})(\text{length of the section})\end{aligned}$$

This Riemann sum becomes a definite integral of the density function in the classic calculus approach. This idea of accumulation of a function is something we will use in most chapters of the remainder of this text.

10.2 Vector-Valued Functions of One Variable

Motivating Questions

- What is a vector-valued function? What do we mean by the graph of a vector-valued function?
- What is a parameterization of a curve in $\mathbb{R}^2$? In $\mathbb{R}^3$? What can the parameterization of a curve tell us?

10.2.1 Introduction

In our previous work, we have seen several examples of curves in space, such as lines in three dimensions and conic sections (in two dimensions). For a line through a fixed point $\vec{r}_0$ in the direction of vector $\vec{v}$, we can express the line parametrically through the vector equation

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}.$$

We call $\vec{r}(t)$ a vector-valued function of one variable (or simply a vector-valued function, when the the single variable aspect is clear in context) because there is one scalar input (t) and the output of $\vec{r}$ is a vector that will vary based on the value of the scalar input. When the output vectors are graphed with initial point at the origin, the terminal points of these vector outputs trace out the line in space. Figure 9.6.5 and Figure 9.6.9 provide illustrations of this idea.

Similar to lines, other curves in space are classified as one-dimensional objects because there is only one way to move along this object. We refer to this direction as forward/backward depending on whether motion is toward larger (forward) or smaller (backward) values of the input variable.

We will similarly express the coordinates of points on a curve in terms of a single variable. This expression will describe movement along this path. For instance, the graph of an ellipse, such as the one shown in Figure 10.2.1, cannot be expressed with y as a function of x. However, we still classify an ellipse as a one-dimensional graph because at any point on the graph, there is only one dimension to move along the graph.

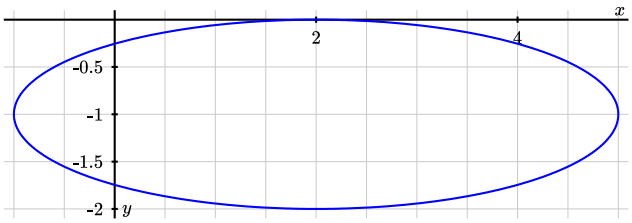


Figure 10.2.1 The ellipse $(x-2)^2/9 + (y+1)^2/1 = 1$

Vector-valued functions of one variable are the perfect vehicle for describing curves in general. We use vectors starting at the origin to identify points in space and the terminal points of these vectors trace out a curve in space. This approach allows us to draw a wide variety of graphs in two- and three-dimensional space. We can also describe curves in $\mathbb{R}^n$ for any n. This same approach to drawing curves will also allow us to represent slices of surfaces in space, which we will discuss more in Activity 10.2.4 and many times in Chapter 11.

Preview Activity 10.2.1 After graduating, you decide to start a self-driving car company because it didn't look too hard based on a few articles you scrolled past online. You decide to call your company *Steer Clear* and start working on the self-driving part of the car. Your first task is to understand how the location tracking equipment you bought online works. There are two parts to your location tracking system: a receiver box and a tracker that you put in the object you want to track.

(a) After putting the tracker in your pocket, you get in a car and drive around the receiver box. When you download the data about your drive, you notice that the software is outputting a vector from the receiver box to the tracker. The software also seems to love math, because these vectors are written in terms of trig functions.

Below are some of the output vectors from the software. Evaluate each of these vectors and draw these vectors with initial point at the origin on the axes in Figure 10.2.2.

- $\langle \cos(0), \sin(0) \rangle$
- $\langle \cos\left(\frac{\pi}{2}\right), \sin\left(\frac{\pi}{2}\right) \rangle$
- $\langle \cos(\pi), \sin(\pi) \rangle$
- $\langle \cos\left(\frac{3\pi}{2}\right), \sin\left(\frac{3\pi}{2}\right) \rangle$

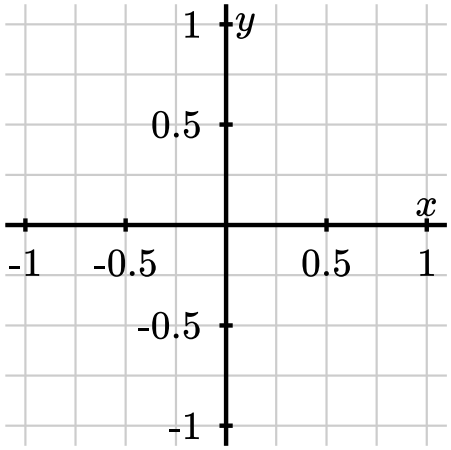


Figure 10.2.2 Two-dimensional axes for plotting points

(b) Below are a few more output vectors from your software. Evaluate each of these vectors and draw each of these vectors with initial point at the origin on the axes in Figure 10.2.2.

- $\langle \cos\left(\frac{\pi}{4}\right), \sin\left(\frac{\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{3\pi}{4}\right), \sin\left(\frac{3\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{5\pi}{4}\right), \sin\left(\frac{5\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{7\pi}{4}\right), \sin\left(\frac{7\pi}{4}\right) \rangle$

(c) Based on the data you are seeing from the two sets of vector outputs of your software, it appears that the output of your software is the set of vectors of the form $\langle \cos(t), \sin(t) \rangle$, where t goes from 0 to 2π. Add to Figure 10.2.2 a sketch the set of *terminal* points (only) of vectors $\langle \cos(t), \sin(t) \rangle$ plotted in standard position (with initial point at the origin). You should allow t to assume values from 0 to 2π. Write a couple sentences about what path your drive took.

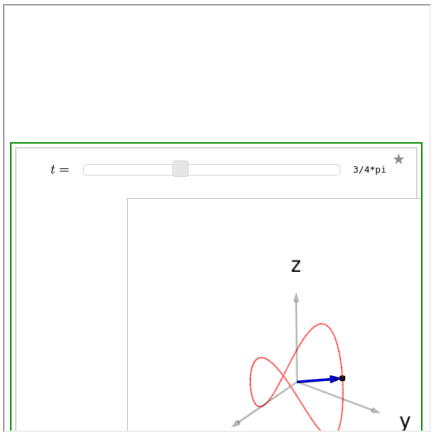
(d) What part of the path you drove is described by $\langle \cos(t), \sin(t) \rangle$, where t goes from 0 to π? What would the path be if you drove along $\langle \cos(t), \sin(t) \rangle$ for t from 0 to 4π?

In the preview activity, we saw how location data for how an object is moving along a curve can be given by vectors with a common initial point. In this section, we will continue to look at the implications of using this as an algebraic description of movement along a curve. In particular, we will look at algebraic rules and definitions for functions using the output vectors as points and we will conclude this section by generating several families of examples.

10.2.2 Vector-Valued Functions

As in Preview Activity 10.2.1, we can think of a point on the curve shown in Figure 10.2.3 as the terminal point of a vector starting at the origin. As the point we are looking at moves along the curve, the vector changes to terminate at the desired point. Figure 10.2.3 shows a curve in space with a vector from the origin to a highlighted point on the curve. Use the slider to change the location

on the curve to see how changing the scalar input gives different outputs with terminal points on the curve.



Standalone
In Context
Embed

Figure 10.2.3 A plot of a curve in space with a position vector shown

This correspondence between a location and a vector starting at the origin will allow us to think of a curve as a collection of terminal points of vectors emanating from the origin. We can describe an object traveling along this curve by defining a function $\vec{r}$ whose input is the variable t and whose output is the vector from the origin to the point on the curve at time t. In so doing, we have introduced a new type of function, one whose input is a scalar and whose output is a vector.

The terminal points of the output vectors of $\vec{r}$ will trace out the curve in space. From this perspective, the x, y, and z coordinates of the points on the curve are functions of time, t,

$$x = x(t), \quad y = y(t), \quad \text{and} \quad z = z(t),$$

and thus we have three coordinate functions that represent the collection of points on the curve. The variable t is called a **parameter** and the equations $x = x(t)$, $y = y(t)$, and $z = z(t)$ are called **parametric equations** (or a **parameterization of the curve**). The function $\vec{r}$ whose output is the vector from the origin to a point on the curve is defined by

$$\vec{r}(t) = \langle x(t), y(t), z(t) \rangle.$$

Definition 10.2.4 A **vector-valued function of one variable** is a function whose input is a real parameter t and whose output is a vector that depends on t. The **graph** of a vector-valued function is the set of all terminal points of the output vectors with their initial points at the origin.

Parametric equations for a curve in space are equations of the form

$$x = x(t), \quad y = y(t), \quad \text{and} \quad z = z(t)$$

that describe the (x, y, z) coordinates of a point on a curve in $\mathbb{R}^3$. ◆

Note that every set of parametric equations determines a vector-valued function of the form

$$\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$$

and every vector-valued function defines a set of parametric equations for a curve. Moreover, we can consider vector-valued functions as parameterizations in $\mathbb{R}^2$, $\mathbb{R}^3$, or any dimension.

As a reminder, in Section 9.6, we determined the parametric equations of a line in space using a point and a direction vector. For a nonlinear example, the curve in Figure 10.2.3 has the parametric equations

$$x(t) = \cos(t), \quad y(t) = \sin(t), \quad \text{and} \quad z(t) = \cos(t)\sin(t).$$

We represent this as the vector-valued function $\vec{r}$ given by

$$\vec{r}(t) = \langle \cos(t), \sin(t), \cos(t)\sin(t) \rangle.$$

Before we start working on parameterizations of curves, you may be asking yourself "Why can't we just have the output of our function be a point? Why do vectors need to be involved?!?" This is a good question because simpler tends to be better. However, using vectors means we will have significantly more tools at our disposal. These vector tools will make calculations easier and help us make geometric sense of measurements throughout the rest of this course. For example, we used a constant vector to indicate the direction of a line and this vector is related to the location vectors along the path. Remember that we already have vector tools that allow us to separate measurements into different pieces (like magnitude or direction), determine how much of a vector is in a particular direction, or determine how much of a vector is orthogonal to another direction.

Activity 10.2.2 The same curve can be represented with different parameterizations. Use appropriate technology to plot each of the curves generated by the following vector-valued functions for values of t from 0 to 2π. For each example, you should write a few sentences to explain how the graphs are alike and how they are different. Be sure to pay attention to how the parameter values are related to different points on the curve.

(a) $\vec{r}(t) = \langle \sin(t), \cos(t) \rangle$

(b) $\vec{r}(t) = \langle \sin(2t), \cos(2t) \rangle$

(c) $\vec{r}(t) = \langle \cos(t+\pi), \sin(t+\pi) \rangle$

(d) $\vec{r}(t) = \langle \cos(t^2), \sin(t^2) \rangle$

A mantra we will use when parameterizing curves is "There is more than one way to walk the same path", which means there is not a unique way to parameterize a curve in space. The examples in Activity 10.2.2 illustrate that a parameterization allows us not only to describe a curve in space as the graph of a vector-valued function of one variable but also to specify the direction and speed at which the graph is traversed as t changes. In the different parameterizations of the circle, we see that we can start at different points and move around the circle in either direction. In Section 10.3, we will begin to investigate the calculus of vector-valued functions. This will enable us to precisely quantify the direction, speed, and acceleration of an object moving along a curve in space. As such, describing curves parametrically will not only allow us to indicate the points on the curve itself, but also to describe motion along the curve.

While the ideas of a path or curve in space are closely related to a parameterization, we will take a moment here to clarify the precise meaning of these terms. A **curve in space** refers to the collection of points on the path. A curve does *not* have a specific direction to be traveled or even a particular way to move along points. A **parameterization** is an explicit description for how to travel along the points on the curve in terms of the parameter t. A common analogy we will use is that a curve in space would be represented by a bike

path or road while a parameterization would be a description of how you took a trip along this bike path or road as a function of time.

In general, we will avoid having parameterizations that pause at any point on the curve or reverse their direction along the curve because many of the measurements we consider in this chapter will be undefined when this occurs. It is common in practice to split parameterizations up into separate pieces that correspond to the intervals between where these pauses or direction changes happen.

The **graph** of a vector-valued function of one variable is the curve traced out by the endpoints of the output vectors, but this graph does not typically have parameter values labeled. This means that there is significant information lost when considering only the graph of a vector-valued function because no information about how the path was traveled/parameterized is available.

In the next subsection, we will cover some ideas that will be useful in parameterizing curves. First, however, we will address a couple of ideas related to the new category of functions we will be studying. Remember that the key idea behind functions is that for each input there is exactly one output assigned to that input. The **domain** of a function is the set of allowed inputs for the function. Sometimes a restricted set will be given that limits what inputs should be considered. Unless otherwise stated, however, the domain of a function will be the largest set of inputs for which the function is defined. Common ways for scalar functions to ***not*** be defined for a particular input include:

- dividing by zero,
- taking the square root of negative numbers, and
- restricted domains from other functions like arcsin or logarithms.

The domain of a vector-valued function of one variable will be the subset of $\mathbb{R}$ where *all* of the component functions are defined. In other words, the domain of a vector-valued function of one variable is the intersection of the domains for the component functions, because that is the set of inputs for which the vector output will be defined.

Example 10.2.5 Let

$$\vec{r}(t) = \left\langle \frac{-1}{t-2},\ \sqrt{25-t^2},\ \ln(t+1) \right\rangle.$$

The domains of the component functions are $\{x \in \mathbb{R} | x \neq 2\}$ for the x-component, $[-5, 5]$ for the y-component, and $(-1, \infty)$ for the z-component. The function $\vec{r}$ has an output for only the parameter values for which *all* three component functions are defined. Thus, the domain of $\vec{r}(t)$ is $(-1, 2) \cup (2, 5]$. ▲

10.2.3 Parameterizing Curves

We now discuss how to parameterize several types of curves in space. An important idea that will be used throughout the rest of the text is the way that different relationships of between coordinates can be expressed or understood, such as expressing one coordinate as a function of the other or relationships such as trig identities between the coordinate values.

Using parametric equations to define vector-valued functions in two dimensions is a more versatile way of describing curves than defining y as a function

of x. In fact, if $y = f(x)$ is a function of x, then we can parameterize the graph of f by

$$\vec{r}(t) = \langle t, f(t) \rangle.$$

Hence, the graph of every single-variable function can be described parametrically. In addition, as we saw in Preview Activity 10.2.1 and Activity 10.2.2, we can use vector-valued functions to represent curves in the plane that cannot be described with y as a function of x (or x as a function of y).

Example 10.2.6 Parameterizing $y = f(x)$. Consider the parabola given by $(x-2)^2 = 8(y+1)$. From Figure 10.2.7, you can see that the graph of $(x-2)^2 = 8(y+1)$ passes the vertical line test. Therefore, we can write the y-coordinate as a function of the x-coordinate. Specifically, we isolate y and obtain $y = f(x) = \frac{(x-2)^2}{8} - 1$. Using the idea above, we set $x = t$ to be our parameter. Now we can express both coordinates of points on our curve as functions of t. In particular, setting $x(t) = t$ means that we can write y as a function of t because $x = t$. Thus $x(t) = t$ and $y(x) = y(t) = \frac{(t-2)^2}{8} - 1$ will parameterize the parabola given by $(x-2)^2 = 8(y+1)$. This corresponds to the vector-valued function $\vec{r}(t) = \langle t, f(t) \rangle = \langle t, \frac{(t-2)^2}{8} - 1 \rangle$.

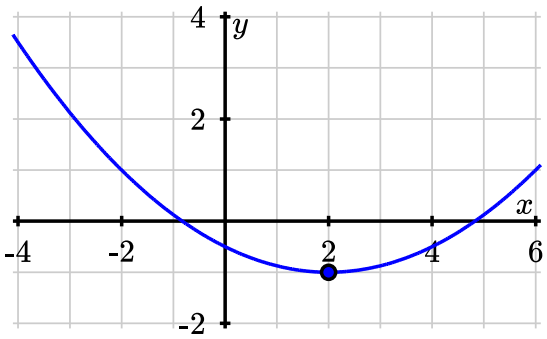


Figure 10.2.7 A plot of the parabola given by $(x-2)^2 = 8(y+1)$

Remember that there is *not* a unique way to parameterize a curve. This means that the choices we made above are not the only way to describe motion on this parabola. You could move along the curve with double the speed (relative to our original parameterization). In this case, the parameterization would be $\vec{r}(t) = \langle 2t, \frac{(2t-2)^2}{8} - 1 \rangle$.

Other choices for the parameterization might include wanting to have the vertex $(2, -1)$ of the parabola correspond to a specific parameter value, such as $t = 0$. One parameterization with this property would be $\vec{r}(t) = \langle t - 2, \frac{t^2-2}{2} \rangle$. ***Remember that there is more than one way to walk the same path.*** ▲

The next activity has you consider a few examples of curves generated by vector-valued functions, including in three dimensions.

Activity 10.2.3 Vector-valued functions can be used to generate many interesting curves. Graph each of the following using an appropriate technological tool, and then write one sentence for each function to describe the behavior of the resulting curve.

(a) $\vec{r}(t) = \langle t\cos(t), t\sin(t) \rangle$

(b) $\vec{r}(t) = \langle \sin(t)\cos(t), t\sin(t) \rangle$

(c) $\vec{r}(t) = \langle \sin(5t), \sin(4t) \rangle$

(d) $\vec{r}(t) = \langle t^2 \sin(t)\cos(t), 0.9t\cos(t^2), \sin(t)\rangle$. (Note that this defines a curve in three-dimensional space.)

Experiment with different formulas for $x(t)$ and $y(t)$ and ranges for t to see what other interesting curves you can generate. Share your best results with peers.

The next example helps illustrate how thinking of curves as graphs of vector-valued functions can help make it easier to algebraically address some ideas about the graphs of functions that are tricker when working from the perspective of $y = f(x)$.

Example 10.2.8

(a) Consider a one-to-one function, such as $f(x) = x^2 + 1$ on $[0, 2]$. The graph of this function is shown in Figure 10.2.9. In this scenario, we can also find the inverse function f^{-1} and graph it as well. Algebraically, this can sometimes be rather complex. However, when we parameterize the function as we did in Example 10.2.6, plotting the inverse function using a vector-valued function becomes routine. The method for doing this will rely on the fact that the graph of the inverse $y = f^{-1}(x)$ the reflection of the graph of $y = f(x)$ across the line $y = x$.

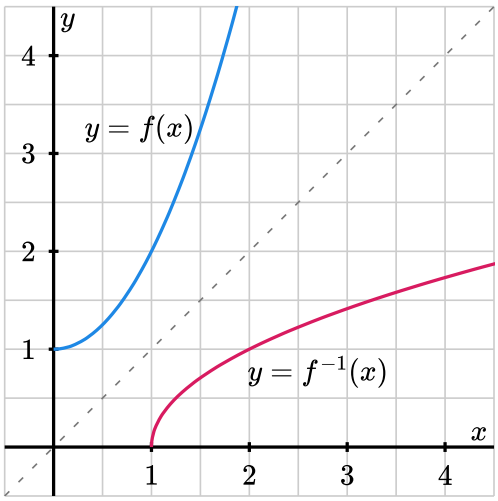


Figure 10.2.9 The graph of a function and its inverse

As in Example 10.2.6, we parameterize the graph of $f(x) = x^2 + 1$ on $[0, 2]$ as $\vec{r}(t) = \langle t, f(t)\rangle = \langle t, t^2 + 1\rangle$ for $0 \le t \le 2$. The points on the graph of the inverse function are determined by swapping the coordinates of the points on the graph of $y = f(x)$, so we swap the components of our vector-valued function that has parameterized the graph and consider $\vec{s}(t) = \langle f(t), t\rangle\rangle = \langle t^2 + 1, t\rangle$. Notice here that while we do not have an algebraic formula that allows us to take a given x and find $f^{-1}(x)$, we do have a nice way of graphing the inverse function with absolutely no algebra required.

(b) Now suppose that we need to think about the parabola with equation $x = y^2 + 1$. If we wanted to graph this using familiar tools from precalculus or single-variable calculus, we would need to graph *two* functions, since this curve does not pass the vertical line test. However, if we think parametrically, we can view y as the parameter here. This is similar to what we did in Example 10.2.6, but now we are thinking of a vector-valued

function $\vec{r}(t) = \langle x(t), y(t) \rangle$ where $y(t) = t$. With that settled, we then have that $x(t) = (y(t))^2 + 1 = t^2 + 1$. This leads us to $\vec{r}(t) = \langle t^2 + 1, t \rangle$.

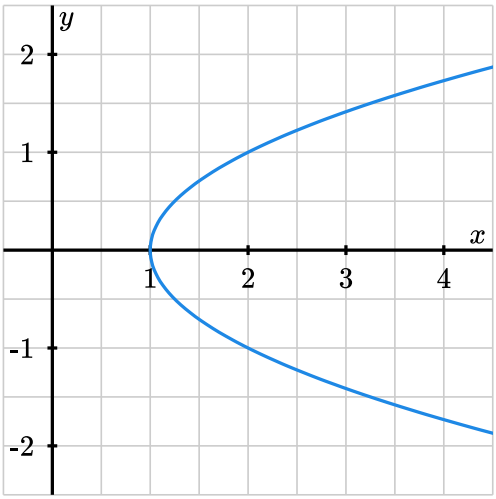


Figure 10.2.10 The graph of $x = y^2 + 1$

▲

The intersection of surfaces in space are often curves. Thus, we may determine parameterizations for these curves defined by satisfying multiple equations. For example, if we want to describe the intersection of the surface $z = \cos(x^2 + y^2)$ and the fundamental plane $y = 1$, then we can set $y = 1$ in $z = \cos(x^2 + y^2)$. We then let x be parameterized by the variable t. This tells us that the intersection of $z = \cos(x^2 + y^2)$ and the fundamental plane $y = 1$ is the curve parameterized as $\langle t, 1, \cos(t^2 + 1) \rangle$. The next activity guides you through similar questions for a more familiar surface. You may find it useful to use the idea of using a variable other than x as the parameter, as we did in the second part of Example 10.2.8

Activity 10.2.4 For this activity we will consider the paraboloid with equation $z = x^2 + y^2$. You do not need to know what the graph of $z = x^2 + y^2$ looks like at this point.

(a) Parameterize the intersection of the paraboloid with the fundamental plane $x = 2$. What type of curve is the intersection?

(b) Parameterize the intersection of the paraboloid with the fundamental plane $y = -1$. What type of curve is the intersection?

(c) Parameterize the intersection of the paraboloid with the fundamental plane $z = 25$. What type of curve is the intersection?

(d) How do your responses change to all three of the preceding questions if you instead consider the surface defined by $z = x^2 - y^2$?

Hint. For generating one of the parameterizations you may want to consider $\sec^2(t) - \tan^2(t) = 1$.

As you saw in Activity 10.2.4, it is convenient to parameterize conic sections and other curves defined by common relationships between coordinates. In Exercise 24, we will generalize our work for parameterizing circles earlier to give a general parameterization of a translated ellipse. These ideas are extended in Exercise 29, which asks you to parameterize hyperbolas.

Summary

- A vector-valued function is a function whose input is a real parameter t and whose output is a vector that depends on t. The graph of a vector-valued function is the set of all terminal points of the output vectors with their initial points at the origin.
- Every vector-valued function provides a parameterization of a curve. In $\mathbb{R}^2$, a parameterization of a curve is a pair of equations $x = x(t)$ and $y = y(t)$ that describes the coordinates of a point (x, y) on the curve in terms of a parameter t. In $\mathbb{R}^3$, a parameterization of a curve is a set of three equations $x = x(t)$, $y = y(t)$, and $z = z(t)$ that describes the coordinates of a point (x, y, z) on the curve in terms of a parameter t.

10.2.4 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Please keep this in mind when working WeBWorK exercises.

1. Find the domain of the vector function

$$\mathbf{r}(t) = \left\langle \ln(19t), \sqrt{t+11}, \frac{1}{\sqrt{18-t}} \right\rangle$$

using .

Domain: _____

2. Suppose parametric equations for the line segment between $(-2, -7)$ and $(0, -10)$ have the form:

$$x = a + bt$$
$$y = c + dt$$

If the parametric curve starts at $(-2, -7)$ when $t = 0$ and ends at $(0, -10)$ at $t = 1$, then find a, b, c, and d.

$a =$ _____,
$b =$ _____,
$c =$ _____,
$d =$ _____.

3. Find a parametrization of the circle of radius 3 in the xy-plane, centered at the origin, oriented clockwise. The point $(3, 0)$ should correspond to $t = 0$. Use t as the parameter for all of your answers.

$x(t) =$ _____
$y(t) =$ _____

4. Find a vector parametrization of the circle of radius 4 in the xy-plane, centered at the origin, oriented clockwise so that the point $(4, 0)$ corresponds to $t = 0$ and the point $(0, -4)$ corresponds to $t = 1$.

$\vec{r}(t) =$ _____

5. Find a parametrization of the curve $x = -4z^2$ in the xz-plane. Use t as the parameter for all of your answers.

$x(t) =$ _____
$y(t) =$ _____
$z(t) =$ _____

6. Find parametric equations for the quarter-ellipse from $(3, 0, 9)$ to $(0, -2, 9)$ centered at $(0, 0, 9)$ in the plane $z = 9$. Use the interval $0 \le t \le \pi/2$.

$x(t) =$ _____
$y(t) =$ _____
$z(t) =$ _____

7. Are the following statements true or false?

(a) The parametric curve $x = (3t+4)^2, y = 5(3t+4)^2 - 9$, for $0 \le t \le 3$ is a line segment.

(b) The line parametrized by $x = 7, y = 5t, z = 6 + t$ is parallel to the x-axis.

(c) A parametrization of the graph of $y = \ln(x)$ for $x > 0$ is given by $x = e^t, y = t$ for $-\infty < t < \infty$.

8. Find a vector function that represents the curve of intersection of the paraboloid $z = 7x^2 + 2y^2$ and the cylinder $y = 5x^2$. Use the variable t for the parameter.

$\mathbf{r}(t) = \langle t,$_____, _____$\rangle$

9. A bicycle wheel has radius R. Let P be a point on the spoke of a wheel at a distance d from the center of the wheel. The wheel begins to roll to the right along the the x-axis. The curve traced out by P is given by the following parametric equations:

$x = 21\theta - 13\sin(\theta)$
$y = 21 - 13\cos(\theta)$
What must we have for R and d?
R= _____
d = _____

10. Find a vector function that represents the curve of intersection of the paraboloid $z = 9x^2 + 3y^2$ and the cylinder $y = 4x^2$. Use the variable t for the parameter.

$\mathbf{r}(t) = \langle t,$_____, _____$\rangle$

11. (a) Find a vector-parametric equation $\vec{r}_1(t) = \langle x(t), y(t), z(t)\rangle$ for the shadow of the circular cylinder $x^2 + z^2 = 5$ in the xz-plane.

Shadow: $\vec{r}_1(t) =$ _____
for $0 \le t \le 2\pi$.

(b) Find a vector-parametric equation for intersection of the circular cylinder $x^2 + z^2 = 5$ and the plane $2x + 9y + 6z = 1$.

Intersection: $\vec{r}_2(t) =$ _____
for $0 \le t \le 2\pi$.

12. Find a vector-parametric equation for the intersection of the paraboloids $y = 3x^2 + 7z^2$ and $y = 10 - 7x^2 - 3z^2$.

Intersection: $\vec{r}(t) =$ _____
for $0 \le t \le 2\pi$.

13. Find a parametrization of the curve $x = -4z^2$ in the xz-plane. Use t as the parameter for all of your answers.

$x(t) =$ _____
$y(t) =$ _____
$z(t) =$ _____

14. Find a vector parametric equation for the parabola $y = x^2$ from the origin to the point $(3, 9)$ using t as a parameter.

$\vec{r}(t) =$ _____ for _____ $\le t \le$ _____

15. Find a parameterization for the circle of radius 5 in the xy-plane, centered at the origin, clockwise.

$x(t) =$ _____,
$y(t) =$ _____,
$z(t) =$ _____

16. Find a parameterization $(x(t), y(t), z(t))$ for the curve $y = 6 - 2x^3$ that is parallel to the xy-plane. Your parametrized curve should pass through the point $(0, 6, -4)$ when $t = -1$, and have $x'(-1) \neq 0$.

$x(t) =$ _____,
$y(t) =$ _____,
$z(t) =$ _____

17. The ellipse

$$\frac{x^2}{4^2} + \frac{y^2}{8^2} = 1$$

can be drawn with parametric equations. If

$$x = r\cos(t)$$

then r = _____
and y = _____

18. Which is a parametric equation for the curve $y = 8 - 6x$?

◯ $c(t) = (t, 8 + t)$

◯ $c(t) = (t, 6 + t)$

◯ $c(t) = (8t, 6t)$

◯ $c(t) = (t, 8 - 6t)$

19. Which is a parametric equation for the curve $16 = (x - 9)^2 + (y + 5)^2$?

◯ $c(t) = (9 + 4\cos(t), 4\sin(t) - 5)$

◯ $c(t) = (4\cos(t) - 5, 9 + 4\sin(t))$

◯ $c(t) = (9 + 16\cos(t), 16\sin(t) - 5)$

◯ $c(t) = (16\cos(t) - 5, 9 + 16\sin(t))$

20. The function $\mathbf{r}(t)$ traces a circle. Determine the radius, center, and plane containing the circle

$\mathbf{r}(t) = -9\mathbf{i} + (13\cos(t))\,\mathbf{j} + (13\sin(t))\,\mathbf{k}$
Plane : $x =$ _____
Circle's Center : (_____ , _____ , _____)
Radius : _____

21. Find a parametrization, using $\cos(t)$ and $\sin(t)$, of the following curve:
The intersection of the plane $y = 5$ with the sphere $x^2 + y^2 + z^2 = 50$
$\mathbf{r}(t) = \langle$ _____, _____, _____ $\rangle$

22. A standard parameterization for the unit circle is $\langle\cos(t), \sin(t)\rangle$, for $0 \leq t \leq 2\pi$.

(a) Find a vector-valued function $\vec{r}$ that describes a point traveling along

the unit circle so that at time $t = 0$ the point is at $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and travels clockwise along the circle as t increases.

(b) Find a vector-valued function $\vec{r}$ that describes a point traveling along the unit circle so that at time $t = 0$ the point is at $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and travels counter-clockwise along the circle as t increases.

(c) Find a vector-valued function $\vec{r}$ that describes a point traveling along the unit circle so that at time $t = 0$ the point is at $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ and travels clockwise along the circle as t increases.

(d) Find a vector-valued function $\vec{r}$ that describes a point traveling along the unit circle so that at time $t = 0$ the point is at $(0, 1)$ and makes one complete revolution around the circle in the counter-clockwise direction on the interval $[0, \pi]$.

23. You might recall the pythagorean identity from trigonometry given by

$$\cos^2(\theta) + \sin^2(\theta) = 1$$

We want to use this identity to parameterize the circle with center (h, k) and radius $r > 0$ given by

$$(x - h)^2 + (y - k)^2 = r^2$$

(a) Reorganize the general equation of a circle

$$(x - h)^2 + (y - k)^2 = r^2$$

to be of the form

$$(_____)^2 + (_____)^2 = 1$$

(b) Because identities work for any value, we can use the pythagorean identity to parameterize the general equation of a circle. Fill in the blanks below with each of the respective quantities from the previous task

$$\cos(t) = _____$$
$$\sin(t) = _____$$

(c) Solve each of the two equations from the previous part for x and y to get parametric equations for a generalized circle.

(d) Write a couple sentences about what r, h, and k mean in the parametric equations for a generalized circle.

24. In Exercise 10.2.4.23, we saw that to parametrize the implicit equation of a circle with center (h, k) and radius $r > 0$ given by

$$(x - h)^2 + (y - k)^2 = r^2$$

we can use the relationship between sin and cos. We found a parametric equations for the generalized circle can be expressed by

$$x(t) = r\cos(\theta) + h \qquad y(t) = r\sin(\theta) + k$$

The conic section most closely related to the circle is the ellipse. The

general equation of an ellipse (centered at the origin) is given by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

We can continue to make use of the relationship between sin and cos to discover parametric equations for an ellipse. In fact, without the a and b in the equation things would work perfectly. To remedy this (that is, get rid of a, b) we simply multiply by sin and cos by these values. Consider the equations

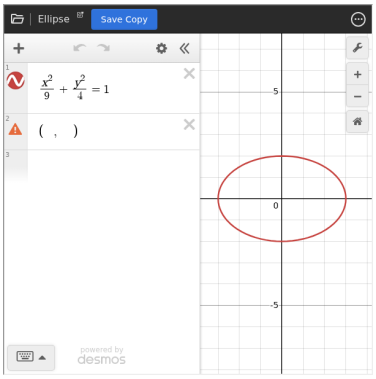
$$x(t) = a\cos(t) \qquad y(t) = b\sin(t).$$

(a) Substitute the equations for $x(t), y(t)$ above into

$$\frac{x^2}{a^2} + \frac{y^2}{b^2}$$

and simplify.

(b) Consider the ellipse given by $\frac{x^2}{9} + \frac{y^2}{4} = 1$. What are the parametric equations for this ellipse? Graph them below to ensure you obtain the exact same graph.



Standalone
In Context

Figure 10.2.11 A Desmos Interactive for Exercise 10.2.4.24

(c) In Figure 10.2.11, update the equation in x and y as well as the parametric equations for the ellipse $\frac{x^2}{16} + \frac{y^2}{25} = 1$.

(d) Write a few sentences about the orientation of the parametric equations given as your answer for the previous couple of tasks. In particular, you should talk about the following: at what point do you "start", i.e. what point corresponds to $t = 0$? How do you move around the ellipse (clockwise or counterclockwise)?

(e) Update your equations in x and y from part (c), as well as the parametric equations that were your answer to part (c), so that the ellipse has a center of $(-2, 1)$.

(f) Since we know that

$$(\cos(t))^2 + (\sin(t))^2 = 1 \quad \text{for every } t \in \mathbb{R}$$

and the form of a translated ellipse is given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1,$$

set

$$\frac{(x-h)^2}{a^2} = (\cos(t))^2$$

and solve for x.

Set up a similar equation involving y and $\sin(t)$ then solve for y to get a general set of parametric equations for the translated ellipse.

25. Let a and b be positive real numbers. You have probably seen the equation $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ that generates an ellipse, centered at (h, k), with a horizontal axis of length $2a$ and a vertical axis of length $2b$.

(a) Write a few sentences to explain why the vector function $\vec{r}$ defined by $\vec{r}(t) = \langle a\cos(t), b\sin(t)\rangle$, $0 \le t \le 2\pi$ is a parameterization of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

(b) Find a parameterization of the ellipse $\frac{x^2}{4} + \frac{y^2}{16} = 1$ that is traversed clockwise.

(c) Find a parameterization of the ellipse $\frac{(x+3)^2}{4} + \frac{(y-2)^2}{9} = 1$.

(d) Determine the xy-equation of the ellipse that is parameterized by

$$\vec{r}(t) = \langle 3 + 4\sin(2t), 1 + 3\cos(2t)\rangle.$$

26. In this exercise, we will examine the graph of $z = 3x^2 + 4y^2 - 2$. You do not need to know what this graph looks like, but we will look at the intersection of this graph with several types of fundamental planes.

(a) Determine a vector-valued function $\vec{r}$ that parameterizes the curve which is the $x = 2$ trace of $z = 3x^2 + 4y^2 - 2$. Plot the resulting curve. Do likewise for $x = -2, -1, 0$, and 1.

(b) Determine a vector-valued function $\vec{r}$ that parameterizes the curve which is the $y = 2$ trace of $z = 3x^2 + 4y^2 - 2$. Plot the resulting curve. Do likewise for $y = -2, -1, 0$, and 1.

(c) Determine a vector-valued function $\vec{r}$ that parameterizes the curve which is the $z = 2$ contour of $z = 3x^2 + 4y^2 - 2$. Plot the resulting curve. Do likewise for $z = -2, -1, 0$, and 1.

(d) Use the traces and contours you've just investigated to create a wireframe plot of the surface generated by $z = 3x^2 + 4y^2 - 2$. In addition, write a couple sentences to describe the characteristics of the surface.

27. Recall that any line in space may be represented parametrically by a vector-valued function.

(a) Find a vector-valued function $\vec{r}$ that parameterizes the line through $(-2, 1, 4)$ in the direction of the vector $\vec{v} = \langle 3, 2, -5\rangle$.

(b) Find a vector-valued function $\vec{r}$ that parameterizes the line of intersection of the planes $x + 2y - z = 4$ and $3x + y - 2z = 1$.

(c) Determine the point of intersection of the lines given by

$$x = 2 + 3t,\ y = 1 - 2t,\ z = 4t,$$

$$x = 3 + 1s,\ y = 3 - 2s,\ z = 2s.$$

Then, find a vector-valued function $\vec{r}$ that parameterizes the line that passes through the point of intersection you just found and is perpendicular to both of the given lines.

28. For each of the following, describe the effect of the parameter s on the parametric curve for t in the interval $[0, 2\pi]$.

(a) $\vec{r}(t) = \langle \cos(t), \sin(t) + s \rangle$

(b) $\vec{r}(t) = \langle \cos(t) - s, \sin(t) \rangle$

(c) $\vec{r}(t) = \langle s\cos(t), \sin(t) \rangle$

(d) $\vec{r}(t) = \langle s\cos(t), s\sin(t) \rangle$

(e) $\vec{r}(t) = \langle \cos(st), \sin(st) \rangle$

29. Parameterizing a Translated Hyperbola. Another conic section that will be useful to quickly parameterize is the hyperbola. The translated hyperbola is given by

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \tag{10.2.1}$$

for horizontal orientation and

$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1 \tag{10.2.2}$$

for vertical orientation.

(a) Find a trig identity (using the variable t) of the form

$$(_____)^2 - (_____)^2 = 1$$

Hint. Look for one involving secant and tangent.

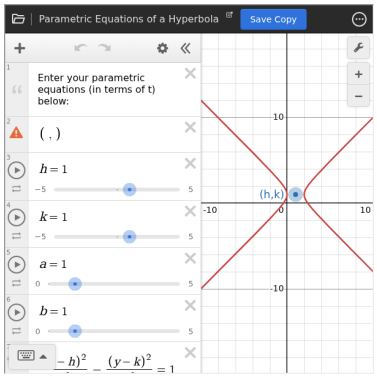
(b) Set the first blank element of your trig identity ($(_____)^2$) from the previous part equal to

$$\frac{(x-h)^2}{a^2}$$

and then solve for x.

(c) Set up a similar equation involving y and the trig function from the second blank of part 10.2.4.29.a then solve for y to get a general set of parametric equations for the translated hyperbola.

(d) Substitute in your parametric equations for the translated hyperbola (your answers for parts (b) and (c) above) into the Desmos Interactive below. You should check that your equations trace the same graph as the translated hyperbola. You should move the sliders for h, k, a, and b to convince yourself that this set of parametric equations works for all cases (of the horizontal orientation).



Standalone
In Context

Figure 10.2.12 A Desmos Interactive for Exercise 10.2.4.29

30. For each of the following equations, give a parameterization and be sure state the bounds of your parameterization.

(a) $\frac{(x-2)^2}{9} + \frac{(y+1)^2}{16} = 1$

(b) $\frac{(y+3)^2}{4} - \frac{(x+1)^2}{9} = 1$

(c) $(y-7)^2 = -4(x+1)$

10.3 Calculus of Vector-Valued Functions of One Variable

Motivating Questions

- What do we mean by the derivative of a vector-valued function and how do we efficiently calculate it?
- What does the derivative of a vector-valued function measure?
- What do we mean by the integral of a vector-valued function and how do we efficiently calculate it?
- How do we describe the motion of a projectile if the only force acting on the object is acceleration due to gravity?

10.3.1 Introduction

In this book, we will use the structure of the Classic Calculus Approach to investigate properties of new kinds of functions, equations, and graphs. In this section, we will investigate the meaning of limits, derivatives, and integrals of vector-valued functions of one variable by generalizing the intuition and concepts from single variable calculus. Additionally, we will connect these measurements to properties of curves in space as the graphs of these vector-valued functions.

Preview Activity 10.3.1 As the only employee of *Steer Clear*, you have decided that you need to understand how the timing of position measurements will change different properties related to your self-driving car. You decide to drive in a figure eight path described by $\vec{r}_8(t) = \langle \cos(t), \sin(2t) \rangle$ for $0 \leq t \leq 2\pi$. A plot of this path is given in Figure 10.3.1.

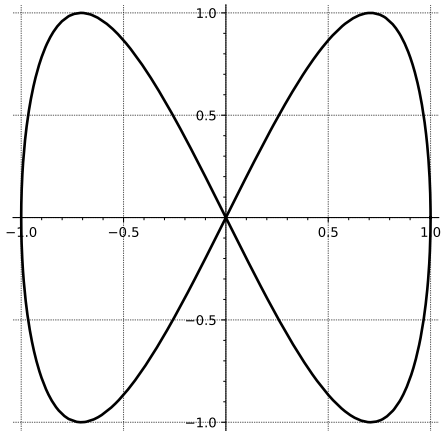


Figure 10.3.1 A plot of the figure eight curve given by $\vec{r}_8(t) = \langle \cos(t), \sin(2t) \rangle$ for $0 \leq t \leq 2\pi$.

(a) In order to understand how often your software should collect location data, you decide to look at your position for a few different times. Calculate the following, rounding the component values to three decimal places. Draw the output vectors of $\vec{r}_8$ on Figure 10.3.1 in standard position.

i. $\vec{r}_8(3)$ ii. $\vec{r}_8(3.1)$ iii. $\vec{r}_8(3.14)$

The tips of these vectors correspond to the locations that would be sampled if you wanted to know the location of your car at $\vec{r}_8(\pi)$ but collected data every second, every tenth of a second, and every hundredth of a second, respectively.

(b) Write a couple of sentences to describe both geometrically and algebraically what happens to the output of $\vec{r}_8(t)$ as $t \to \pi$.

(c) Calculate $\vec{r}_8(\pi)$ and $\vec{r}_8(\pi) - \vec{r}_8(3)$. Sketch $\vec{r}_8(\pi)$ and $\vec{r}_8(3)$ in standard position in Figure 10.3.1. Also plot $\vec{r}_8(\pi) - \vec{r}_8(3)$, positioned in such a way to illustrate that it is difference of these two position vectors.

Hint. Remember that $\vec{r}_8(\pi)$, $\vec{r}_8(3)$, and $\vec{r}_8(\pi) - \vec{r}_8(3)$ will form a triangle as in Figure 9.3.7.

(d) Compute $\frac{\vec{r}_8(\pi) - \vec{r}_8(3)}{\pi - 3}$ and explain how this calculation is different than the result of the previous step.

(e) How would you expect $\frac{\vec{r}_8(\pi) - \vec{r}_8(3.1)}{\pi - 3.1}$ to be different than $\frac{\vec{r}_8(\pi) - \vec{r}_8(3)}{\pi - 3}$? Use this idea write about what is measured by $\frac{\vec{r}_8(\pi) - \vec{r}_8(\pi - h)}{h}$ if we look at smaller and smaller values of h. Remember to be specific about what aspects of our curve or drive are being measured.

In part (b) of the Preview Activity, we examined what the limit of a vector-valued function will look like. Specifically, the limit of a vector-valued function of one variable has the same meaning as in your first calculus course. The limit of a function is a measure of what the *output* approaches as the inputs get closer to a particular value. We will not state the formal version of limits of vector-valued functions of one variable, but rather will appeal to the important idea that limits can be defined and calculated componentwise. If $\vec{r}(t) = \langle x_1(t), \ldots, x_n(t) \rangle$, then the limit of $\vec{r}$ as $t \to a$ exists if all of the limits of $x_i(t)$ exist as $t \to a$. In other words, the limit of a vector-valued function of

one variable is the same as the limits of the n component functions organized as a vector. In $\mathbb{R}^3$, we can express this algebraically as

$$\lim_{t\to a}\langle f(t), g(t), h(t)\rangle = \left\langle \lim_{t\to a} f(t), \lim_{t\to a} g(t), \lim_{t\to a} h(t)\right\rangle.$$

10.3.2 The Derivative

In single variable calculus, we define the derivative, f', of a given function f by

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

provided the limit exists. At a given value of a, $f'(a)$ measures the instantaneous rate of change of f, and also tells us the slope of the tangent line to the curve $y = f(x)$ at the point $(a, f(a))$. The definition of the derivative extends naturally to vector-valued functions and curves in space.

Definition 10.3.2 The **derivative** of a vector-valued function $\vec{r}$ is defined to be

$$\vec{r}\,'(t) = \lim_{h\to 0}\frac{\vec{r}(t+h)-\vec{r}(t)}{h}$$

for those values of t at which the limit exists. We also use the notation $\frac{d\vec{r}}{dt}$ and $\frac{d}{dt}[\vec{r}(t)]$ for $\vec{r}\,'(t)$. ◆

We will take a moment here to note that both the single variable calculus definition of derivative and Definition 10.3.2 have the form of the classic calculus approach described above. In particular, both definitions are of the form

$$\text{derivative} = \lim_{\text{stepsize}\to 0}\frac{\text{change in output of function}}{\text{change in input of function}}$$

Let $\vec{r}$ be the vector-valued function whose graph is shown in Figure 10.3.3, and let h be a scalar that represents a small change in time. The vector $\vec{r}(t)$ is the blue vector in Figure 10.3.3 and $\vec{r}(t+h)$ is the green vector.

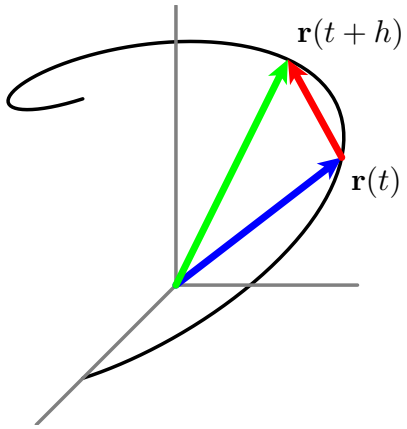


Figure 10.3.3 A single difference quotient of a vector-valued function

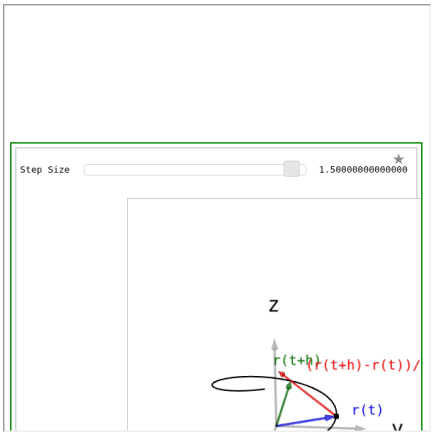
If $\vec{r}(t)$ determines the **position** of an object at time t, then $\frac{\vec{r}(t+h)-\vec{r}(t)}{h}$ represents the average rate of change in the position of the object over the interval $[t, t+h]$. We also call this the **average velocity** of the object on the interval$[t, t+h]$. It may feel a bit surprising to think of a rate of change as a vector quantity. However, describing how the vector quantity position

changes with respect to time requires knowing both about the direction in which position changes and the magnitude of that change. Furthermore, when computing the difference quotient, the vector subtraction in the numerator is done componentwise. The factor of $1/h$ is also multiplied into each component, so we can view the difference quotient's components as each being the average rate of change for that particular component.

The derivative

$$\vec{r}\,'(t) = \lim_{h \to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}$$

is the instantaneous rate of change of $\vec{r}(t)$ at time t for those values of t for which the limit exists. Therefore, the **instantaneous velocity** of the object at time t is $\vec{v}(t) = \vec{r}\,'(t)$. Figure 10.3.4 allows you to reduce the step size in the difference quotient to approach 0. In doing so, you should notice that $\vec{r}\,'(t)$ is tangent to the graph of $\vec{r}(t)$ at time t. We will return to this important characteristic later in the section.



Standalone
In Context
Embed

Figure 10.3.4 A plot of a curve in space with a difference in position vectors shown

Similarly,

$$\vec{v}\,'(t) = \vec{r}\,''(t) = \lim_{h \to 0} \frac{\vec{v}(t+h) - \vec{v}(t)}{h}$$

is the instantaneous rate of change of the velocity of the object at time t, for those values of t for which the limit exists. Therefore, the **acceleration** of the moving object is given by $\vec{a}(t) = \vec{v}\,'(t)$.

Both the velocity and acceleration are *vector quantities*: they have magnitude and direction. In single-variable calculus, we distinguish the velocity of an object, which can be positive, negative, or zero, from its speed, which must be nonnegative. The speed of an object in the single-variable case is the absolute value of its velocity. The analogous idea for an object having movement described by a vector-valued function is that the *magnitude* of the object's velocity vector is the **speed** of the object. This allows the speed $\|\vec{v}(t)\|$ to again be a nonnegative scalar quantity.

10.3.3 Computing Derivatives

As we learned in single variable calculus, computing derivatives using the definition is often difficult because the limit involved is an indeterminate form. Fortunately, properties of limits and vectors make it straightforward to calculate the derivative of a vector-valued function. This will be similar to how we developed shortcut differentiation rules in single-variable calculus. If we analyze the steps in Definition 10.3.2, we identify three steps

1. vector subtraction ($\vec{r}(t+h) - \vec{r}(t)$),
2. scalar multiplication (multiply by $\frac{1}{h}$),
3. evaluate limit as $h \to 0$, and

each of these is all evaluated *componentwise* on a vector-valued function. Therefore, familiar differentiation rules from single-variable calculus can be applied to each component of $\vec{r}(t)$ to compute the derivative $\vec{r}\,'$.

Key Idea 10.3.5 Efficient calculation of the derivative of a vector-valued function. *If $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$, then*

$$\frac{d}{dt}\vec{r}(t) = x'(t)\hat{\imath} + y'(t)\hat{\jmath} + z'(t)\hat{k}$$

for those values of t at which x, y, and z are differentiable.

Proof. To see why, recall that the limit of a sum is the sum of the limits, and that we can remove constant factors from limits. Thus, as we observed after Preview Activity 10.3.1, if $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$, it follows that

$$\begin{aligned}
\vec{r}\,'(t) &= \lim_{h\to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h} \\
&= \lim_{h\to 0} \frac{[x(t+h) - x(t)]\hat{\imath} + [y(t+h) - y(t)]\hat{\jmath} + [z(t+h) - z(t)]\hat{k}}{h} \\
&= \left(\lim_{h\to 0} \frac{x(t+h) - x(t)}{h}\right)\hat{\imath} + \left(\lim_{h\to 0} \frac{y(t+h) - y(t)}{h}\right)\hat{\jmath} \\
&\quad + \left(\lim_{h\to 0} \frac{z(t+h) - z(t)}{h}\right)\hat{k} \\
&= x'(t)\hat{\imath} + y'(t)\hat{\jmath} + z'(t)\hat{k}.
\end{aligned}$$

Thus, we can calculate the derivative of a vector-valued function by differentiating each of its components. ■

Our next activity gives you a chance to practice with componentwise differentiation and also asks you to think about when limits of vector-valued functions do not exist.

Activity 10.3.2 For each of the following vector-valued functions, state any values of a for which $\lim_{t\to a} \vec{r}(t)$ will *not* exist and find $\vec{r}\,'(t)$.

(a) $\vec{r}(t) = \langle \cos(t), t\sin(t), \ln(t)\rangle$.

(b) $\vec{r}(t) = \langle t^2 + 3t, e^{-2t}, \dfrac{t}{t^2+1}\rangle$.

(c) $\vec{r}(t) = \langle \tan(t), \cos(t^2), te^{-t}\rangle$.

(d) $\vec{r}(t) = \left\langle \sqrt{t^4+4}, \frac{2}{t^2+t}, e^{2t}\sin(-2t)\right\rangle$.

In first-semester calculus, we developed several important differentiation rules, including the constant multiple, product, quotient, and chain rules. For instance, we can formally state the product rule as

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

There are several analogous rules for vector-valued functions, including a product rule for each of the products we have defined involving vectors. These

rules, which are easily verified by componentwise application of the familiar single-variable calculus rules, are summarized below.

Properties of Derivatives of Vector-Valued Functions.

Let f be a differentiable real-valued function of a real variable t. Let $\vec{r}$ and $\vec{s}$ be differentiable vector-valued functions of the real variable t. Then

Sum Rule	$\frac{d}{dt}\left[\vec{r}(t)+\vec{s}(t)\right]=\vec{r}\,'(t)+\vec{s}\,'(t)$
Scalar Multiplication Product Rule	$\frac{d}{dt}\left[f(t)\vec{r}(t)\right]=f(t)\vec{r}\,'(t)+f'(t)\vec{r}(t)$
Dot Product Rule	$\frac{d}{dt}\left[\vec{r}(t)\cdot\vec{s}(t)\right]=\vec{r}\,'(t)\cdot\vec{s}(t)+\vec{r}(t)\cdot\vec{s}\,'(t)$
Cross Product Rule	$\frac{d}{dt}\left[\vec{r}(t)\times\vec{s}(t)\right]=\vec{r}\,'(t)\times\vec{s}(t)+\vec{r}(t)\times\vec{s}\,'(t)$
Chain Rule	$\frac{d}{dt}\left[\vec{r}\,(f(t))\right]=f'(t)\;\vec{r}\,'(f(t))$

When applying these properties, pay attention to whether the quantities involved are either scalars or vectors. For example, $\vec{r}(t)\cdot\vec{s}(t)$ defines a scalar function because we have taken the dot product of two vector-valued functions. However, $\vec{r}(t)\times\vec{s}(t)$ defines a vector-valued function since we have taken the cross product of two vector-valued functions. Also notice that while we have *three* product rules for the differentiation of vector-valued functions, we have only one chain rule and no quotient rule. The lack of a quotient rule should make sense, since we have not defined what it would mean to divide a vector by another. That we have a single chain rule arises because so far, we have only discussed how to have a single input to a vector-valued function. Thus, we cannot compose one vector-valued function with another vector-valued function. You might also notice that the order in which the factors in the Chain Rule are written is a bit different to what you would write for single-variable functions. This is because $f'(t)$ is a *scalar*, and convention says we should write scalar-vector multiplication with the scalar first.

In the next several sections of this chapter, we will examine how to use calculus tools and vector measurements to calculate different properties of curves in space and the associated velocity and acceleration vectors that come from their parameterizations. Exercise 36 offers practice on using calculus tools on vector-valued functions and previews some ideas that will be used in later sections of this chapter.

10.3.4 Tangent Lines

One of the most important ideas in first-semester calculus is that a differentiable function is **locally linear**: that is, when viewed up close, the graph of a differentiable function looks like a line. Indeed, when we zoom in sufficiently close to a particular point, the curve is indistinguishable from its tangent line. You can look at Figure 9.6.1 for an interactive visualization of this important idea.

In the same way, we expect that a smooth curve in three-dimensional space will be locally linear if the graph of the curve looks like a line when you zoom in

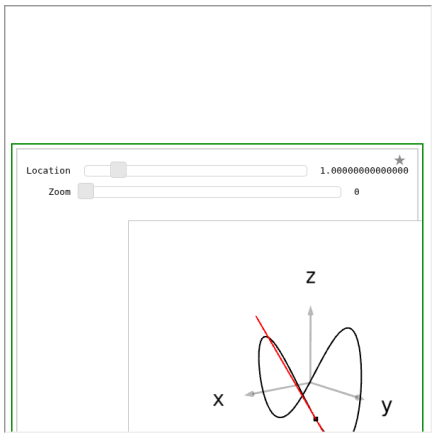
sufficiently far. In Activity 10.3.3, we investigate how to find the tangent line to such a curve. Recall from our work in Section 9.6 that the vector equation of a line that passes through the point at the tip of the vector $\vec{L}_0 = \langle x_0, y_0, z_0 \rangle$ in the direction of the vector $\vec{u} = \langle a, b, c \rangle$ can be written as

$$\vec{L}(t) = \vec{L}_0 + t\vec{u}.$$

The line described by $\vec{L}$ can also be written in parametric form as

$$x(t) = x_0 + at, \quad y(t) = y_0 + bt, \quad z(t) = z_0 + ct.$$

Figure 10.3.6 shows a plot of a curve, a point on that curve, and the line tangent to the curve at that point. You can change the point at which the tangent line is shown by moving the "Location" slider. By moving the "Zoom" slider to the right, you can show smaller and smaller regions around the highlighted point to see that the tangent line approximates the curve very well on sufficiently small neighborhoods of the point. If you move the zoom slider all the way to the left, the whole curve and axes will be shown.



Standalone
In Context
Embed

Figure 10.3.6 A plot of a curve in space with the tangent line at a point drawn in red

Activity 10.3.3

(a) Let

$$\vec{r}(t) = \cos(t)\hat{\imath} - \sin(t)\hat{\jmath} + t\hat{k}.$$

Sketch the curve using some appropriate tool and make a drawing by hand that labels the point at the terminal point of $\vec{r}(\pi)$.

(b) Recall that we discussed earlier that the vector $\vec{r}\,'(a)$ is tangent to the graph of $\vec{r}(t)$ at the point where $t = a$. Find a direction vector for the line tangent to the graph of $\vec{r}$ at the point where $t = \pi$.

(c) Find the parametric equations of the line tangent to the graph of $\vec{r}$ when $t = \pi$.

(d) On your plot of the curve $\vec{r}(t)$, sketch the tangent line corresponding to $t = \pi$ and highlight the role of $\vec{r}\,'(\pi)$ on your plot.

Our work in Activity 10.3.3 can be generalized to give a method for finding an equation of the line tangent to a curve described by a vector-valued function.

Tangent Line to Curve Described by a Vector-Valued Function.

Given a differentiable vector-valued function $\vec{r}$, the tangent line to the curve at the input value a is given by

$$\vec{L}(t) = \vec{r}(a) + t\vec{r}\,'(a). \tag{10.3.1}$$

The tangent line is determined entirely by a given point (provided by the function $\vec{r}$) and a direction vector (provided by the derivative, $\vec{r}\,'$), with both the function and its derivative evaluated at $t = a$. Note how analogous the formula for $\vec{L}(t)$ is to the tangent line approximation from single-variable calculus: in that context, for a given function $y = f(x)$ at a value $x = a$, we found that the tangent line can be expressed by the linear function $y = L(x)$ whose formula is

$$L(x) = f(a) + f'(a)(x - a).$$

Equation (10.3.1) for the tangent line $\vec{L}(t)$ to the vector-valued function $\vec{r}(t)$ is nearly identical. Indeed, because there are multiple parameterizations for a single line, it is even possible to write the parameterization as

$$\vec{L}(t) = \vec{r}(a) + (t - a)\vec{r}\,'(a). \tag{10.3.2}$$

For example, in Equation (10.3.1), $\vec{L}(0) = \vec{r}(a)$, so the line's parameterization "starts" at $t = 0$. When we write the parameterization in the form of Equation (10.3.2), $\vec{L}(a) = \vec{r}(a)$, so the line's parameterization "starts" at $t = a$. Both of these ways of parameterizing the line tangent to the graph of a vector-valued function will be useful in our future work.

10.3.5 Integrating a Vector-Valued Function

In single-variable calculus, you learned that the definite integral of a (scalar-valued) function $f(x)$ on the interval $[a, b]$ is defined as

$$\int_a^b f(x)\ dx = \lim_{\Delta x_i \to 0} \sum_{i=0}^{n} f(x_i^*)\Delta x_i,$$

where Δx_i is the width of the interval from $[x_{i-1}, x_i]$ and x_i^* is a point in $[x_{i-1}, x_i]$. The definite integral measures the net signed area between the graph of $f(x)$ and the x-axis on the interval from $x = a$ to $x = b$. The definite integral uses our classic calculus approach to measure the accumulation of the output of f over the interval $[a, b]$.

In practice, you probably did not use this Riemann sum definition to calculate any definite integrals, since the Fundamental Theorem of Calculus allowed you to use associated antiderivatives to efficiently calculate the value of a definite integral. The Fundamental Theorem of Integral Calculus states that if f is a continuous function on $[a, b]$ and $F(x)$ is an antiderivative of f (that is, $\frac{dF}{dx} = f(x)$), then

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

The family of functions that are antiderivatives of $f(x)$ is also called the **indefinite integral** of $f(x)$ and is denoted $\int f(x)\ dx$.

We can apply all of the ideas above related to definite and indefinite integrals to vector-valued functions of one variable by performing each operation separately on each component.

Integrating a Vector-Valued Function.

If $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$ with $t \in [a, b]$, then

$$\int_a^b \vec{r}(t)\ dt = \left(\int_a^b x(t)\ dt\right)\hat{\imath} + \left(\int_a^b y(t)\ dt\right)\hat{\jmath} + \left(\int_a^b z(t)\ dt\right)\hat{k},$$

provided that all three of the integrals on the right-hand side exist. An **antiderivative** of a vector-valued function $\vec{r}$ is a vector-valued function $\vec{R}$ such that

$$\vec{R}\,'(t) = \vec{r}(t).$$

The **indefinite integral** $\int \vec{r}(t)\ dt$ of a vector-valued function $\vec{r}$ is the most general antiderivative of $\vec{r}$ and represents the family of all antiderivatives of $\vec{r}$. The indefinite integral of $\vec{r}(t)$ is denoted by

$$\int \vec{r}\ dt = \int \langle x(t), y(t), z(t)\rangle\ dt = \left\langle \int x(t)\ dt, \int y(t)\ dt, \int z(t)\ dt \right\rangle.$$

The area under the curve idea does ***not*** generalize when we think about the integral of $\vec{r}(t)$. One way to understand this is that the integral of $\vec{r}$ is with respect to the input variable t and there is not a t-axis on the graph of $\vec{r}(t)$. However, as described in the introduction to this section, the concept that *will* generalize to vector-valued functions is the definite integral as a measurement of the accumulation of the output of a function. Specifically, in Recovering Position from Velocity, we stated how we can recover the position function by integrating the velocity function. This is a way of viewing position as the accumulation of the rate of change of position. If we apply our argument for the accumulation of rates of change to each component separately, then we can integrate the velocity componentwise to obtain the components of the position vector separately. Additionally, we can integrate acceleration to get the velocity as a function of t. Whether we are integrating velocity to get position or acceleration to get velocity, it will be necessary to have a particular point (in position or velocity) that will allow us to solve for the particular integration constant (the "$+C$" from single-variable calculus).

Example 10.3.7 For this example, we consider the motion of a projectile in two dimensions where the only force on the projectile is gravity. This type of problem applies to sports like archery or shotput as well as to military applications like mortar or artillery placements. It also has applications to firefighting or fireworks construction. This kind of simplified physics problem is common in introductory courses because there is only one force involved, the force is constant (does not change based on location, speed, or time), and the force is in only one coordinate direction. Even this description is a greatly simplified version of how gravity acts in a greater sense.

Newton's second law of motion says that that sum of the forces acting on an object is equal to the mass of the object times the acceleration vector. In our problem, the only force is gravity pulling the object in the negative y-direction. We want to completely determine the path traveled by an object that is launched from a fixed position at a given angle from the horizontal with a given initial velocity. This information can be stated as follows with the initial values shown in Figure 10.3.8:

- The sum of forces is given by $m\vec{a}(t) = -mg\hat{\jmath}$ (the only force is gravity pulling down with a constant force).

- The acceleration is given by $\vec{a}(t) = -g\hat{j}$.
- The initial position is $\vec{r}(0) = \langle x_0, y_0 \rangle$.
- The initial velocity is $\vec{v}(0) = \frac{d\vec{r}}{dt}(0) = \langle v_0 \cos(\theta), v_0 \sin(\theta) \rangle$, where v_0 is the initial speed.
- We want to find position as a function of time, $\vec{r}(t)$.

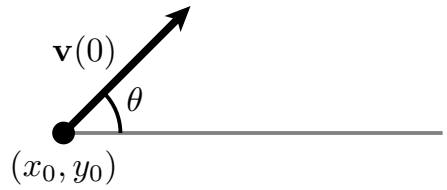


Figure 10.3.8 Projectile motion

We can simplify $m\vec{a}(t) = -mg\hat{j}$ to find $\vec{a}(t) = \langle 0, -g \rangle$.

Next, we integrate the acceleration to find the velocity as a function of t. The indefinite integral of $\vec{a}(t) = \langle 0, -g \rangle$ with respect to t will be $\langle 0, -gt \rangle + \vec{C}$, where $\vec{C}$ is a constant vector. We need to pick the constant such that the initial velocity is $\vec{v}(0) = \langle v_0 \cos(\theta), v_0 \sin(\theta) \rangle$. This will ensure that the initial speed is v_0 and that the initial velocity has an angle of θ above the horizontal. Because $\vec{v}(0) = \langle 0, -g(0) \rangle + \vec{C} = \vec{C} = \langle v_0 \cos(\theta), v_0 \sin(\theta) \rangle$, we find the following for the velocity as a function of time:

$$\vec{v}(t) = \langle 0, -gt \rangle + \langle v_0 \cos(\theta), v_0 \sin(\theta) \rangle = \langle v_0 \cos(\theta), v_0 \sin(\theta) - gt \rangle$$

Now we can integrate our velocity function to get the position as a function of t. The indefinite integral of $\langle v_0 \cos(\theta), v_0 \sin(\theta) - gt \rangle$ will be

$$\langle v_0 \cos(\theta)t, v_0 \sin(\theta)t - \frac{1}{2}gt^2 \rangle + \vec{D}.$$

Again we need to find the integration constant $\vec{D}$ that satisfies the initial position $\vec{r}(0) = \langle x_0, y_0 \rangle$. This gives the following vector-valued function for the position of the object as a function of time.

$$\vec{r}(t) = \left\langle v_0 \cos(\theta)t + x_0, -\frac{g}{2}t^2 + v_0 \sin(\theta)t + y_0 \right\rangle$$

▲

In the next activity, we look at relationships between position, velocity, and acceleration in a situation that is more complicated than the constant acceleration of Example 10.3.7. The process of using integration and differentiation will remain the same, but the results will not be nearly as simple.

Activity 10.3.4 Suppose a moving object in space has its velocity given by

$$\vec{v}(t) = (-2\sin(2t))\hat{\imath} + (2\cos(t))\hat{\jmath} + \left(1 - \frac{1}{1+t}\right)\hat{k}.$$

A graph of the position of the object for times t in $[-0.5, 3]$ is shown in Figure 10.3.9. Suppose further that the object is at the point $(1.5, -1, 0)$ at time $t = 0$.

(a) Determine $\vec{a}(t)$, the acceleration of the object at time t.

(b) Determine $\vec{r}(t)$, position of the object at time t.

(c) Compute the position, velocity, and acceleration vectors of the object at time $t = 1$ and plot these vectors using Figure 10.3.9.

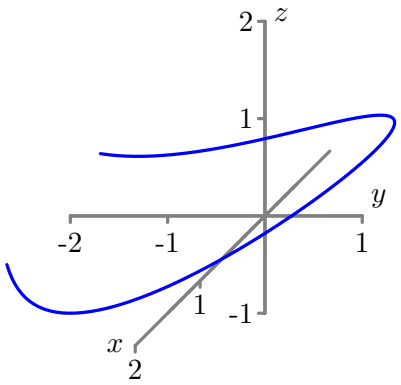


Figure 10.3.9 The position graph for the function in Activity 10.3.4

(d) Give the vector equation for the tangent line, $\vec{L}(t)$, that is tangent to the graph of $\vec{r}(t)$ at $t = 1$.

The previous activity shows a case where the motion of our object is driven by forces that either turn the object or change the speed of the object in the z-coordinate direction. We will spend the rest of this chapter looking at measurements related to motion along curves like this. Key to these measurements will be the combined use of our new calculus for vector-valued functions with the vector tools from the previous chapter.

Summary

- If $\vec{r}$ is a vector-valued function, then the derivative of $\vec{r}$ is defined by

$$\vec{r}\,'(t) = \lim_{h \to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}$$

for those values of t at which the limit exists, and is computed componentwise by the formula

$$\vec{r}\,'(t) = x'(t)\hat{\imath} + y'(t)\hat{\jmath} + z'(t)\hat{k}$$

for those values of t at which x, y, and z are differentiable, where $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$.

- The derivative $\vec{r}\,'(t)$ of the vector-valued function $\vec{r}$ tells us the instantaneous rate of change of $\vec{r}$ with respect to time, t, which can be interpreted as a direction vector for the line tangent to the graph of $\vec{r}$ at the point $\vec{r}(t)$, or also as the instantaneous velocity of an object traveling along the graph defined by $\vec{r}(t)$ at time t.

- An antiderivative of $\vec{r}$ is a vector-valued function $\vec{R}$ such that $\vec{R}\,'(t) = \vec{r}(t)$. The indefinite integral $\int \vec{r}(t)\, dt$ of a vector-valued function $\vec{r}$ is the general antiderivative of $\vec{r}$ (which is a collection of all of the antiderivatives of $\vec{r}$, with any two antiderivatives differing by at most a constant vector). Moreover, if $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$, then

$$\int \vec{r}(t)\, dt = \left(\int x(t)\, dt\right)\hat{\imath} + \left(\int y(t)\, dt\right)\hat{\jmath} + \left(\int z(t)\, dt\right)\hat{k}.$$

- If an object is launched from a point (x_0, y_0) with initial velocity v_0 at an angle θ with the horizontal, then the position of the object at time t is given by

$$\vec{r}(t) = \left\langle v_0 \cos(\theta)t + x_0, -\frac{g}{2}t^2 + v_0 \sin(\theta)t + y_0 \right\rangle$$

provided that that the only force acting on the object is the acceleration g due to gravity.

10.3.6 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Please keep this in mind when working WeBWorK exercises.

1. Find the limit:

$$\lim_{t\to 0}\langle \frac{e^{-5t}-1}{t}, \frac{t^4}{t^5-t^4}, \frac{1}{10+t}\rangle$$

$\langle$ _____, _____, _____ $\rangle$

2. Let $\mathbf{r}(t) = (\sqrt{t+2})\,\mathbf{i} + \left(\frac{t^2-9}{t-3}\right)\mathbf{j} + \sin(5\pi t)\,\mathbf{k}$.

Then
$\lim_{t\to 1} \mathbf{r}(t)$= _____ **i**+ _____**j**+ _____**k**.

3. Let $\mathbf{r}(t) = (\sqrt{t+3})\,\mathbf{i} + \left(\frac{t^2-0}{t-0}\right)\mathbf{j} + \sin(-\pi t)\,\mathbf{k}$.

Then
$\lim_{t\to 1} \mathbf{r}(t)$= _____ **i**+ _____**j**+ _____**k**.

4. If $\mathbf{r}(t) = \cos(2t)\mathbf{i} + \sin(2t)\mathbf{j} - 7t\mathbf{k}$, compute:

A. The velocity vector $\mathbf{v}(t)$ = _____ **i**+ _____ **j**+ _____ **k**

B. The acceleration vector $\mathbf{a}(t)$ = _____ **i**+ _____ **j**+ _____ **k**

Note: the coefficients in your answers must be entered in the form of expressions in the variable t; e.g. "5 cos(2t)"

5. Given that the acceleration vector is $\mathbf{a}(t) = (-4\cos(2t))\,\mathbf{i} + (-4\sin(2t))\,\mathbf{j} + (4t)\,\mathbf{k}$, the initial velocity is $\mathbf{v}(0) = \mathbf{i} + \mathbf{k}$, and the initial position vector is $\mathbf{r}(0) = \mathbf{i} + \mathbf{j} + \mathbf{k}$, compute:

A. The velocity vector $\mathbf{v}(t)$ = _____ **i**+ _____ **j**+ _____ **k**

B. The position vector $\mathbf{r}(t)$ = _____ **i**+ _____ **j**+ _____ **k**

Note: the coefficients in your answers must be entered in the form of expressions in the variable \emph{t}; e.g. "5 cos(2t)"

6. Evaluate
$\int_0^9 (t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k})dt$ = _____ **i**+ _____**j**+ _____**k**.

7. Find parametric equations for line that is tangent to the curve $x = \cos t,\ y = \sin t,\ z = t$ at the point
$(\cos(\frac{5\pi}{6}), \sin(\frac{5\pi}{6}), \frac{5\pi}{6})$.
Parametrize the line so that it passes through the given point at t=0.

All three answers are required for credit.

$x(t) =$ _____

$y(t) =$ _____

$z(t) =$ _____

8. If $\mathbf{r}(t) = \cos(-5t)\mathbf{i} + \sin(-5t)\mathbf{j} + 6t\mathbf{k}$

compute $\mathbf{r}'(t)=$ _____ **i**+ _____**j**+ _____ **k**

and $\int \mathbf{r}(t)\,dt=$ _____ **i**+ _____**j**+ _____ $\mathbf{k} + \mathbf{C}$

with **C** a constant vector.

9. For the given position vectors $\mathbf{r}(t)$,

compute the (tangent) velocity vector $\mathbf{r}'(t)$ for the given value of t .

A) Let $\mathbf{r}(t) = (\cos t, \sin t)$.

Then $\mathbf{r}'(\frac{\pi}{4})=$ (_____ , _____)?

B) Let $\mathbf{r}(t) = (t^2, t^3)$.

Then $\mathbf{r}'(3)=$ (_____ , _____)?

C) Let $\mathbf{r}(t) = e^t\mathbf{i} + e^{-3t}\mathbf{j} + t\mathbf{k}$.

Then $\mathbf{r}'(4)=$ _____ **i**+ _____**j**+ _____ **k** ?

10. Suppose $\vec{r}(t) = \cos(\pi t)\,\boldsymbol{i} + \sin(\pi t)\,\boldsymbol{j} + 3t\boldsymbol{k}$ represents the position of a particle on a helix, where z is the height of the particle.

(a) What is t when the particle has height 6?

$t =$ _____

(b) What is the velocity of the particle when its height is 6?

$\vec{v} =$ _____

(c) When the particle has height 6, it leaves the helix and moves along the tangent line at the constant velocity found in part (b). Find a vector parametric equation for the position of the particle (in terms of the original parameter t) as it moves along this tangent line.

$L(t) =$ _____

11. Suppose the displacement of a particle in motion at time t is given by the parametric equations

$x(t) = (4t-1)^2, \quad y(t) = 7, \quad z(t) = 768t^4 - 256t^3.$

(a) Find the speed of the particle when $t = 1$.

Speed = _____

(b) Find t when the particle is stationary.

$t =$ _____

12. Find the derivative of the vector function

$\mathbf{r}(t) = t\mathbf{a} \times (\mathbf{b} + t\mathbf{c})$, where

$\mathbf{a} = \langle -1, -4, -1\rangle$, $\mathbf{b} = \langle -4, -2, -3\rangle$, and $\mathbf{c} = \langle 1, -3, -3\rangle$.

$\mathbf{r}'(t) = \langle$ _____, _____, _____ $\rangle$

13. Let $\mathbf{c}_1(t) = (e^t, \sin(-5t), t^3)$, and $\mathbf{c}_2(t) = (e^{-4t}, \cos(-4t), 3t^3)$

$\dfrac{d}{dt}\left[\mathbf{c}_1(t) \cdot \mathbf{c}_2(t)\right] =$_____

$\dfrac{d}{dt}\left[\mathbf{c}_1(t) \times \mathbf{c}_2(t)\right] =$ _____**i** +

_____**j** +

_____**k**

14. A gun has a muzzle speed of 90 meters per second. What angle of elevation should be used to hit an object 190 meters away? Neglect air resistance and use $g = 9.8\,\text{m/sec}^2$ as the acceleration of gravity.

Answer: _____ radians

15. A child wanders slowly down a circular staircase from the top of a tower. With x, y, z in feet and the origin at the base of the tower, her position t minutes from the start is given by

$$x = 30\cos t, \quad y = 30\sin t, \quad z = 90 - 5t.$$

(a) How tall is the tower?
height = _____ ft
(b) When does the child reach the bottom?
time = _____ minutes
(c) What is her speed at time t?
speed = _____ ft/min
(d) What is her acceleration at time t?
acceleration = _____ ft/min^2

16. Find the derivative of the vector function

$$\mathbf{r}(t) = \ln(19 - t^2)\,\mathbf{i} + \sqrt{2+t}\,\mathbf{j} + 6e^{-3t}\,\mathbf{k}$$

$\mathbf{r}'(t) = \langle$ _____, _____, _____ $\rangle$

17. If $\mathbf{r}(t) = \cos(-5t)\mathbf{i} + \sin(-5t)\mathbf{j} + 4t\mathbf{k}$
compute $\mathbf{r}'(t)$= _____ **i**+ _____**j**+ _____ **k**
and $\int \mathbf{r}(t)\,dt$= _____ **i**+ _____**j**+ _____ **k** + **C**
with **C** a constant vector.

18. For the given position vectors $\mathbf{r}(t)$,
compute the (tangent) velocity vector $\mathbf{r}'(t)$ for the given value of t .
A) Let $\mathbf{r}(t) = (\cos t, \sin t)$.
Then $\mathbf{r}'(\frac{\pi}{4})$= (_____ , _____)?
B) Let $\mathbf{r}(t) = (t^2, t^3)$.
Then $\mathbf{r}'(2)$= (_____ , _____)?
C) Let $\mathbf{r}(t) = e^t\mathbf{i} + e^{-2t}\mathbf{j} + t\mathbf{k}$.
Then $\mathbf{r}'(2)$= _____ **i**+ _____**j**+ _____ **k** ?

19. Find the velocity and acceleration vectors for the motion described by

$$x = 3 + 5t^2, \quad y = 1 - 5t^2, \quad z = 7 + 4t^2.$$

$\vec{v}(t)$ = _____
$\vec{a}(t)$ = _____

20. For the given position vectors $\mathbf{r}(t)$,
compute the (tangent) velocity vector $\mathbf{r}'(t)$ for the given value of t .
A) Let $\mathbf{r}(t) = (\cos 2t, \sin 2t)$.
Then $\mathbf{r}'(\frac{\pi}{4})$= (_____ , _____)?
B) Let $\mathbf{r}(t) = (t^2, t^3)$.
Then $\mathbf{r}'(4)$= (_____ , _____)?
C) Let $\mathbf{r}(t) = e^{2t}\mathbf{i} + e^{-4t}\mathbf{j} + t\mathbf{k}$.
Then $\mathbf{r}'(0)$= _____ **i**+ _____**j**+ _____ **k** ?

21. The position function of a particle is given by $\mathbf{r}(t) = \langle -2t^2, 4t, t^2 - 5t\rangle$.
At what time is the speed minimum? _____

22. Find the velocity $\vec{v}(t)$ and the speed $\|\vec{v}(t)\|$ of a particle with position

$$x = 5t^2 - 30t + 45, \quad y = 6, \quad z = 4t^3 - 18t^2.$$

$\vec{v}(t) =$_____

speed $\|\vec{v}(t)\| =$ _____

Find any times at which the particle stops.

$t =$ _____

(Give your answer as a time or list of times, comma separated.)

23. Find the velocity, acceleration, and speed of a particle with position function

$\mathbf{r}(t) = \langle 4\cos t, 2t, 4\sin t\rangle$

$\mathbf{v}(t) = \langle$_____,_____, _____ $\rangle$

$\mathbf{a}(t) = \langle$_____,_____, _____ $\rangle$

$|\mathbf{v}(t)| =$_____

24. Find the velocity, acceleration, and speed of a particle with position function

$\mathbf{r}(t) = \langle 6t\sin t, 6t\cos t, -5t^2\rangle$

$\mathbf{v}(t) = \langle$_____,_____, _____ $\rangle$

$\mathbf{a}(t) = \langle$_____,_____, _____ $\rangle$

$|\mathbf{v}(t)| =$_____

25. If $\mathbf{r}(t) = \cos(9t)\mathbf{i} + \sin(9t)\mathbf{j} + 8t\mathbf{k}$, compute:

A. The velocity vector $\mathbf{v}(t) =$ _____ **i**+ _____ **j**+ _____ **k**

B. The acceleration vector $\mathbf{a}(t) =$ _____ **i**+ _____ **j**+ _____ **k**

Note: the coefficients in your answers must be entered in the form of expressions in the variable t; e.g. "5 cos(2t)"

26. Suppose $x(t) = 3t^2$ and $y(t) = t^3 - 2t$.

(a) Find $x'(t)$ and $y'(t)$.

$x'(t) =$ _____

$y'(t) =$ _____

(b) Use $x'(t)$ and $y'(t)$ to determine when $x(t)$ and $y(t)$ are increasing and decreasing. Then, find the intervals of t for which the motion of the curve is up, down, right, and left in the xy-plane. (Your answers below should be intervals of t values.)

Up and to the right: _____

Up and to the left: _____

Down and to the right: _____

Down and to the left: _____

27. Suppose a curve is traced by the parametric equations

$$x = 2\big(\sin(t) + \cos(t)\big)$$

$$y = 16 - 6\cos^2(t) - 12\sin(t)$$

as t runs from 0 to π. At what point (x, y) on this curve is the tangent line horizontal?

$x =$ _____

$y =$ _____

28. Find the parametric equations for the tangent line to the curve

$$x = t^2 - 1, y = t^5 + 1, z = t^5$$

at the point (0, 2, 1). Use the variable t for your parameter.

$x =$_____,
$y =$_____,
$z =$_____

29. Find the parametric equations for the tangent line to the curve

$$x = t^3 - 1, y = t^3 + 1, z = t^2$$

at the point (26, 28, 9). Use the variable t for your parameter.

$x =$_____,
$y =$_____,
$z =$_____

30. Find the equation for the line tangent to the parametric curve:

$$x = t^3 - 9t$$
$$y = 9t^2 - t^4$$

at the points where $t = 3$ and $t = -3$.

For $t = 3$, the tangent line (in form $y = mx + b$) is

$y =$ _____.

For $t = -3$, the tangent line is

$y =$ _____.

31. Find the velocity and position vectors of a particle with acceleration $\mathbf{a}(t) = \langle 0, 0, 6\rangle$, and initial conditions $\mathbf{v}(0) = \langle 3, -5, -4\rangle$ and $\mathbf{r}(0) = \langle 4, 3, 5\rangle$.

$\mathbf{v}(t) = \langle$_____, _____,_____$\rangle$
$\mathbf{r}(t) = \langle$_____, _____,_____$\rangle$

32. Evaluate

$\int_{-2}^{3}(t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k})dt =$ _____ .

33. Evaluate

$\int_0^8(t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k})dt =$ _____ **i**+ _____**j**+ _____**k**.

34. If $\mathbf{r}(t) = \cos(5t)\mathbf{i} + \sin(5t)\mathbf{j} + 6t\mathbf{k}$

compute $\mathbf{r}'(t)=$ _____ **i**+ _____**j**+ _____ **k**

and $\int \mathbf{r}(t)\,dt=$ _____ **i**+ _____**j**+ _____ $\mathbf{k} + \mathbf{C}$

with **C** a constant vector.

35. Compute the derivative of each of the following functions in two different ways: (1) use the rules provided in the theorem stated just after Activity 10.3.2, and (2) rewrite each given function so that it is stated as a single function (either a scalar function or a vector-valued function with three components), and differentiate component-wise. Compare your answers to ensure that they are the same.

(a) $\vec{r}(t) = \sin(t)\langle 2t, t^2, \arctan(t)\rangle$

(b) $\vec{s}(t) = \vec{r}(2^t)$, where $\vec{r}(t) = \langle t + 2, \ln(t), 1\rangle$

(c) $f(t) = \langle\cos(t), \sin(t), t\rangle \cdot \langle -\sin(t), \cos(t), 1\rangle$

(d) $\vec{r}(t) = \langle\cos(t), \sin(t), t\rangle \times \langle -\sin(t), \cos(t), 1\rangle$

36. The left side of Figure 10.3.10 shows the curve described by the vector-valued function $\vec{r}$ defined by

$$\vec{r}(t) = \left\langle 2t - \frac{1}{2}t^2 + 1, t - 1 \right\rangle.$$

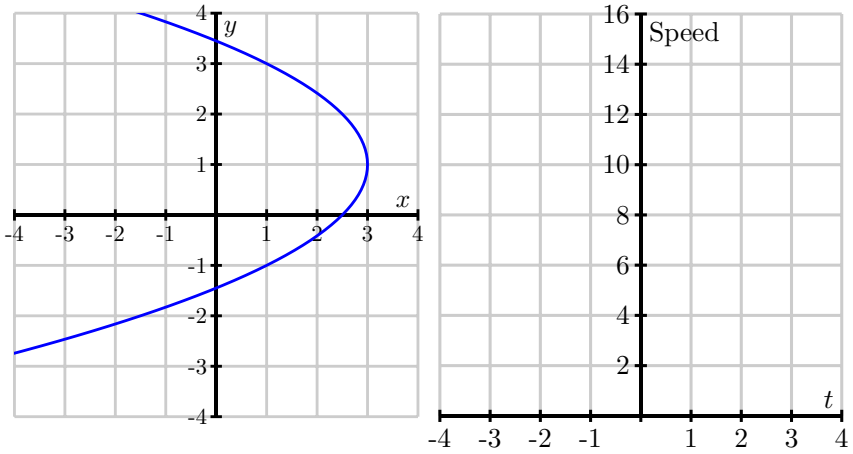


Figure 10.3.10 The curve $\vec{r}(t) = \left\langle 2t - \frac{1}{2}t^2 + 1, t - 1 \right\rangle$ and its speed.

(a) Find the object's velocity $\vec{v}(t)$.

(b) Find the object's acceleration $\vec{a}(t)$.

(c) Indicate on the left of Figure 10.3.10 the object's position, velocity and acceleration at the times $t = 0, 2, 4$. Draw the velocity and acceleration vectors with their tails placed at the object's position.

(d) Recall that the speed is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$. Find the object's speed and graph it as a function of time t on the right of Figure 10.3.10. When is the object's speed the slowest? When is the speed increasing? When is it decreasing?

(e) What seems to be true about the angle between $\vec{v}$ and $\vec{a}$ when the speed is at a minimum? What is the angle between $\vec{v}$ and $\vec{a}$ when the speed is increasing? when the speed is decreasing?

(f) Since the square root is an increasing function, we see that the speed increases precisely when $\vec{v} \cdot \vec{v}$ is increasing. Use the product rule for the dot product to express $\frac{d}{dt}(\vec{v} \cdot \vec{v})$ in terms of the velocity $\vec{v}$ and acceleration $\vec{a}$. Use this to explain why the speed is increasing when $\vec{v} \cdot \vec{a} > 0$ and decreasing when $\vec{v} \cdot \vec{a} < 0$. Compare this to part d.

(g) Show that the speed's rate of change is

$$\frac{d}{dt} \|\vec{v}(t)\| = \text{comp}_{\vec{v}}\vec{a}.$$

37. Consider the two vector-valued functions given by

$$\vec{r}(t) = \left\langle t + 1, \cos\left(\frac{\pi}{2}t\right), \frac{1}{1+t} \right\rangle$$

and

$$\vec{w}(s) = \left\langle s^2, \sin\left(\frac{\pi}{2}s\right), s \right\rangle.$$

(a) Determine the point of intersection of the curves generated by $\vec{r}(t)$ and $\vec{w}(s)$. To do so, you will have to find values of a and b that result in $\vec{r}(a)$ and $\vec{w}(b)$ being the same vector.

(b) Use the value of a you determined in (a) to find a vector form of the tangent line to $\vec{r}(t)$ at the point where $t = a$.

(c) Use the value of b you determined in (a) to find a vector form of the tangent line to $\vec{w}(s)$ at the point where $s = b$.

(d) Suppose that $z = f(x, y)$ is a function that generates a surface in three-dimensional space, and that the curves generated by $\vec{r}(t)$ and $\vec{w}(s)$ both lie on this surface. Note particularly that the point of intersection you found in (a) lies on this surface. In addition, observe that the two tangent lines found in (b) and (c) both lie in the tangent plane to the surface at the point of intersection. Use your preceding work to determine the equation of this tangent plane.

38. For each given function $\vec{r}$, determine $\int \vec{r}(t)\, dt$. In addition, recalling the Fundamental Theorem of Calculus for functions of a single variable, also evaluate $\int_0^1 \vec{r}(t)\, dt$ for each given function $\vec{r}$. For each of your calculations, you should write a sentence that states whether your result is a vector or a scalar and describes what is being measured.

(a) $\vec{r}(t) = \left\langle \cos(t), \frac{1}{t+1}, te^t \right\rangle$

(b) $\vec{r}(t) = \langle \cos(3t), \sin(2t), t \rangle$

(c) $\vec{r}(t) = \left\langle \frac{t}{1+t^2}, te^{t^2}, \frac{1}{1+t^2} \right\rangle$

39. A *central force* is one that acts on an object so that the force $\vec{F}$ is parallel to the object's position $\vec{r}$. Since Newton's Second Law says that an object's acceleration is proportional to the force exerted on it, the acceleration $\vec{a}$ of an object moving under a central force will be parallel to its position $\vec{r}$. For instance, the Earth's acceleration due to the gravitational force that the sun exerts on the Earth is parallel to the Earth's position vector as shown in Figure 10.3.11.

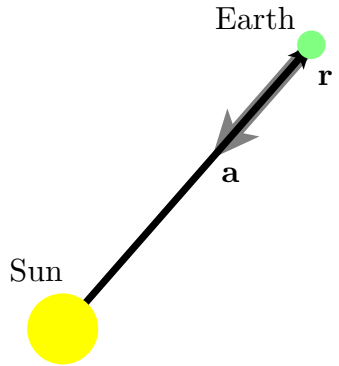


Figure 10.3.11 A central force.

(a) If an object of mass m is moving under a central force, the angular momentum vector is defined to be $\vec{L} = m\vec{r} \times \vec{v}$. Assuming the mass is constant, show that the angular momentum is constant by showing

that

$$\frac{d\vec{L}}{dt} = \vec{0}.$$

(b) Explain why $\vec{L} \cdot \vec{r} = 0$.

(c) Explain why we may conclude that the object is constrained to lie in the plane passing through the origin and perpendicular to $\vec{L}$.

10.4 Arc Length

Motivating Questions

- How can a definite integral be used to measure the length of a curve in 2- or 3-space?
- Is it possible to parameterize any curve with constant speed?

10.4.1 Introduction

In the previous sections, we introduced vector-valued functions, which provided us with a useful tool for thinking about curves in three-dimensional space. We have also seen how the calculus operations of limits, derivatives, and integrals apply to these functions. In single-variable calculus, you likely learned that displacement (net change in position) and total distance traveled for an object moving in one-dimension will not be the same if the object changes direction. For motion along a curve in two or more dimensions, we have a similar distinction. For instance, consider the curve in Figure 10.4.1.

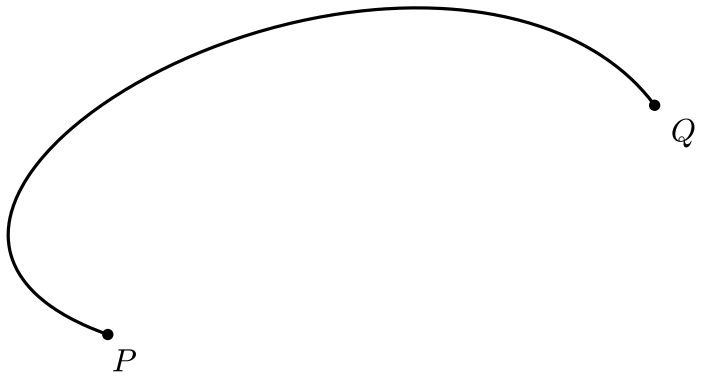


Figure 10.4.1 A curve C from the point P to the point Q

Remember that the distance between two points in space is given by

$$|PQ| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}.$$

However, if this curve is a portion of a road on which we drive a car, the distance traveled as measured by the change in the car's odometer value from point P to point Q will be different than the straight-line distance $|PQ|$ calculated by the distance formula. You likely saw how to calculate the length of a portion of a curve $y = f(x)$ in your single-variable calculus class. In this section, we will use the idea of vector-valued functions to find the lengths of a larger family of curves.

Preview Activity 10.4.1 In Preview Activity 10.2.1 and Preview Activity 10.3.1 we saw how the position reported by the location tracking system (LTS) of a self-driving car corresponds to a vector-valued function that describes the location of the car at a given time. We want to relate the location at different times to properties about how the car is being driven. In order to attract investors and drive shareholder value in *Steer Clear*, you decide to use the location tracking system's output to build a customized navigation and telemetry tool for a self-driving car. The first element you will need to build is a way to use the LTS output to calculate how far the car has driven in a given time.

(a) In your testing center (an abandoned parking garage), you drive your car around different levels of your testing center and note that the LTS recorded that the car had coordinates of $\langle 10, -5, 0\rangle$ at $t = 0$ and $\langle -8, 10, 12\rangle$ at $t = 30$ seconds. The units on each component of the coordinate vectors are meters. What is the distance between initial ($t = 0$) and final ($t = 30$) positions of the car? This distance is the length of the *displacement* vector of your total trip.

(b) Do you think the distance the car actually traveled is greater than, less than, or exactly the same as your answer to part (a)? Write a couple of sentences to explain your reasoning.

(c) In order to get more information for your telemetry system, you decide to pull the location data from your drive up the parking ramp every 10 seconds. The LTS gives the following data:

- $\langle 10, -5, 0\rangle$ at $t = 0$
- $\langle 6, 0, 6\rangle$ at $t = 10$
- $\langle -10, 2, 10\rangle$ at $t = 20$
- $\langle -8, 10, 12\rangle$ at $t = 30$

Calculate the distance between successive times and estimate how far your car went on the drive up the parking ramp.

(d) Do you think your estimate from part (c) is greater than, less than, or equal to the actual distance traveled? Write a couple of sentences to explain your reasoning.

(e) You look into the documentation on your location tracking system and see that you can specify how often the software will output location data. You decide that getting the location every half second seems like a good idea. How many data points will that correspond to for your drive up the ramp? Write a couple of sentences to describe what steps you would take with this new location data to calculate a better estimate on the distance traveled by your car. How might you go about finding the *exact* distance traveled by your car?

10.4.2 Arc Length

The central question we want to answer in this section is "Given a curve C from P to Q, what is the length of C?" We will answer this question for well-behaved curves by applying the classic calculus approach to the ideas you investigated in the Preview Activity. For step 1 of the CCA, we approximate the length of the curve C by computing the length of line segments between points on C. We will call these intermediate points $\vec{r}_0, \vec{r}_1, ..., \vec{r}_n$. Remember that we are

specifying location using vectors graphed in standard position. Using vectors to specify position means that the corresponding line segments between successive positions can be represented by the vector difference $\Delta\vec{r}_i = \vec{r}_i - \vec{r}_{i-1}$.

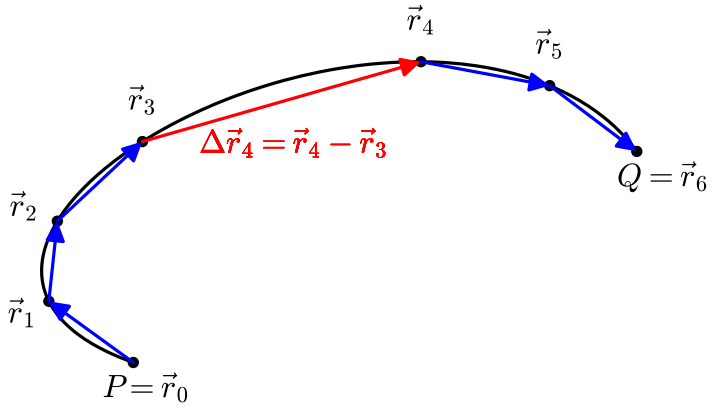


Figure 10.4.2 A curve C oriented from the point P to the point Q split into 6 parts with endpoints $\vec{r}_0, \vec{r}_1, ..., \vec{r}_6$

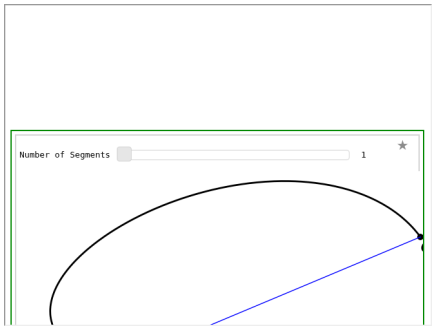
This leads to an approximation for the length of C given by the sum of the lengths of the vectors $\Delta\vec{r}_i = \vec{r}_i - \vec{r}_{i-1}$. Using Δx_i, Δy_i, and Δz_i to denote the change in each component, the magnitude of $\Delta\vec{r}_i$ is determined by the distance formula as we saw in the previous chapter:

$$\|\Delta\vec{r}_i\| = \sqrt{\Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2}.$$

The length of the curve can now be approximated by the sum of these lengths, which gives

$$\begin{aligned} L &\approx \sum_{i=1}^{n} \|\Delta\vec{r}_i\| \\ &\approx \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2}. \end{aligned}$$

Use the slider on Figure 10.4.3 to change the number of segments used in the approximation of the arc length of C. Notice that as the number of segments increases, the difference between the actual length of the curve and the line segments gets smaller. You can also see how the sum of the lengths of the blue segments will approach the true length of the curve, which is approximately 6.3286.



Standalone
In Context
Embed

Figure 10.4.3 A plot of curve in space

Conceptually, you can use Figure 10.4.3 to see how more intermediate points will give rise to a better estimate for the true length of the curve, but we have not quantified how the approximation changes on a finer scale (step 2 of the CCA). In order to complete step 2 of the CCA, we will need to incorporate steps in our parameter t into this calculation.

In general, it is easiest to generate the positions $\vec{r}_i$ by evaluating a parameterization at different t-values. We will use $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ as a parameterization of C to specify points between P and Q. Specifically, if $\vec{r}(t)$ is a parameterization of C with $P = \vec{r}(a)$ and $Q = \vec{r}(b)$, we can divide C into n parts given by $\vec{r}_i = \vec{r}(t_i)$ where $t_i = a + i\Delta t$ with $\Delta t = \frac{b-a}{n}$. This divides C into parts that correspond to equal steps in the parameter t. (Note that these parts will, in general, *not* correspond to parts with equal length.)

This process leads to a specific approximation used in the classic calculus approach. Since this approximation approaches the actual length as we use smaller and smaller steps in the parameter, we define the arc length as follows:

Definition 10.4.4 The arc length of C, a curve in space, is given by

$$L = \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2}$$

where $\Delta\vec{r}_i = \vec{r}_i - \vec{r}_{i-1} = \langle \Delta x_i, \Delta y_i, \Delta z_i\rangle$ such that $\vec{r}_i = \vec{r}(t_i)$ where $t_i = a + i\Delta t$ with $\Delta t = \frac{b-a}{n}$. ◆

We now write out the second step of the CCA described above more explicitly in order to understand how to efficiently calculate the arc length. Since we want to shrink the step size Δt to get better approximations, we now rewrite the approximation to allow us to see how the approximation changes as we consider smaller values of Δt. To do this, note that assuming $x(t)$ is differentiable, since $x'(t_i) \approx \Delta x_i/\Delta t$, we can write that $\Delta x_i \approx x'(t_i)\Delta t$. This applies equally well for both y and z, so this leads us to

$$\begin{aligned} L &\approx \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2} \\ &\approx \sum_{i=1}^{n} \sqrt{(x'(t_i)\Delta t)^2 + (y'(t_i)\Delta t)^2 + (z'(t_i)\Delta t)^2} \\ &\approx \sum_{i=1}^{n} \sqrt{(x'(t_i))^2 + (y'(t_i))^2 + (z'(t_i))^2}\ \Delta t. \end{aligned}$$

This refines the approximation and gives an explicit statment that completes the second step in the classic calculus approach.

Note that we have set up the approximation as a Riemann sum. In particular, this approximation is a Riemann sum of the function $\sqrt{(x'(t_i))^2 + (y'(t_i))^2 + (z'(t_i))^2}$ which will simplify as Δt goes to 0. Thus the limit of the Riemann sum from the definition of arc length gives a definite integral

$$\begin{aligned} L &= \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2} \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(x'(t_i))^2 + (y'(t_i))^2 + (z'(t_i))^2}\ \Delta t \\ &= \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}\ dt. \end{aligned}$$

The following theorem summarizes how the definition of arc length can be evaluated as a definite integral using a parameterization.

Theorem 10.4.5 The length of a curve. *If $\vec{r}(t)$ parameterizes a curve C on an interval $[a, b]$ for which $\vec{r}\,'(t)$ exists at each point, then the* length *L of C, is given by*

$$L = \int_a^b \|\vec{r}\,'(t)\| \, dt. \tag{10.4.1}$$

Note that formula (10.4.1) applies to curves in any dimensional space. Moreover, this formula has a natural interpretation: If $\vec{r}(t)$ records the position of a moving object, then $\vec{r}\,'(t) = \vec{v}(t)$ is the object's velocity and $\|\vec{r}\,'(t)\|$ its speed. Formula (10.4.1) says that we can integrate the speed of an object traveling over the curve to find the distance traveled by the object. Thinking in terms of units, if we are measuring distance in meters and time in seconds, the units on $\|\vec{r}\,'(t)\|$ will be meters per second and the units on dt will be seconds. Thus, the units on the integral will be meters, which is exactly what we expect when finding the *length* of a curve or distance traveled by an object along a curve. This formula is analogous to the formula for finding the arc length of a graph from single-variable calculus.

Example 10.4.6 In this example, we will use a parameterization of a line to compute the arc length of the line between $(1, 2, 3)$ and $(2, 0, 5)$. An easy-to-find parameterization will have the base point $(1, 2, 3)$ and direction vector given by the difference between our points, $\langle 2-1, 0-2, 5-3\rangle = \langle 1, -2, 2\rangle$. This leads to a parameterization of the line as $\vec{r}(t) = \langle 1, 2, 3\rangle + t\langle 1, -2, 2\rangle$. Notice that with this parameterization, restricting to $0 \le t \le 1$ gives a domain for $\vec{r}(t)$ for the line segment from $(1, 2, 3)$ when $t = 0$ to $(2, 0, 5)$ when $t = 1$. Note that the speed of this parameterization is constant, $\|\vec{r}\,'(t)\| = \sqrt{1^2 + (-2)^2 + 2^2} = 3$.

We can use Theorem 10.4.5 to calculate the length of this path.

$$L = \int_a^b \|\vec{r}\,'(t)\| \, dt. = \int_0^1 3 \, dt = 3.$$

This example was simple enough that we didn't actually need to use Theorem 10.4.5 because we had a straight line path which we could have calculated the length of using the distance formula. This example will be revisited in Example 10.4.8. ▲

Activity 10.4.2 In this activity, we will use parameterizations to find the length of some common curves.

(a) Parameterize a circle of radius 3 centered at the origin. Give bounds on the parameter.

(b) Use your parameterization from the previous part in the definite integral of Theorem 10.4.5 to calculate the circumference of a circle of radius 3.

(c) Find the exact length of the spiral defined by $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), t\rangle$ on the interval $[0, 2\pi]$.

(d) Explain why your result for the length of the spiral is larger than the circumference of the circle of the same radius.

Notice here that while we focused on the circle of radius 3, changing to the more general radius of R for the circle gives a parameterization of $\langle R\cos(t), R\sin(t)\rangle$ with $0 \le t \le 2\pi$. The work to use the arc length integral to find the circumference of this circle requires only small modifications to the work you did in this activity.

Exercise 19 will apply Definition 10.4.4 and Theorem 10.4.5 to arrive at a formula for the length of a graph given by $y = f(x)$ that may be familiar from single-variable calculus courses.

10.4.3 Moving with Unit Speed

In Activity 10.4.2, we considered parameterizations of a circle and a helix. The computation of the arc length in these problems was made easier by the fact that the speed of an object with motion described by the parameterizations used did not depend on the time t. That is, the speed of an object moving according to the parameterization was *constant*. Moving with constant speed is always feasible. For example, we can set our self-driving car to move at a constant speed of 60 miles per hour, and the car then moves with constant speed. (Of course, this is probably only recommended on a road where no obstacles or other vehicles will appear.) However, finding a parameterization with constant speed may involve some complex algebra, as we will see while working through this section.

We will focus on finding a parameterization with unit speed. That is, one for which $\|\vec{v}(t)\| = \|\vec{r}\,'(t)\| = 1$ for all values of t. We call such a parameterization a **unit speed parameterization** of the curve. The key idea involves recognizing that, in general, distance along the curve is not the same as the parameter being used. If an object moves with constant speed one, then the numerical value of the distance it travels *is* the numerical value of the time elapsed. For example, if you walk at exactly 1 meter per second for 47 seconds, how far have you gone? If you walk at exactly 1 meter per second, how long will it take you to travel 47 meters? The answer to both is 47 but the units are different. Remember that we do not have to walk in a straight line (i.e., with constant direction) to keep speed constant.

As a more concrete illustration of this idea, consider the curve defined by the parabola $y = x^2/2$ with $x \in [0, 2]$. We can parameterize this curve by $\vec{r}(t) = \langle t, t^2/2\rangle$ for $t \in [0, 2]$. Figure 10.4.7 shows a plot of this parabola with equally-spaced points in the parameter t on the left. You can see in the plot on the left that equally-spaced points based on parameter values are *not* separated by equal distances along the curve. On the right, however, the points are equally-spaced by arc length. By this we mean that the distance *along the curve* between the point labeled $s = 0.5$ and the point labeled $s = 1.0$ is 0.5. Similarly, the distance along the curve between the point labeled $s = 0.5$ and the point labeled $s = 2.5$ is 2. For this reason, a unit speed parameterization is also often called a **parameterization by arc length**.

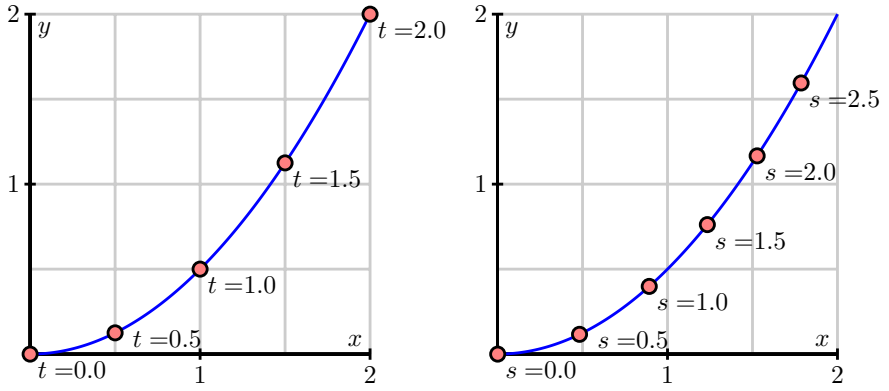


Figure 10.4.7 A plot of $y = x^2/2$ with $x \in [0, 2]$ with points equally spaced in the parameter value (on left) and in arc length (on right)

Terminology. Many other textbooks will refer to the unit speed parameterization of a curve as the arc length parameterization. Conceptually, we motivated this parameterization by wanting to move with a constant speed of 1, which is why we have opted to use the name "unit speed parameterization". After seeing the motivation for wanting to move with unit speed, you can see the correspondence between the arc length and the parameter necessary to give this new description.

In many ways, a parameterization with unit speed is a natural parametrization. Consider an interstate highway cutting across a state. One way to parametrize the curve defined by the highway is to drive along the highway and record our position at every time t, thus creating a function $\vec{r}(t)$. If we encounter an accident or road construction, however, this parametrization might not be at all relevant to another person driving the same highway. A parameterization using unit speed is like using the mile markers on the side of the road to specify our position on the highway. If we know how far we've traveled along the highway, we know exactly where we are. Another way to think about this is that by driving at exactly one mile per hour, we can put mile markers down every hour.

Example 10.4.8 In this example, we look at a linear path and explore how to write a parameterization that moves with unit speed. We will look at the parameterization of the line through $(1, 2, 3)$ and $(2, 0, 5)$. We are using these values to make the calculations less abstract but there is nothing special about these values and the methods will work for any linear parameterization. In Example 10.4.6, we calculate that the length of the path of the parameterization in this case is 3.

We want to convert this to a parameterization with unit speed. To think of this another way, we are aiming to have the parameter value when we reach the point $(2, 0, 5)$ be equal to the distance from the base point $(1, 2, 3)$. This will mean we seek a parameterization $\vec{r}_1(s)$ where the domain $0 \leq s \leq 3$ corresponds to the line segment from $(1, 2, 3)$ to $(2, 0, 5)$. The speed of our original parameterization is

$$\|\vec{r}\,'(t)\| = \left\|\frac{d\vec{r}}{dt}\right\| = \|\langle 1, -2, 2\rangle\| = \sqrt{9} = 3.$$

This means that we need to adjust our parameterization to travel 3 times slower, which can be accomplished by replacing t in our parameterization with $s = t/3$. In other words, we adjust the parameter used to make sure the speed with respect to the new parameter will be one. This adjustment is easy to

do because the amount we needed to change the parameter did not change at different times.

Our proposed unit speed parameterization is

$$\vec{r}_1(s) = \vec{r}\left(\frac{t}{3}\right)\langle 1,2,3\rangle + \left(\frac{t}{3}\right)\langle 1,-2,2\rangle.$$

Computing the speed here gives

$$\|\vec{r}_1(s)\| = \left\|\frac{1}{3} = \langle 1,-2,2\rangle\right\| = \frac{\sqrt{9}}{3} = 1.$$

Algebraically, we can describe this change by thinking of $\vec{r}(t)$ as the "old" parameterization with parameter t, and we aim to replace t with a function $s(t)$ so that $\|\vec{r}_1\,'(s(t))\| = 1$ for all t. For this linear path, the choice of $s(t) = t/3$ means that applying the chain rule to $\|\vec{r}_1\,'(s)\|$ gives

$$\|\vec{r}_1\,'(s(t))\| = \left\|\frac{d\vec{r}}{ds}\right\|\left|\frac{ds}{dt}\right| = 3\frac{1}{3} = 1. \tag{10.4.2}$$

▲

The left part of equation (10.4.2) shows the important idea needed to find a unit speed parameterization: we can adjust the speed of the parameterization. This requires that we pick a transformation $s(t)$ that satisfies

$$\left\|\frac{d\vec{r}}{ds}\right\|\left|\frac{ds}{dt}\right| = 1.$$

Algebraically, this is ***very*** difficult to write out in closed form. However, it will be possible because the arc length traveled as a function of the parameter cannot decrease. If we make a two-dimensional plot with the parameter value on the horizontal axis and the arc length traveled on the vertical axis, then we have a graph that passes the horizontal line test. This means that this function has an inverse! An example of this kind of plot is shown in Figure 10.4.9. The point highlighted on the graph indicates that at the parameter value $t = 4$, the arc length from the point with $t = 0$ (or the distance traveled by an object following the curve from $t = 0$ to $t = 4$) is approximately 48.134. In other words, when finding a parameterization by arc length or unit speed parameterization, we will need to arrive at the same point on the curve when the parameter value is 48.134.

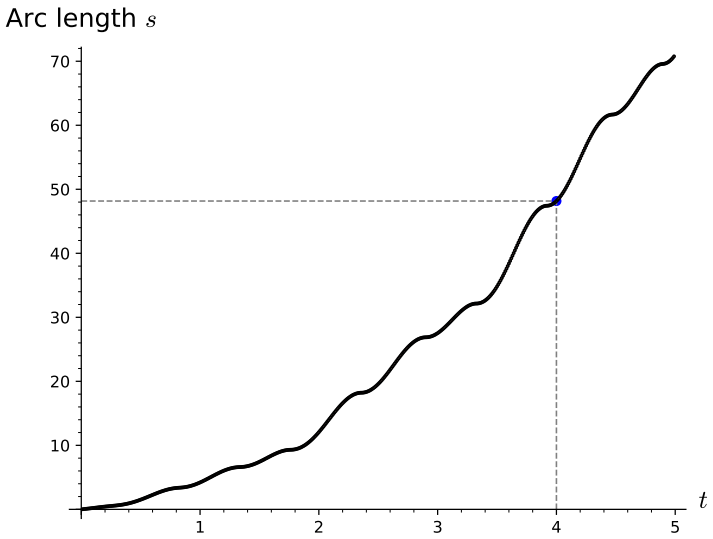


Figure 10.4.9 A plot of the arc length traveled as a function of the parameter value to that point for $\vec{r}(t) = \langle 3t\cos(4t)\sin(3-2t), t-\cos(2+t)\rangle$

In the next example, we will work through how to transform a parameterization with variable speed into one that has unit speed.

Example 10.4.10 In this example, we will consider the curve C parameterized by

$$\vec{r}(t) = \left\langle t^2, \frac{8}{3}t^{3/2}, 4t \right\rangle$$

for $t \geq 0$. Our goal is to find a unit speed parameterization of C. The first step is to find the speed of our current parameterization:

$$\begin{aligned}
\|\vec{r}\,'(t)\| &= \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} \\
&= \sqrt{(2t)^2 + \left(4t^{1/2}\right)^2 + (4)^2} \\
&= \sqrt{4t^2 + 16t + 16} \\
&= \sqrt{(2t+4)^2} \\
&= 2t + 4.
\end{aligned}$$

An important element of this example is that our speed is not constant but still algebraically easy enough to work with for later calculations. Note that because we have stated that $t \geq 0$ we only need to consider the positive square root calculation.

Next, we will calculate $s(a)$, the arc length traveled from $t = 0$ to $t = a$ along the curve C in order to help us figure out how to transform our parameterization into one with unit speed. We use Theorem 10.4.5 to compute the arc length of C as t goes from 0 to a.

$$\begin{aligned}
s(a) &= \int_0^a \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}\, dt \\
&= \int_0^a 2t + 4\, dt \\
&= \left(t^2 + 4t\right)\Big|_0^a \\
&= a^2 + 4a
\end{aligned}$$

This computation shows us that the arc length traveled from $t = 0$ to $t = a$ along the curve C is $a^2 + 4a$. Remember that we want to transform our parameterization to have unit speed, which will mean that the arc length up to a point and the parameter value corresponding to that point will be the same value. We want to state a function that has an input of arc length along the curve C and outputs the time at which the original parameter reaches that location. In other words, we want the inverse function of $s(a) = a^2 + 4a$, because the output of $s(a)$ is arc length up to the point corresponding to the parameter value a. In Figure 10.4.11, we show the graph of the function $s(a)$.

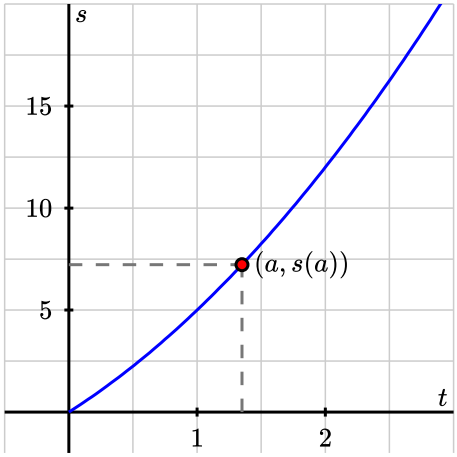


Figure 10.4.11 The arc length function for the curve $\vec{r}(t)$

Since $a \geq 0$, we can solve the equation $s = a^2 + 4a$ (or $a^2 + 4a - s = 0$) for a to obtain $a(s) = \frac{-4+\sqrt{16+4s}}{2} = -2 + \sqrt{4+s}$. Note again that by restricting to positive parameter values, we only need to consider one solution. Another way of thinking about this is that the point labeled $(a, a^2 + 4a)$ in Figure 10.4.11 could also be labeled $(-2 + \sqrt{4+s}, s)$.

Let's take a moment to further discuss the relationship between our two functions and their composition. The original parameterization $\vec{r}(t)$ has inputs of parameter values (t) and outputs corresponding to locations on our curve C. The function value $s(a)$ gives the arc length along C from $t = 0$ to $t = a$, so the inputs to this function are parameter values and the outputs are arc lengths. The new function $a(s)$ is the inverse of s, and thus its input is arc length. (Specifically, the input is the arc length traveled from $t = 0$ to $t = a$ along the curve C.) The output from $a(s)$ is a parameter value of the original parameterization corresponding to when the arc length is s.

We can now compose our result of $a = -2 + \sqrt{4+s}$ with the original parameterization to get our unit speed parameterization. Specifically, we claim that $\vec{r}(a(s))$ will be a unit speed parameterization. Remember that a unit speed parameterization of a curve is equivalent to the idea that "parameter value = time elapsed = arc length traveled".

The parameterization $\vec{r}(a(s))$ is a unit speed parameterization because the parameter s *is* the amount of arc length traveled. In terms of the mile marker analogy, the value of s is the value you see on the mile marker. The input of the composition $\vec{r}(a(s))$ will be the arc length traveled. The "inner" function, $a(s)$, outputs the parameter value *of the original parameterization* where that arc length is achieved. (Keep in mind Figure 10.4.11, which illustrates the inverse relationship between a and s.) This means that $\vec{r}(a(s))$ outputs the location on C where an arc length of s is achieved.

To illustrate the concepts of the previous paragraph concretely, we return

to the algebra of our problem. For our example, the composition $\vec{r}(a(s))$ gives the following as the unit speed parameterization:

$$\vec{r}_1(s) = \left\langle \left(-2+\sqrt{4+s}\right)^2, \frac{8}{3}\left(-2+\sqrt{4+s}\right)^{3/2}, 4\left(-2+\sqrt{4+s}\right) \right\rangle.$$

Exercise 18 asks you to compute the speed of $\vec{r}_1(s)$ to verify that the speed is in fact 1 for all values of s.

The crucial step, both algebraically and conceptually, in our work was the inverse function of $s(a)$, which was relatively simple for this example. In general, this can be difficult to write out explicitly. ▲

These examples illustrate a general method for finding a unit speed parameterization. Evaluating an arc length integral *and* finding the formula for the inverse of a function can each be difficult. Although this process is always theoretically possible, it is frequently not practical to parameterize a curve in terms of arc length. However, we can guarantee that such a parameterization exists, and this observation plays an important role in the next section.

10.4.4 The Driver or The Road?

In Activity 10.4.3, we will look at several measurements as we consider different drivers completing one lap around the track seen in Figure 10.4.12.

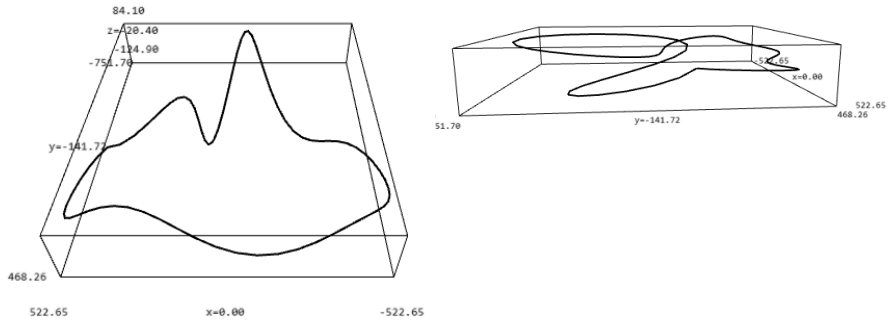


Figure 10.4.12 A plot of the racetrack with scale in meters

All of the drivers are going the same way around the track, all start at the same location at $t = 0$, all complete one lap of the track, and all of the cars have perfect grip of the road (the race cars are never sliding). The picture of the racetrack above is given so you have an example to help you think about the tasks in this problem, not because any particular feature of the track needs to be considered for the following activity.

Activity 10.4.3 Is it a property of the driver or the road? In this activity, we want to determine if the different measurements we have described are a property of the driver or the road. A measurement is a **property of the driver** provided that the values of that measurement *can be different* for different drivers when measured at the same location on the racetrack. A measurement is a **property of the road** provided that different drivers *must have the same values* when measured at the same location on the racetrack.

(a) Let's start by looking at a couple of easy measurements. Is the time elapsed a property of the driver or the road? Be sure to explain your answer.

Hint. Remember that we are *not* asking if the driver or road can control time. We are asking if the elapsed time is the same for all drivers at a

fixed point on the racetrack or if drivers can have a different elapsed time to a particular location on the racetrack.

(b) Is position (the location of the car on the racetrack) a property of the driver or the road? Be sure to explain your answer.

Hint. You should make sure you understand that the measurements for different drivers must be made at the same location on the track, as stated above. Your explanation sentence may sound a little silly.

(c) *Speedometer Reading.*

Now that we are warmed up, let's look at some more interesting measurements. The car's speedometer reading measures how fast (as a scalar) the car is moving. Is the car's speedometer reading a property of the driver or the road? Be sure to explain your answer.

(d) What vector calculus quantity is the speedometer reading?

Hint. The vector calculus quantities we have covered are time, position, velocity, acceleration, speed, and arc length.

(e) *Odometer Reading.*

The racecar's odometer measures the distance traveled by the car. Every car's odometer is set to be zero at the start of the race. Is the car's odometer reading a property of the driver or the road? Be sure to explain your answer.

When we are using the analogy of the racetrack for a curve in space, each different driver corresponds to a different parameterization of the curve. If a measurement is a property of the driver, that measurement can take different values for different parameterizations. If a measurement is a property of the road, then every parameterization *must* have the same measured value when considered at the same location on the curve. In short, parameterizations are drivers and the curve in space is the racetrack.

Summary

- The integration process shows that the length L of a smooth curve defined by $\vec{r}(t)$ on an interval $[a, b]$ is

$$L = \int_a^b \|\vec{r}'(t)\| \, dt.$$

- A parameterization with unit speed is useful because when we move with unit speed, the parameter value is equal to the arc length traveled to that point.

10.4.5 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Please keep this in mind when working WeBWorK exercises.

1. Find the length of the curve

$$x = -(5 + t), \quad y = 3 - 5t, z = 4 + 2t,$$

for $2 \leq t \leq 4$.

length = _____

(Think of second way that you could calculate this length, too, and see that you get the same result.)

2. Find the length of the given curve:

$$\mathbf{r}(t) = (-4t, -5\sin t, -5\cos t)$$

where $-5 \le t \le 4$.

3. Consider the path $\mathbf{r}(t) = (6t, 3t^2, 3\ln t)$ defined for $t > 0$.

Find the length of the curve between the points $(6, 3, 0)$ and $(24, 48, 3\ln(4))$.

4. Starting from the point $(-2, -1, 5)$, reparametrize the curve $\mathbf{x}(t) = (-2 - 2t, -1 + 2t, 5 - 2t)$ in terms of arclength.

$\mathbf{y}(s) = ($ _____, _____,_____ $)$

5. Consider the path $\mathbf{r}(t) = (10t, 5t^2, 5\ln t)$ defined for $t > 0$.

Find the length of the curve between the points $(10, 5, 0)$ and $(20, 20, 5\ln(2))$.

6. Find the arclength of the curve $\mathbf{r}(t) = \langle 7\sqrt{2}t, e^{7t}, e^{-7t}\rangle$, $0 \le t \le 1$

7. Find the arclength of the curve $\mathbf{r}(t) = \langle 8t^2, 2\sqrt{8}\ t, \ln t\rangle$, for $1 \le t \le 4$.

L = _____

8. Find the length of the curve $\mathbf{r}(t) = \mathbf{i}\ + 9t^2\mathbf{j} + t^3\mathbf{k}$, $0 \le t \le \sqrt{45}$

L = _____

9. Calculate the length of the path over the given interval.

$$(9t^2, 7t^2 - 1),\ 0 \le t \le 4$$

10. Calculate the length of the path over the given interval.

$$c(t) = (3t^2, 4t^3),\ 1 \le t \le 3$$

11. Calculate the length of the path over the given interval.

$$(\sin 3t, \cos 3t),\ 0 \le t \le \pi$$

12. Calculate the length of the path over the given interval.

$$(\sin\theta - \theta\cos\theta, \cos\theta + \theta\sin\theta),\ 0 \le \theta \le 3$$

13. Find the length of the spiral $c(t) = (t\cos t, t\sin t)$ for $0 \le t \le 3\pi$ to three decimal places. Hint: use the formula

$$\int \sqrt{1+t^2}\,dt = \frac{1}{2}t\sqrt{1+t^2} + \frac{1}{2}\ln\left(t + \sqrt{1+t^2}\right) + C$$

14. Compute the length of the curve $\mathbf{r}(t) = 5t\,\mathbf{i} + 9t\,\mathbf{j} + (t^2 - 1)\,\mathbf{k}$ over the interval $0 \le t \le 4$.

HINT: use the formula

$$\int \sqrt{t^2+a^2}\,dt = \frac{1}{2}t\sqrt{t^2+a^2} + \frac{1}{2}a^2\ln\left(t + \sqrt{t^2+a^2}\right) + C$$

L = _____

15. Consider the curve $\mathbf{r} = (e^{1t}\cos(1t), e^{1t}\sin(1t), e^{1t})$.

Compute the arclength function $s(t)$: (with initial point $t = 0$).

16. Remember that parameterizing in terms of arc length is the same as the unit speed parameterization.

Starting from the point $(2, -3, 5)$, reparametrize the curve $\mathbf{r}(t) = (2 + 3t, -3 + 1t, 5 - 2t)$ in terms of arclength.

$\mathbf{r}(t(s)) = ($ _____, _____, _____ $)$

17. Remember that parameterizing in terms of arc length is the same as the unit speed parameterization.

Starting from the point $(3, 4, -1)$ reparametrize the curve

$$\mathbf{r}\,(t) = (3 + 1t)\,\mathbf{i} + (4 + 1t)\,\mathbf{j} + (-1 + 1t)\,\mathbf{k}$$

in terms of arclength.

$\mathbf{r}\,(t\,(s)) =$ _____ $\mathbf{i}+$ _____ $\mathbf{j}+$ _____ $\mathbf{k}$

18. Verify that the parameterization $\vec{r}_1(s)$ of the curve C in Example 10.4.10 has unit speed.

19. We can adapt the arc length formula to curves in 2-space that define y as a function of x as the following activity shows.

Let $y = f(x)$ define a smooth curve in 2-space. Parameterize this curve and use Equation (10.4.1) to show that the length of the curve defined by f on an interval $[a, b]$ is

$$\int_a^b \sqrt{1 + [f'(t)]^2}\, dt.$$

20. In this exercise, we will make the unit speed parameterization for the curve parameterized by $\vec{r}(t) = \langle \frac{3}{2}t^2, 2t^2, \frac{5}{3}t^3 \rangle$ for $t \in [0, a]$.

(a) Calculate the speed for the parameterization $\vec{r}(t) = \langle \frac{3}{2}t^2, 2t^2, \frac{5}{3}t^3 \rangle$. Factor out as large a power of t out of the square root as possible.

(b) Use your speed calculation from above to compute the arc length on our curve C as t goes from 0 to a. Your answer should be in terms of a.

Hint. After factoring out the power of t, you should have a nice substitution problem with $u = t^2 + 1$.

(c) Your previous step should have created a function $s(a)$ which outputs the arclength traveled on C for t from 0 to a. Find the inverse function, $a(s)$, that outputs the time past 0 that is needed to travel s along the curve.

(d) Verify that the composition of $\vec{r}(a(s))$ is a parameterization of C with unit speed.

21. In this exercise, we will look at transforming a parameterization for part of the graph of $y = x^2$ to move with unit speed. While the curve is relatively simple, this conversion to a unit speed parameterization can be very difficult (or impossible even!) to write out algebraically.

(a) Parameterize the curve given by $y = x^2$ for $-2 \le x \le 2$

(b) Compute the speed for your parameterization and set up the integral (from (10.4.1)) to calculate the arclength for the given portion of the parabola.

(c) We want to compute the arclength to some value $a \in [-2, 2]$ so that

we can find the inverse of the function $s(a)$. So we need to compute the integral with a as the upper bound (rather than -2). Write out this integral.

You can either evaluate this integral with trig subsitutions and identities involving hyperbolic functions or use computer algebra system to evaluate this integral. Either way, confirm that the arclength along the curve $y = x^2 x$ from $x = -2$ to $x = a$ is given by

$$s = \frac{a}{2}\sqrt{1+4a^2} + \frac{1}{4}\sinh^{-1}(2a) + \sqrt{17} + \frac{1}{4}\sinh^{-1}(4)$$

(d) Write a few sentences about why you would not want to try to solve for a as a function of s.

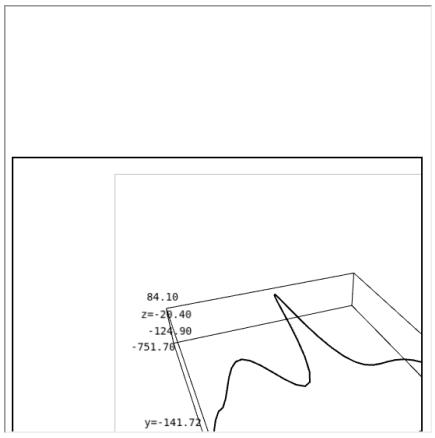
10.5 The TNB Frame

Motivating Questions

- How can we measure the direction of travel on a parametrized curve?
- How can we measure the direction of turning on a parameterized curve?
- How can we measure the axis of rotation along a parameterized curve?

10.5.1 Introduction

In Section 10.2, we saw how a vector-valued function of one variable will graphically correspond to a curve in space. While the application of derivatives, integrals, and limits apply to these functions componentwise, we will need to apply our calculus and vector tools *together* to measure important properties of these curves. Conceptually, you should think about the parameterization of the curve in space as a description on how to travel through those points in space. Remember that parameterizations are not unique: there is more than one way to walk the same path!



Standalone
In Context
Embed

Figure 10.5.1 A plot of a racetrack

For example, consider the curve in space that we are exploring to be a race track as shown in Figure 10.5.1. Not every driver will travel along this racetrack in the same way, but every driver has to go through these same points in space, since we assume that everyone will stay on the track and not take shortcuts.

In the vocabulary of this chapter, any parameterization of this curve has to contain the same points, but can possibly go through these (x, y, z)-coordinates at different parameter values. More generally, some properties we measure will depend on the particular way that a driver goes along the course and some properties will be the same for all drivers when measured at the same location on the track.

Preview Activity 10.5.1 As CEO and Head of Engineering at *Steer Clear*, you decide that you are almost ready to start testing your self-driving car out on the road. Your navigation and telemetry software will use the location tracking system (LTS) to determine all of the important information about how the car is moving and how adjustments need to be made. Before you start programming your software, you decide to drive on a quiet, country road and collect data from the LTS to use as test data for your programming. In other words, you will drive on a section of road you already have mapped out in order to check that your software is calculating the correct information. The map in Figure 10.5.2 below shows the path you plan to take on the country road (with the direction given by the arrows on the plot). Notice that at points B and F, the road crosses itself to go in a different direction.

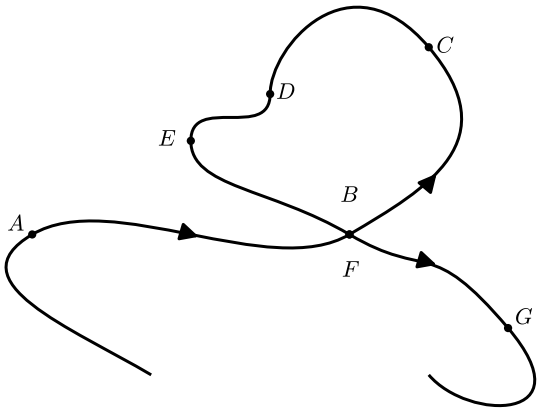


Figure 10.5.2 The path taken on a country road to collect data with your LTS

(a) At each of the labeled points on the curve, draw a vector in the direction of travel. Write a sentence about how you are determining this vector at the various points.

(b) At each point, decide whether the car is turning left or right. Write a sentence about how you are determining this turning based on the plot given.

Preview Activity 10.5.1 highlights the intuitive nature of measuring the direction of travel and the direction of turning given a plot of a curve in space. We will now look at how to use location information in the form of a parameterization of a curve to precisely measure the direction of travel and the direction of turning at a particular point on the curve.

10.5.2 Direction of travel

Based on our work in Subsection 10.3.2, if $\vec{r}(t)$ is a parameterization of a curve in space, then the velocity vector, $\vec{v}(t) = \vec{r}\,'(t)$, is in the direction of travel. Since the speed, $\|\vec{v}(t)\|$, is not always 1 (the velocity is not always a unit vector), we define the **direction of travel** using a unit vector as follows.

Definition 10.5.3 The **unit tangent vector** to a curve, denoted $\vec{T}(t)$, is defined as

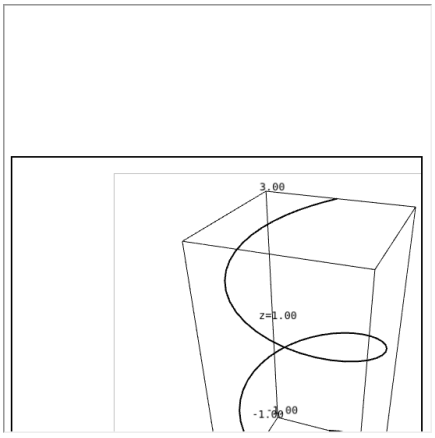
$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|}$$

and measures the direction of travel along a parameterized curve. ◆

By definition, the unit tangent vector exists as long as the velocity exists and is not the zero vector. If the velocity vector is $\vec{0}$, then the object is not moving, thus the direction of travel would not make sense at that instant. If there is a jump or discontinuity in the derivative of $\vec{r}$ so that $\vec{v}$ does not exist, the direction of travel at that instant does not make sense because there is not a consistent way to define how the object is moving.

The definition of $\vec{T}$ allows us to separate the velocity of a moving object into its magnitude (speed) and direction (unit tangent). Using a unit vector allows us to express the direction of travel as a vector that is separate from the length of any related measurements. Remember that $\vec{T}$ changes along the curve and is calculated as a function of the parameter.

Example 10.5.4 For our first example, we will look computing the unit tangent vector of a helix. In particular, we will look at the helix traced out by $\vec{r}(t) = \langle \cos(t), \sin(t), \frac{t}{\pi} \rangle$, which is shown in Figure 10.5.5.



Standalone
In Context
Embed

Figure 10.5.5 A helix

Based on the parameterization $r(t)$ for this helix, the velocity vector is $\vec{v}(t) = \langle -\sin(t), \cos(t), \frac{1}{\pi} \rangle$. Thus, the speed is

$$\text{speed}(t) = \|\vec{v}(t)\| = \sqrt{(-\sin(t))^2 + (\cos(t))^2 + \left(\frac{1}{\pi}\right)^2} = \sqrt{1 + \frac{1}{\pi^2}}.$$

One of the useful and interesting ideas of this example is that the speed is constant. That is, no matter what parameter value we consider, the speed is

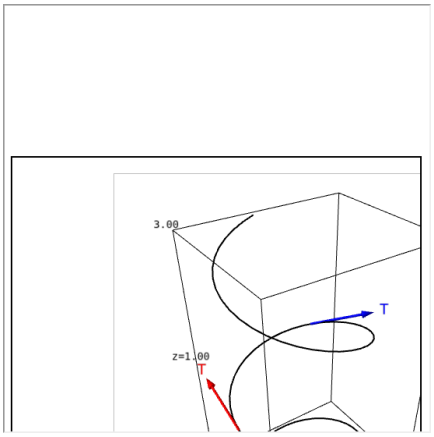
always $\sqrt{1+\frac{1}{\pi^2}}$. Using Definition 10.5.3, we find

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}}\langle -\sin(t), \cos(t), \frac{1}{\pi}\rangle.$$

To compute some specific examples, we have

$$\vec{T}\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}}\left\langle -1, 0, \frac{1}{\pi}\right\rangle$$
$$\vec{T}\left(\frac{5\pi}{4}\right) = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}}\left\langle \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{\pi}\right\rangle.$$

In Figure 10.5.6, $\vec{T}(\frac{\pi}{2})$ is plotted in red and $\vec{T}(\frac{5\pi}{4})$ is plotted in blue. Both of these vectors are unit length and tangent to the curve in the direction of travel for that instant.



Standalone
In Context
Embed

Figure 10.5.6 A helix with the unit tangent vector plotted at two locations

Note that we can also find $\vec{T}(\frac{\pi}{2})$ by computing $\vec{v}(\frac{\pi}{2}) = \langle -1, 0, \frac{1}{\pi}\rangle$ and shrinking this vector to be a unit vector. Because our definitions and theorems work for all values of t, we can apply them at a specific value of t. This has the added advantage that computations tend to be manageable when done by hand. The downside to this calculation is that you only have information for the one parameter value. This will only useful for finding information at a fixed value of t because you are not concerned with how the measurement is done for other parameter values. ▲

The next activity gives you an opportunity to practice with the velocity and unit tangent vectors and speed.

Activity 10.5.2

(a) Find $\vec{v}$, speed, and $\vec{T}$ as functions of t for the line parameterized as $\vec{r}(t) = \langle 3t-1, 2-2t, 5+t\rangle$. Write a few sentences about why your results make sense and why in for this particular curve $\vec{T}$ does not vary based on the parameter value.

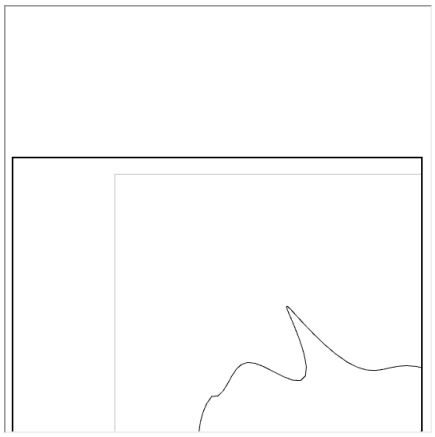
(b) Find $\vec{v}$, speed, and $\vec{T}$ for the curve with parameterization $\vec{r}(t) = \langle t, t^2, t^3\rangle$.

(c) Consider a curve for which we do not know the parameterization. However, we do know that at a point P, the tangent line to the curve through

P can be parameterized as $\langle 2-7t, 3t+1, -4t-1\rangle$. What are you able to say about $\vec{T}$ for this curve at the point P?

10.5.3 The Direction of Turning

Recalling our analogy in which we think of a parameterization as describing how someone drives along a racetrack, we say that $\vec{T}$ measures the direction that the car is heading at any given instant. If we use the curve plotted in Figure 10.5.7 as an example of the racetrack, then $\vec{T}(2) \approx \langle -0.8813, -0.4691, 0.0568\rangle$. The plot in Figure 10.5.7 shows how the direction of travel corresponds to the curve at the point $\vec{r}(2)$. What direction is the car turning at $t = 2$? You can see by the plot that the car will be be turning left, but that is a direction relative to the car. We wish to measure the direction that the car is turning at this instant in terms of the locations given by the parameterization in an xyz-coordinate system?



Standalone
In Context
Embed

Figure 10.5.7 A plot of a racetrack with $\vec{T}(2)$ plotted in red

We first consider $\vec{T}\,' = \frac{d\vec{T}}{dt}$, which measures how quickly $\vec{T}$ is changing in terms of the parameter. The vector $\vec{T}\,' = \frac{d\vec{T}}{dt}$ will measure the rate of change of the direction of travel. Since vectors have magnitude and direction, the derivative of a vector-valued function measures both the change in the magnitude of $\vec{T}$ *and* the change in the direction of $\vec{T}$. Since $\vec{T}$ will always have length 1 (when $\vec{T}$ exists), *all* of the change in $\vec{T}$ corresponds to a change in direction. Thus, the direction of $\vec{T}' = \frac{d\vec{T}}{dt}$ will be exactly what we are looking for: $\vec{T}\,'$ will point in the direction of turning. For now, we don't care about how fast the car is turning; we want to find the direction of the turning. This leads us to define the unit normal vector, denoted $\vec{N}$, as the unit vector in the direction of $\vec{T}\,'$.

Definition 10.5.8 The **unit normal vector** to a curve is denoted by $\vec{N}(t)$. This vector is computed as

$$\vec{N}(t) = \frac{\vec{T}\,'(t)}{\left\|\vec{T}\,'(t)\right\|}.$$

The unit normal vector measures the direction of turning along a parameterized curve. ◆

We are able to define the unit normal without the classic calculus approach because our vector tools allowed us to make sense of this measurement rather

than requiring a new tool.

Based on the definition of the unit normal vector, there are a few different characteristics a curve can have that would cause $\vec{N}$ to not exist. The next activity leads you to consider three different cases:

1. $\vec{N}$ will not exist if $\vec{T}$ does not exist
2. $\vec{N}$ will not exist if $\vec{T}\,'$ does not exist
3. $\vec{N}$ will not exist if $\vec{T}\,' = \vec{0}$

Activity 10.5.3

(a) Draw an example of a curve such that $\vec{T}$ does not exist at a point on your curve. Identify the point where $\vec{T}$ fails to exist. Explain both why the direction of travel does not exist at that point based on your plot and why the direction of turning will not exist at the same point.

(b) Draw an example of a curve such that $\vec{T}\,' = \vec{0}$ at a point on your curve. Explain why $\vec{T}\,' = \vec{0}$ based on your plot and explain why $\vec{N}$ does not exist at the same point.

Example 10.5.9 In this example, we will calculate the unit normal vector for a helical path given by $\vec{r}(t) = \langle \cos(t), \sin(t), \frac{t}{\pi} \rangle$. This is the same path as was used in Example 10.5.4 and was plotted in Figure 10.5.5.

From our work in Example 10.5.4, we calculated

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \left\langle -\sin(t), \cos(t), \frac{1}{\pi} \right\rangle$$

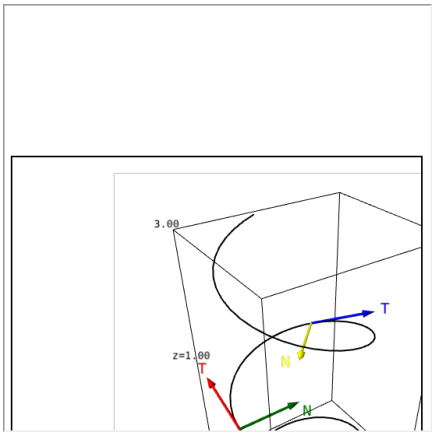
Using the definition of $\vec{N}$, we need to compute the derivative of $\vec{T}$ with respect to t:

$$\vec{T}\,'(t) = \frac{d\vec{T}}{dt} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle -\cos(t), -\sin(t), 0 \rangle$$

Note that in this example, the derivative of $\vec{T}$ is easily calculated for all time values but this will rarely be the case. Further, making a unit vector in the direction of $\vec{T}(t)' = \frac{d\vec{T}}{dt}$ is routine and gives us the following calculation:

$$\begin{aligned} \vec{N}(t) &= \frac{\vec{T}\,'(t)}{\left\|\vec{T}'(t)\right\|} \\ &= \frac{\frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle -\cos(t), -\sin(t), 0 \rangle}{\frac{1}{\sqrt{1+\frac{1}{\pi^2}}}} \\ &= \langle -\cos(t), -\sin(t), 0 \rangle \end{aligned}$$

We can calculate $\vec{N}(\frac{\pi}{2}) = \langle 0, -1, 0 \rangle$ and $\vec{N}(\frac{5\pi}{4}) = \langle \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0 \rangle$. If we add these vectors to the plot of the helix with the unit tangent vectors, as in Figure 10.5.6, we can see that the direction of turning (as measured by $\vec{N}$) is pointing toward the center of the helix. It should make sense the the direction you are turning to stay on the helix is toward the inside of the circular trace. Note that the unit tangent vector has a vertical component which corresponds to going "up" the helix but there is no vertical component to $\vec{N}$. This is because there is vertical motion, but there is no *turning* in the vertical direction. That is, the z-component of $\vec{T}$ is unchanging, which means that the vertical component of $\vec{T}\,'$ must be be zero.



Standalone
In Context
Embed

Figure 10.5.10 A plot of the helix with $\vec{T}$ and $\vec{N}$ plotted for a couple points

▲

Example 10.5.9 is the exceptional case where $\vec{N}$ is easily calculated directly from the definition and the calculations are fairly clean. In the next activity, we will look at a curve with a simple polynomial parameterization that will demonstrate how easily the calculation of $\vec{N}$ can go off the rails.

Activity 10.5.4 In general, calculating $\vec{N}$ as a function of the parameter t ends up being very difficult because there are multiple compositions of functions involved which means there is a nesting of chain rules involved in the definition. In this activity, we will examine how to go through the direct calculation of $\vec{N}$ for a curve parameterized by $\vec{r}(t) = \langle t, t^2, t^3 \rangle$.

(a) Calculate velocity $\vec{v}(t)$ and speed(t) for the parameterization $\vec{r}(t) = \langle t, t^2, t^3 \rangle$.

(b) Now calculate $\vec{T}(t)$ for the path given by $\vec{r}(t) = \langle t, t^2, t^3 \rangle$.

(c) Calculate the first component of $\vec{T}\,' = \frac{d\vec{T}}{dt}$.

Hint. Remember that the derivative with respect to t is computed componentwise.

(d) Calculate the second component of $\vec{T}\,'$.

(e) Calculate the third component of $\vec{T}\,'$.

(f) Put together your work for the three components of $\vec{T}\,'$.

(g) In the previous part, you likely decided that trying to simplify your formulas for $\frac{d\vec{T}}{dt}$ was more complicated than you had space for on your paper, and note that we still would need to calculate $\left\| \frac{d\vec{T}}{dt} \right\|$. Instead of doing more intricate algebraic computations, describe how you would calculate $\vec{N}$ if you had nice formulas for $\frac{d\vec{T}}{dt}$ and $\left\| \frac{d\vec{T}}{dt} \right\|$.

As talked about in Subsection 10.5.2, the unit tangent vector $\vec{T}$ measures the direction of *travel* along a parameterized curve. The unit normal vector $\vec{N}$ measures the direction that an object is *turning* in order to stay on the curve. In Figure 10.5.10, we observed that the unit tangent and unit normal vectors at each point along the curve were orthogonal to each other.

Question 10.5.11 Can $\vec{N}$ have any part parallel to $\vec{T}$ or must $\vec{T}$ and $\vec{N}$ be orthogonal? ▲

We first answer this question with a conceptual, geometric argument. By Definition 10.5.8, the unit normal vector is parallel to $\vec{T}\,' = \frac{d\vec{T}}{dt}$. The derivative of the unit tangent vector will measure the change in $\vec{T}$, and since $\vec{T}$ is a vector, this includes change in both the magnitude and the direction of $\vec{T}$. Since the length of $\vec{T}$ does not change, *all* of $\vec{T}\,'$ represents change in the direction of $\vec{T}$. This is an intuitive argument for why $\vec{T}$ and $\vec{N}$ are orthogonal. We next state this formally for future reference and present a formal, algebraic argument.

Key Idea 10.5.12 *The unit tangent vector $\vec{T}$ and the unit normal vector $\vec{N}$ are orthogonal so that*

$$\vec{T} \cdot \vec{N} = 0.$$

Proof. Here we give a rigorous, algebraic argument. However, this argument does not provide insight into the measurements or relationship between $\vec{T}$ and $\vec{N}$ that the paragraph before this Key Idea presents. For this algebraic argument, we consider at the derivative of $\left\|\vec{T}\right\|^2$. Because the length of $\vec{T}$ is constant, we must have that $\frac{d}{dt}\left(\left\|\vec{T}\right\|^2\right) = 0$. To simplify $\frac{d}{dt}\left(\left\|\vec{T}\right\|^2\right)$, we use properties of the dot product related to the magnitude, the derivative product rule for the dot product, and finally the commutative property of the dot product. Hence, we see

$$\begin{aligned}\frac{d}{dt}\left(\left\|\vec{T}\right\|^2\right) &= \frac{d}{dt}\left(\vec{T} \cdot \vec{T}\right) \\ &= \vec{T}' \cdot \vec{T} + \vec{T} \cdot \vec{T}' \\ &= 2\vec{T} \cdot \vec{T}'\end{aligned}$$

If $0 = \frac{d}{dt}\left(\left\|\vec{T}\right\|^2\right) = 2\vec{T} \cdot \vec{T}'$, then $\vec{T}$ and $\vec{T}'$ are orthogonal. Since $\vec{N}$ is parallel to $\vec{T}'$, we must also have that $\vec{T}$ and $\vec{N}$ are orthogonal. ■

10.5.4 The Binormal Vector

Since the title of this section is The TNB Frame and we have defined $\vec{T}$ and $\vec{N}$, you likely are expecting the definition of a vector called $\vec{B}$, which we do in this subsection. Since we saw in Key Idea 10.5.12 that the unit tangent and unit normal vectors are orthogonal, this final vector will be defined in a way that completes a right-handed coordinate system with $\vec{T}$ and $\vec{N}$.

Definition 10.5.13 The **binormal vector**, denoted $\vec{B}$, for a parameterized curve is defined as

$$\vec{B} = \vec{T} \times \vec{N}.$$

◆

This definition of $\vec{B}$ creates a three-dimensional right-handed coordinate system that is relative to the motion along the curve. Notice that while we had to divide by magnitude to ensure that $\vec{T}$ and $\vec{N}$ were unit vectors, the binormal vector is a unit vector because $\left\|\vec{T} \times \vec{N}\right\| = \left\|\vec{T}\right\| \left\|\vec{N}\right\| \sin(\theta)$ where θ is the angle between $\vec{T}$ and $\vec{N}$. Since $\vec{T}$ and $\vec{N}$ are unit vectors and orthogonal to each other, we see that $\left\|\vec{T}\right\| \left\|\vec{N}\right\| \sin(\theta) = 1$. The binormal vector $\vec{B}$ gives the axis of rotation for the motion along the curve. This measure of rotation

follows our right-handed measure of rotation. If you put the fingers of your right hand in the direction of travel ($\vec{T}$) and curl your fingers in the direction of turning ($\vec{N}$), then your thumb will correspond to the axis of rotation based on how you are moving along the curve.

The **TNB frame** is also called the **Frenet frame** and is frequently used in fields such as aviation and space travel because much of the important flight information is expressed in terms of relative measurements to the craft and its motion. The plane determined by the vectors $\vec{N}(t)$ and $\vec{B}(t)$ is called the *normal plane* to the curve at t and consists of all lines that are orthogonal to the tangent line to the curve at t. The plane determined by $\vec{T}(t)$ and $\vec{N}(t)$ is called the *osculating plane* and is the plane the comes closest to containing the curve for inputs near t. In the remaining sections of this chapter, we will return to these ideas of what is happening on the osculating plane.

Example 10.5.14 In this example, we complete the calculation of the TNB frame for the helix given by $\vec{r}(t) = \langle \cos(t), \sin(t)\frac{t}{\pi} \rangle$. From Example 10.5.4 and Example 10.5.9 we have the following:

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \left\langle -\sin(t), \cos(t), \frac{1}{\pi} \right\rangle$$

$$\vec{N} = \langle -\cos(t), -\sin(t), 0 \rangle$$

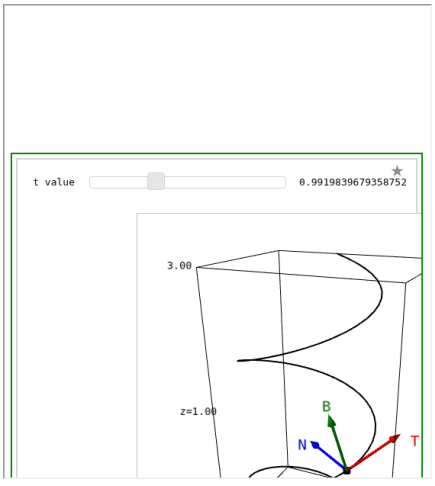
Using the definition for $\vec{B}$, we have

$$\vec{B} = \vec{T} \times \vec{N} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \left\langle \frac{1}{\pi}\sin(t), -\frac{1}{\pi}\cos(t), \sin(t)^2 + \cos(t)^2 \right\rangle$$

Factoring out $\frac{1}{\pi}$ from each component and simplifying the coefficient gives

$$\vec{B} = \frac{1}{\sqrt{\pi^2+1}} \langle \sin(t), -\cos(t), \pi \rangle.$$

Figure 10.5.15 shows the helix with the TNB frame depicted at a single point. You can adjust the location of the TNB frame by moving the slider.



Standalone
In Context
Embed

Figure 10.5.15 A plot of the helix with the TNB frame plotted

Note that as you change the t-value and move around the helix, the unit tangent points in the direction of travel, the unit normal points in the direction the object is turning (always to the left while going up the helix), and the

binormal vector is the axis of rotation for the $\vec{T}$ and $\vec{N}$ vectors. If you curl the fingers on your right hand from the direction of $\vec{T}$ toward the direction of $\vec{N}$, notice that your thumb will point in the direction of $\vec{B}$. This sets up the TNB frame as a right-handed coordinate system that changes relative to the motion of the object. ▲

10.5.5 The Driver or the Road: TNB Frame

To conclude this section, we consider whether the measurements of $\vec{T}$, $\vec{N}$, and $\vec{B}$ are properties of the driver or the road. A parameterized curve has a stated orientation (direction of travel). Remember that a measurement involving a parameterized curve is a property of the driver if two different parameterizations can have different measurements at the same location on the curve. It is important to remember that this does not mean that you look at the same parameter value on the curve, but rather the same *location* on the curve. A measurement on a parameterized curve is a property of the road if *every* parameterization of this curve must have the same measurement for a fixed location on the curve. These are the same ideas that were first discussed in Activity 10.4.3. In the analogy of a parameterization being a description for how a particular race car travels around a track as a function of time, a measurement is a property of the driver if that measurement can have different values for different drivers when measured at the same point on the racetrack. A measurement is a property of the road if that measurement *must* have the same value for *all* drivers when measured at the same point on the racetrack.

Activity 10.5.5

(a) *Direction of Travel.*

Each race car has a 1 meter arrow attached to the hood of the car which points straight ahead. This arrow will measure the three dimensional vector that is the direction of travel. Is the direction of travel a property of the driver or the road? Be sure to explain your answer.

(b) *Turn Signal.*

The turn signal is measured by checking to see if the steering wheel is turned left or right. We are not measuring how much the steering wheel is turned in either direction, just whether the wheel is turned left or right. If the steering wheel is not turned either way at the instant we measure , then we say there is no turn signal because the car isn't turning at that instant. Is the turn signal a property of the driver or the road? Be sure to explain your answer.

(c) *Direction of Turning.*

The direction of turning is measured by using a 1 meter arrow that sticks perpendicularly out of the left side of the car when the steering wheel is turned left and out of the right side of the car when the steering wheel is turned right. We are not measuring how much the wheel is turned in either direction, just whether the wheel is turned left or right. If the steering wheel is not turned either way at the instant we measure the direction of turning, then we say there is no direction of turning because the car isn't turning at that instant. Is the direction of turning as a vector in $\mathbb{R}^3$ a property of the driver or the road? Be sure to explain your answer.

(d) Is $\vec{B}$ a property of the driver or the road? Justify your ideas.

Hint. Use the definition of $\vec{B}$ and your answers to part (a) and part (c).

Summary

- The unit tangent vector, $\vec{T}(t) = \frac{\vec{v}}{\|\vec{v}\|}$, measures the direction of travel for an object along a parameterized curve.
- The unit normal vector, $\vec{N}(t) = \frac{\vec{T}\,'}{\|\vec{T}\,'\|}$, measures the direction that an object is turning in order to stay on a parameterized curve.
- The binormal vector, $\vec{B}(t) = \vec{T} \times \vec{N}$, is the direction of the axis for which $\vec{T}$ and $\vec{N}$ are rotating around.

10.5.6 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Please keep this in mind when working WeBWorK exercises.

Because we do not have a generally efficient way of calculating $\vec{N}$ from a parameterization or a curve, we will return to more exercises involving $\vec{N}$ after the discussion in Section 10.7. Additionally, there are several exercises in Section 10.6 that ask you to compute and interpret $\vec{N}$.

1. Find the unit tangent vector at the indicated point of the vector function

$$\mathbf{r}(t) = e^{19t}\cos t\,\mathbf{i} + e^{19t}\sin t\,\mathbf{j} + e^{19t}\,\mathbf{k}$$

$\mathbf{T}(\pi/2) = \langle$ _____, _____, _____ $\rangle$

2. Consider the vector function

$\mathbf{r}(t) = \langle t, t^2, t^3\rangle$

Compute

$\mathbf{r}'(t) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{T}(1) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{r}''(t) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{r}'(t) \times \mathbf{r}''(t) = \langle$ _____, _____, _____ $\rangle$

3. Consider the vector function

$\mathbf{r}(t) = \langle t, t^2, t^8\rangle$

Compute

$\mathbf{r}'(t) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{T}(1) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{r}''(t) = \langle$ _____, _____, _____ $\rangle$

$\mathbf{r}'(t) \times \mathbf{r}''(t) = \langle$ _____, _____, _____ $\rangle$

4. For the given position vectors $\mathbf{r}(t)$ compute the unit tangent vector $\mathbf{T}(t)$ for the given value of t .

A) Let $\mathbf{r}(t) = (\cos 2t, \sin 2t)$.
Then $\mathbf{T}(\frac{\pi}{4})$= (_____ , _____)
B) Let $\mathbf{r}(t) = (t^2, t^3)$.
Then $\mathbf{T}(3)$= (_____ , _____)
C) Let $\mathbf{r}(t) = e^{2t}\mathbf{i} + e^{-3t}\mathbf{j} + t\mathbf{k}$.
Then $\mathbf{T}(-5)$= _____ **i**+ _____**j**+ _____ **k** .

5. For the given position vectors $\mathbf{r}(t)$ compute the unit tangent vector $\mathbf{T}(t)$ for the given value of t .
A) Let $\mathbf{r}(t) = (\cos 2t, \sin 2t)$.
Then $\mathbf{T}(\frac{\pi}{4})$= (_____ , _____)
B) Let $\mathbf{r}(t) = (t^2, t^3)$.
Then $\mathbf{T}(1)$= (_____ , _____)
C) Let $\mathbf{r}(t) = e^{2t}\mathbf{i} + e^{-t}\mathbf{j} + t\mathbf{k}$.
Then $\mathbf{T}(0)$= _____ **i**+ _____**j**+ _____ **k** .

6. Calculate $\mathbf{r}'(t)$, $\mathbf{T}(t)$, and $\mathbf{T}(3)$ where

$$\mathbf{r}(t) = \langle 2t^4, 4t^4 \rangle.$$

$\mathbf{r}'(t) = \langle$ _____, _____$\rangle$.
$\mathbf{T}(t) = \langle$ _____, _____$\rangle$.
$\mathbf{T}(3) = \langle$ _____, _____$\rangle$.

7. Calculate $\mathbf{r}'(t)$ and $\mathbf{T}(t)$, where

$$\mathbf{r}(t) = \langle 4 + 4t, 1 - 4t, 2t - 6 \rangle.$$

$\mathbf{r}'(t) = \langle$ _____, _____, _____$\rangle$.
$\mathbf{T}(t) = \langle$ _____, _____, _____$\rangle$.

8. Find the unit tangent vector at the indicated point of the vector function

$$\mathbf{r}(t) = e^{14t}\cos t\,\mathbf{i} + e^{14t}\sin t\,\mathbf{j} + e^{14t}\,\mathbf{k}$$

$\mathbf{T}(\pi/2) = \langle$ _____, _____, _____ $\rangle$

9. Consider the moving particle whose position at time t in seconds is given by the vector-valued function $\vec{r}$ defined by $\vec{r}(t) = 5t\hat{\imath} + 4\sin(3t)\hat{\jmath} + 4\cos(3t)\hat{k}$. Use this function to answer each of the following questions.

(a) Find the unit tangent vector, $\vec{T}(t)$, to the space curve traced by $\vec{r}(t)$ at time t. Write one sentence that explains what $\vec{T}(t)$ tells us about the particle's motion.

(b) Determine the speed of the particle moving along the space curve with the given parameterization.

(c) Find the exact distance traveled by the particle on the time interval $[0, \pi/3]$.

(d) Find the average velocity of the particle on the time interval $[0, \pi/3]$.

(e) Determine the parameterization of the given curve with respect to arc length.

10.6 Curvature

Motivating Questions

- How can you measure how fast a path is turning (regardless of the parameterization)?
- What is the radius of curvature for a path at a given location?

10.6.1 Introduction

In Section 10.5, we defined $\vec{T}$ and $\vec{N}$ to measure the direction of travel and the direction of turning for an oriented curve in space. Both $\vec{T}$ and $\vec{N}$ are unit vectors that are used to capture *direction* aspects of motion. In the next couple of sections, we will look at a few different ways to measure the *magnitude* of various aspects of our motion and turning on a curve in space.

Given a parameterization of a curve in space, the speed will give us an idea about how fast an object moves *along the curve* at any given parameter value. We learned in the previous section that $\vec{N}$ tells us the *direction of turning*, but we have not yet developed a way to measure *how much turning* is happening at a given location. For example, consider the three curves shown in Figure 10.6.1. We would like to be able to measure how much turning is happening at the indicated point on each curve. Right now, you probably intuitively would say there is no turning in the first curve, a small amount in the second, and considerably more in the third. However, how can we assign a number to measure this amount of turning?

Figure 10.6.1 Three curves demonstrating variation in amount of turning

In this section, we focus on measuring how quickly *a curve* is turning. Keeping our running example in mind, we can think of this as measuring a property of *the road*. How quickly a curve is turning is a different measurement from how fast *an object* is turning as it moves along a path. That is a property of *the driver* to which we will return in Section 10.7.

Preview Activity 10.6.1 As Chief Engineer and CEO at *Steer Clear*, you have done great work in transforming information from your location tracking system (LTS) into information on the position, velocity, and distance traveled by your car. In order to make the "self-driving" part of the self-driving car, you will need to compare information from the LTS (which describes how your car is moving) to GPS location data for the network of roads (which describes the paths your car should be taking). Since everyone on the road drives differently, you realize that you need to measure how quickly a stretch of road is turning in different places, which leads you to contact a friend who is a civil engineer for help understanding highway construction specifications.

(a) Figure 10.6.2 shows a map of a section of road on your testing route for

your self-driving car with points labeled P_0 and P_1 which are 10 meters apart (along the road). Draw a vector in the direction of travel at both P_0 and P_1. Use these vectors to describe how the direction of travel is changing along the path from P_0 to P_1.

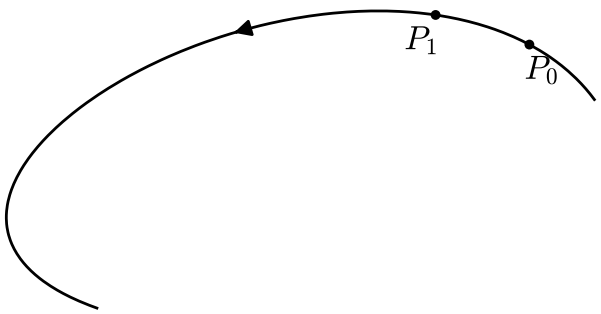


Figure 10.6.2 A plot of the section of road with points P_0 and P_1 separated by 10 meters

(b) Figure 10.6.3 shows a map of a section of road on your testing route for your self-driving car with points labeled Q_0 and Q_1 which are 10 meters apart (along the road). Draw a vector in the direction of travel at both Q_0 and Q_1. Use these vectors to describe how the direction of travel is changing along the path from Q_0 to Q_1.

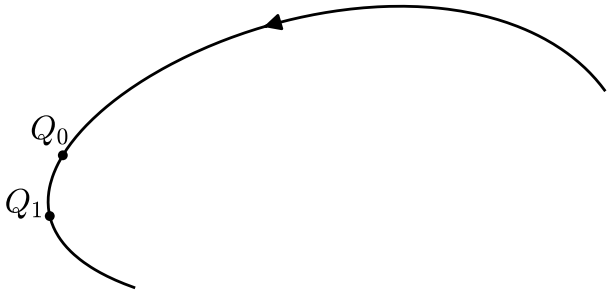


Figure 10.6.3 A plot of the section of road with points Q_0 and Q_1 separated by 10 meters

(c) Is the direction of travel changing faster over the path from P_0 to P_1 or over the path from Q_0 to Q_1. Write a few sentences to explain your reasoning and connect to your arguments and plots for the first two tasks.

The Preview Activity showed how looking at the change in $\vec{T}$, the direction of travel, for small steps in terms of arc length along our curve will allow us to measure the rate at which a curve is turning. This is different than measuring how quickly an object traveling along the path needs to turn in order to stay on the path. The arguments in the Preview Activity depend only on intrinsic geometric properties of the path (road) rather than how an object (car) moves along that path.

10.6.2 Calculating Curvature

Hopefully it is not surprising that the unit speed parameterization of a curve $\vec{r}_1(s)$ is unique. One way to think about this is that if two drivers follow the same path, each at a speed of exactly 1 meter per second, then for any point on the path, the two drivers both reach that point in the same amount of time. For instance, it will take both drivers exactly 15 seconds to reach the point

that is 15 meters down the path. (Remember the most useful property of a unit speed parameterization is that the parameter value for $\vec{r}_1$ is equal to the numeric value of the distance traveled along the curve, albeit with different units.) Conceptually, you can see how using the unit speed parameterization allows us to measure geometric properties of the curve independent of how an object moves along the path.

We have already explored how $\vec{N}$ points in the direction of turning. Now, however, we are looking for a *scalar* quantity that measures *how quickly* a curve turns. To measure how quickly a path turns, we examine how quickly the direction of travel changes in terms of the unit speed parameterization. We apply the classic calculus approach to this measurement in order to obtain the exact value of this measurement.

For step 1 of the CCA, we want approximate how quickly the curve turns by looking at the difference quotient for two nearby direction of travel vectors. Suppose $\vec{r}_1(s)$ is the unit speed parameterization of a curve C and consider a section of C from $\vec{r}_1(s_0)$ to $\vec{r}_1(s_0 + \Delta s)$. We first consider how different the direction of travel is over this curve of length Δs. (Remember, the change in parameter value here corresponds to the arc length along the curve because $\vec{r}_1$ is a unit speed parameterization.) Figure 10.6.4 shows a plot of a sample curve and two points separated by an arc length of Δs.

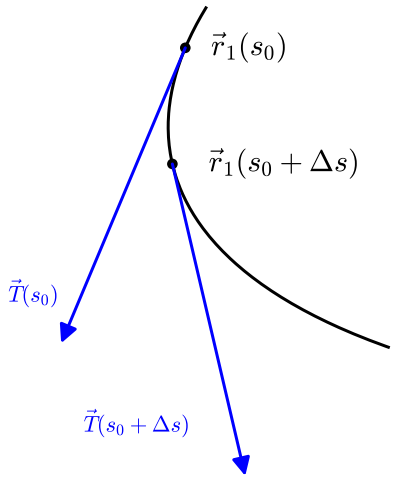


Figure 10.6.4 A plot of a curve with $\vec{T}$ shown for two points

We want to know how fast the direction of travel is changing at the point $\vec{r}_1(s_0)$. To do this, we approximate the change in the direction of travel on an interval around this point and take the limit as the size of the interval goes to zero. We compute the difference between the directions of travel at the beginning and end of our interval by putting both $\vec{T}(s_0 + \Delta s)$ and $\vec{T}(s_0)$ starting at the same location. This allows us to use vector subtraction to find $\vec{T}(s_0 + \Delta s) - \vec{T}(s_0)$, which is the change in the direction of travel over the given interval. This vector subtraction is shown using the triangle of vectors in Figure 10.6.5.

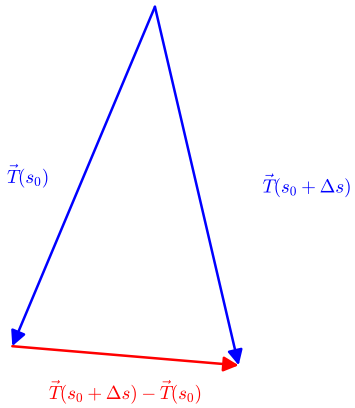


Figure 10.6.5 A plot of a curve with $\vec{T}$ shown for two points

So for a step of size Δs in arc length along our curve, we have a change in direction of $\vec{T}(s_0+\Delta s)-\vec{T}(s_0)$ in the direction of travel. In order to approximate the rate of change for the direction of travel, we will need to look at the ratio

$$\frac{\vec{T}(s_0+\Delta s)-\vec{T}(s_0)}{\Delta s}.$$

Remember that the unit normal vector $\vec{N}$ points in the direction of turning. We want to measure as a *scalar* how quickly the direction of turning is changing along our path. To do this, we compute the magnitude of this difference quotient as an approximation of rate of change for the direction of travel on the interval $[s_0, s_0+\Delta s]$. This gives

$$\left\| \frac{\vec{T}(s_0+\Delta s)-\vec{T}(s_0)}{\Delta s} \right\|,$$

which completes the first step of the classic calculus approach measuring how quickly (as a scalar) the path changes direction.

The CCA step 2 is taken care of easily because we made our earlier approximation in terms of a step in our parameter/arc length along the curve. We now consider smaller step sizes in the parameter, which corresponds to smaller changes in arc length along the curve because we have used the unit speed parameterization. In doing so, we obtain a better approximation of how fast the direction of travel is changing along the path at the point $\vec{r}_1(s_0)$.

Taking the limit as $\Delta s \to 0$, completes step three of the classic calculus approach to measure how quickly a curve is turning at the point $P=\vec{r}_1(s_0)$. Calculating, we have

$$\lim_{\Delta s\to 0}\left\| \frac{\vec{T}(s_0+\Delta s)-\vec{T}(s_0)}{\Delta s} \right\| = \left\| \frac{d\vec{T}}{ds} \right\|.$$

This leads us to the following definition:

Definition 10.6.6 If C is a smooth space curve and s is the parameter for the unit speed parameterization of C, then the **curvature**, κ, of C is

$$\kappa = \kappa(s) = \left\| \frac{d\vec{T}}{ds} \right\|$$

◆

Remember that the parameter s for the unit speed parameterization C is the arc length traveled along the curve. Curvature is denoted by κ, the lowercase Greek letter "kappa".

Activity 10.6.2 In this activity we will use our previous work in finding unit speed parameterizations of lines and circles to make sense of the definition of curvature.

(a) Recall that in Example 10.4.8 we found that the unit speed parameterization of a line through the points $(1, 2, 3)$ and $(2, 0, 5)$ to be $\vec{r}_1(s) = \langle 1, 2, 3\rangle + \frac{s}{3}\langle 1, -2, 2\rangle$. Calculate $\vec{T}(s)$, $\frac{d\vec{T}}{ds}$, and $\left\|\frac{d\vec{T}}{ds}\right\|$ for the unit speed parameterization of this line. Note that because we are using the unit speed parameterization, our parameter t is the same as the arc length traveled s. In other words, $t = s$ for this parameterization.

(b) Write a few sentences to explain why your result for the calculation of the curvature $\kappa = \left\|\frac{d\vec{T}}{ds}\right\|$ for this line makes sense. Exercise 15 asks you to do the calculations to confirm this property of an arbitrary line.

(c) Recall from Activity 10.4.2 have a parameterization of a circle in 2-space of radius R centered at the origin as $\langle R\cos(t), R\sin(t)\rangle$. A quick calculation will verify that this parameterization has constant speed and that speed is R.

Modifying this parameterization by multiplying the parameter by $\frac{1}{\text{speed}} = \frac{1}{R}$ gives the following parameterization:

$$\vec{r}(s) = \left\langle R\cos\left(\frac{s}{R}\right), R\sin\left(\frac{s}{R}\right)\right\rangle$$

Compute the speed of this parameterization to verify that this has unit speed. This will also show that $\vec{r}\,'(s) = \vec{v}(s) = \vec{T}(s)$. (Recall that this fact is not true in general!)

(d) Calculate $\vec{T}(s)$, $\frac{d\vec{T}}{ds}$, and $\left\|\frac{d\vec{T}}{ds}\right\|$ for the unit speed parameterization of a circle of radius R centered at the origin.

(e) Write a few sentences to explain why your result for the calculation of $\kappa = \left\|\frac{d\vec{T}}{ds}\right\|$ of circle will be constant. You should also address why circles with larger radii will have have smaller curvature.

Notice that a unit speed parameterization will have the property that $\vec{r}_1\,'(s) = \vec{v}(s) = \vec{T}(s)$. While this fact is not true in general, it does make the calculation and interpretation of $\vec{T}$ as simple as possible. It is also helpful to observe that a unit speed parameterization has the property that $\vec{T}'(s) = \vec{a}(s)$, the acceleration vector of the curve. Thus, when we have a unit speed parameterization for a path, we can think of the curvature of the path at a point as being the magnitude of the acceleration vector.

The definition of curvature relies on our ability to parameterize curves with unit speed. Since finding unit speed parametrization is frequently difficult to compute directly, the following theorem gives a way of computing the curvature of a curve in terms of *any* parametrization $\vec{r}(t)$.

Theorem 10.6.7 *If $\vec{r}$ is a parameterization of a curve C, and if $\vec{r}\,'(t) \neq \vec{0}$ and*

$\vec{r}''(t)$ exists, then the curvature κ of C can be calculated by

$$\kappa(t) = \frac{\left\|\vec{T}'(t)\right\|}{\|\vec{r}'(t)\|}.$$

Further, we can efficiently calculate curvature in terms of the velocity and acceleration functions as

$$\kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3}. \tag{10.6.1}$$

Proof. We will give the outline of a proof of the first result here in order to emphasize that this result comes from a chain rule argument related to the change of parameters going to a unit speed parameterization.

Let $\vec{r}(t)$ be some parameterization of the curve C. To calculate the curvature, $\vec{T}$ must have an input corresponding to the parameter of the unit speed parameterization. This means that the input of $\vec{T}$ must measure the arc length of C to the point of interest. We will use a function $s(t)$. The input to this function is the parameter of $\vec{r}(t)$, and the output of $s(t)$ is the arc length on C to the point $\vec{r}(t)$. Note that $s(t)$ is computed using Theorem 10.4.5 and that the derivative $\frac{ds}{dt}$ will be $\|\vec{r}'(t)\|$.

We can write $\vec{T}$ as a function of the parameter t by the composition $\vec{T}(t) = \vec{T}(s(t))$. Taking the derivative of this composition with respect to t uses a chain rule argument:

$$\frac{d\vec{T}}{dt} = \frac{d}{dt}\left(\vec{T}(s(t))\right) = \frac{d\vec{T}}{ds}\frac{ds}{dt}.$$

Note here that $\frac{d\vec{T}}{dt} = \vec{T}'$ is the same derivative of $\vec{T}$ we have seen a few times.

Because $\frac{ds}{dt}$ is a scalar-valued function, we can divide both sides of the equation by it to get

$$\frac{\vec{T}'}{\frac{ds}{dt}} = \frac{d\vec{T}}{ds}.$$

Using the idea above from Theorem 10.4.5, we can substitute $\frac{ds}{dt} = \|\vec{r}'(t)\|$. Computing the magnitude of each side yields

$$\left\|\frac{\vec{T}'}{\|\vec{r}'(t)\|}\right\| = \left\|\frac{d\vec{T}}{ds}\right\| = \kappa. \tag{10.6.2}$$

There is also a nice conceptual interpretation of the equation above. Specifically, if we want to convert a measurement such as curvature or the derivative of $\vec{T}$ between the unit speed parameterization and any other parameterization, we adjust by the speed. Equation (10.6.2) shows that the adjustment between the derivative of $\vec{T}$ with respect to the unit speed parameter and another parameter is to divide by the speed, which may vary depending on time (or position).

We omit the proof of (10.6.1) because the argument involves technical algebraic ideas that do not offer geometric insight into what curvature measures. Exercise 13 guides you through this calculation. ■

Soon we will turn to an example of calculating curvature and then some activities for you to practice. First, however, we consider how we can use (10.6.1) to make sense of when curvature is undefined (or does not make sense). From an algebraic perspective, the curvature cannot be defined when the speed is zero. This should make sense because if you are not moving along the

curve, as indicated by having speed equal to zero, then it does not make sense to measure how quickly the curve is turning. In Section 10.7, we consider why $\|\vec{v}(t) \times \vec{a}(t)\|$ being zero means that there will be zero curvature. This is different than the curvature being undefined!

Example 10.6.8 In this example, we will calculate the curvature of the parabola given by $y = x^2$ for $-2 < x < 2$. We would like to use (10.6.1) to calculate the curvature, so we must parameterize this curve in three dimensions with the z-component constantly zero. This is because the cross product is only defined for three-dimensional vectors. The vector-valued function $\vec{r}(t) = \langle t, t^2, 0\rangle$ with $-2 < t < 2$ will parameterize the curve given by $y = x^2$ for $-2 < x < 2$.

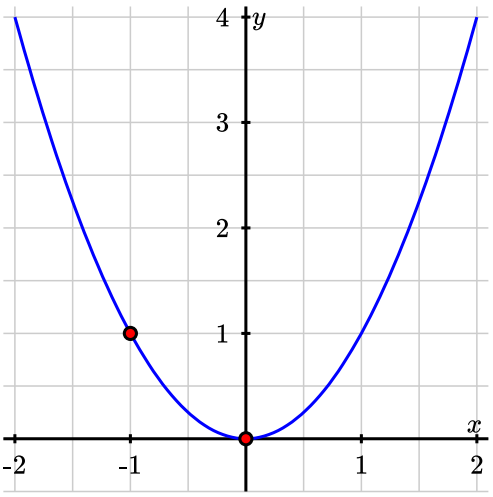


Figure 10.6.9 A plot of $y = x^2$ with points $(0, 0)$ and $(-1, 1)$

Using this parameterization, we compute the following:

$$\begin{aligned}
\vec{v}(t) = \vec{r}\,'(t) &= \langle 1, 2t, 0\rangle \\
\vec{a}(t) = \vec{r}\,''(t) &= \langle 0, 2, 0\rangle \\
\vec{v}(t) \times \vec{a}(t) &= \langle 0, 0, 2\rangle \\
\|\vec{v}(t) \times \vec{a}(t)\| &= 2 \\
\|\vec{v}(t)\| &= \sqrt{1 + 4t^2}
\end{aligned}$$

Using (10.6.1), the curvature is

$$\kappa(t) = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3} = \frac{2}{(1 + 4t^2)^{3/2}}.$$

Comparing the points $(0, 0)$ and $(-1, 1)$ on Figure 10.6.9, you can see that the path bends more at $(0, 0)$ than at $(-1, 1)$, which corresponds to the parabola having larger curvature at $(0, 0)$ than at $(-1, 1)$. Using the formula we obtained above, we can see that at $(0, 0)$, when $t = 0$, the curvature is $\kappa(0) = 2$ while at $(-1, 1)$, when $t = -1$, the curvature is $\kappa(-1) = \frac{2}{(5)^{3/2}} \approx 0.179$. This calculation of curvature matches with our visual description from above. ▲

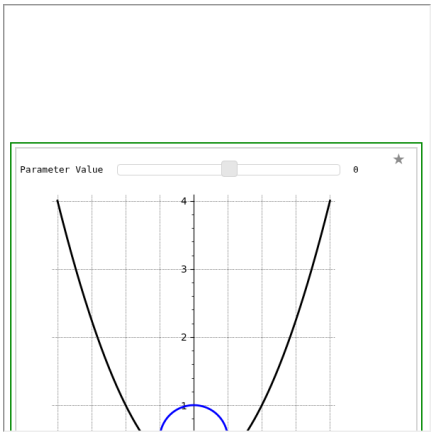
Suppose at the instant $t = x = 0$ you were driving along the path given by $y = x^2$ at unit speed and your steering wheel locked up preventing you from changing how you were turning. Because the steering wheel is held fixed, you will no longer stay on the path given by $y = x^2$. Instead, your car will drive

in a circle, and we can figure out what the radius of that circle would be using curvature.

In Activity 10.6.2, we saw that the curvature of a circle of radius R is $\frac{1}{R}$. Hopefully is not too surprising that the radius of the circle in which the car would turn if the curvature was held at that location would match this reciprocal relationship. In fact, the radius of this circle will be $\frac{1}{\kappa(0)} = \frac{1}{2}$. In general, the **radius of curvature** is defined to be $\frac{1}{\kappa}$ and the circle that is tangent to the curve at the point where the curvature is measured with radius $\frac{1}{\kappa}$ is called the **osculating circle**. The osculating circle can be described as the circle that best approximates the curve at the point since the osculating circle goes through the point and has the same curvature as the original path.

In Figure 10.6.10, you can see a plot of the path $y = x^2$ with the osculating circle drawn at a particular point. Adjust the slider to change the parameter value (and thus point) at which the osculating circle is drawn. Notice how the radius of the circle becomes larger as the point moves farther away from the vertex in either direction.

Remember that curvature and the radius of curvature have a *reciprocal relationship*: large curvature means small radius and small curvature means large radius.



Standalone
In Context
Embed

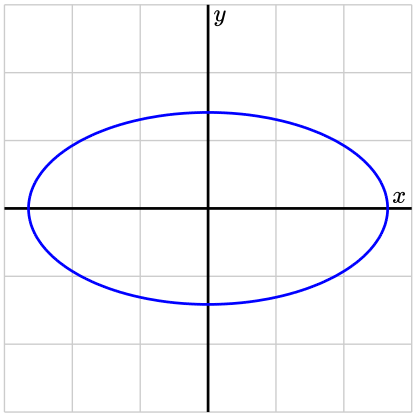
Figure 10.6.10 A plot of $y = x^2$ with the osculating circle drawn tangent to the curve

Activity 10.6.3 In this activity, we calculate the curvature of an ellipse and of a helix and use graphs to help with our understanding.

(a) Consider the ellipse shown below. We have omitted the scale on the axes because we want to think of this as the general ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, which has parameterization

$$\vec{r}(t) = \langle a\cos(t), b\sin(t)\rangle.$$

(The ellipse pictured below has $a > b$.)



On the graph, identify the points at which you believe the curvature to be the largest and the points at which you believe the curvature is smallest.

(b) For the given parameterization of the ellipse, find $\vec{v}(t)$ and $\vec{a}(t)$. Also identify the parameter values that correspond to the points of the ellipse with the largest and smallest x or y coordinates.

(c) Use your calculations for $\vec{v}(t)$ and $\vec{a}(t)$ from part (b) in Equation (10.6.1) to find a formula for the curvature of the ellipse with the given parameterization. Use your formula to find the curvature of the ellipse at the parameter values you identified in part (b) as corresponding to the extreme points of the ellipse.

Hint. Remember that to compute the cross product here, you will need to consider $\vec{v}$ and $\vec{a}$ as three-dimensional vectors.

(d) While it is possible to use single-variable calculus techniques to find the parameter values that maximize and minimize $\kappa(t)$ on the interval $[0, 2\pi]$, the algebra involved masks much of the insight. Use graphing technology such as Desmos to plot the curvature function you found above, preferably with sliders that allow you to adjust the values of a and b. Write a couple of sentences about what you observe about the parameter values that maximize and minimize curvature and how they depend on a and b. Then use the general form of the curvature function from above to find the curvature at these parameter values, simplifying as much as possible.

(e) The standard helix has parameterization $\vec{r}_2(t) = \cos(t)\hat{\imath} + \sin(t)\hat{\jmath} + t\hat{k}$. Find the curvature of the helix using this parameterization.

(f) Write a few sentences to describe why the curvature of the helix given above is constant. You may want to include a plot of the curve.

Before concluding this section, it may be helpful to think about what the units are on curvature. To do this, we again consider the formula that is in terms of velocity and acceleration:

$$\kappa = \frac{\|\vec{v} \times \vec{a}\|}{\|\vec{v}\|^3}$$

If we assume that our parameterization has distance with units of meters and input is thought of as time measured in seconds, then the units on $\vec{v}$ are meters per second and the units on $\vec{a}$ are meters per second-squared. This means the numerator of κ has units meters squared per second-cubed and the denominator is measured in meters-cubed per second-cubed. Simplifying, we

find that curvature has units one over meters.

$$\frac{1}{\text{meters}} = \frac{\frac{\text{meters}^2}{\text{seconds}^3}}{\frac{\text{meters}^3}{\text{seconds}^3}}$$

This aligns with the idea of the radius of curvature being $1/\kappa$, as it results in the radius of curvature having meters as its unit. That makes sense for something that we call a radius!

Activity 10.6.4 In this activity we will look at the measurement of curvature and determine if curvature is a property of the driver or a property of the road. This is a continuation of Activity 10.4.3. As a reminder, a measurement is a property of the driver if the value(s) of that measurement *can* be different for different drivers (when measured at the same location on the racetrack). A measurement is a property of the road when different drivers *must* have the same value(s) (when measured at the same location on the racetrack).

(a) Is curvature a property of the driver or the road? Write a few sentences about your reasoning and be sure to address why every driver may or may not have the same measurement at the same location on the track.

(b) What is the curvature on a straightaway (the race track is a line segment)? What would the radius of curvature be for a straightaway? Write a couple of sentences to justify your answers.

Summary

- We define the curvature κ of a curve in 2- or 3-space to be the rate of change of the magnitude of the unit tangent vector with respect to arc length, or
$$\kappa = \left\| \frac{d\vec{T}}{ds} \right\|$$
- Curvature can be efficiently calculated by the following formula for curves in space:
$$\kappa = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3}$$
- The radius of curvature is $\frac{1}{\kappa}$ and describes the radius of a circle corresponding to how quickly a path is turning.

10.6.3 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Additionally, some WeBWork problems denote the magnitude of vectors using single bars, like $|\vec{v}|$. Please keep this in mind when working WeBWorK exercises.

1. Find the curvature of $y = \sin(3x)$ at $x = \frac{\pi}{4}$.

2. Find the curvature $\kappa(t)$ of the curve $\mathbf{r}(t) = (4\sin t)\,\mathbf{i} + (4\sin t)\,\mathbf{j} + (-1\cos t)\,\mathbf{k}$

3. A factory has a machine which bends wire at a rate of 1 unit(s) of curvature per second. How long does it take to bend a straight wire into a circle of radius 1?

_____ seconds

4. For the curve given by $r(t) = \langle -7\sin(t), 2t, 7\cos(t)\rangle$,

Find the unit tangent

$T(t) = \langle$ _____, _____, _____ $\rangle$

Find the unit normal

$N(t) = \langle$ _____, _____, _____ $\rangle$

Find the curvature

$\kappa(t) =$ _____

5. For the curve given by

$$r(t) = \langle \sin(t) - t\cos(t), \cos(t) + t\sin(t), 5t^2 - 9\rangle$$

Find the unit tangent

$T(t) = \langle$ _____, _____, _____ $\rangle$

Find the unit normal

$N(t) = \langle$ _____, _____, _____ $\rangle$

Find the curvature

$\kappa(t) =$ _____

6. For the curve given by $r(t) = \langle -3t, -5t, 1 + 9t^2\rangle$,

Find the derivative

$r'(t) = \langle$ _____, _____, _____ $\rangle$

Find the second derivative

$r''(t) = \langle$ _____, _____, _____ $\rangle$

Find the curvature at $t = 1$

$\kappa(1) =$ _____

7. For the curve given by $r(t) = \langle -1t, e^{3t}, e^{-3t}\rangle$,

Find the derivative

$r'(t) = \langle$ _____, _____, _____ $\rangle$

Find the second derivative

$r''(t) = \langle$ _____, _____, _____ $\rangle$

Find the curvature at $t = 0$

$\kappa(0) =$ _____

8. Consider the helix $\mathbf{r}(t) = (\cos(1t), \sin(1t), -2t)$. Compute, at $t = \frac{\pi}{6}$:

A. The unit tangent vector $\mathbf{T} =$ (_____, _____, _____)

B. The unit normal vector $\mathbf{N} =$ (_____, _____, _____)

C. The unit binormal vector $\mathbf{B} =$ (_____, _____, _____)

D. The curvature $\kappa =$ _____

9. Find the equation of the osculating circle at the local minimum of

$$f(x) = 4x^3 + 8x^2 + \frac{13}{4}x + 8$$

Equation: _____

10. Calculate $\kappa(t)$ when

$$\mathbf{r}(t) = \langle 5t^{-1}, 1, 6t\rangle$$

$\kappa(t) =$ _____.

11. Let $y = f(x)$ define a curve in the plane. We can consider this curve as a curve in three-space with z-coordinate 0.

(a) Find a parameterization of the form $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ of the curve $y = f(x)$ in three-space.

(b) Use the formula

$$\kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3}$$

to show that

$$\kappa = \frac{|f''(x)|}{[1 + (f'(x))^2]^{3/2}}.$$

12. Consider the single variable function defined by $y = 4x^2 - x^3$.

(a) Find a parameterization of the form $\vec{r}(t) = \langle x(t), y(t)\rangle$ that traces the curve $y = 4x^2 - x^3$ on the interval from $x = -3$ to $x = 3$.

(b) Write a definite integral which, if evaluated, gives the exact length of the given curve from $x = -3$ to $x = 3$. Why is the integral difficult to evaluate exactly?

(c) Determine the curvature, $\kappa(t)$, of the parameterized curve. (Exercise 10.6.3.11 might be useful here.)

(d) Use appropriate technology to approximate the absolute maximum and minimum of $\kappa(t)$ on the parameter interval for your parameterization. Compare your results with the graph of $y = 4x^2 - x^3$. How do the absolute maximum and absolute minimum of $\kappa(t)$ align with the original curve?

13. In this exercise we verify the curvature formula

$$\kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3}.$$

(a) Explain why

$$\|\vec{r}'(t)\| = \frac{ds}{dt}.$$

(b) Use the fact that $\vec{T}(t) = \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$ and $\|\vec{r}'(t)\| = \frac{ds}{dt}$ to explain why

$$\vec{r}'(t) = \frac{ds}{dt}\vec{T}(t).$$

(c) The Product Rule shows that

$$\vec{r}''(t) = \frac{d^2s}{dt^2}\vec{T}(t) + \frac{ds}{dt}\vec{T}'(t).$$

Explain why

$$\vec{r}'(t) \times \vec{r}''(t) = \left(\frac{ds}{dt}\right)^2 (\vec{T}(t) \times \vec{T}'(t)).$$

(d) In Key Idea on $\vec{T}$ and $\vec{N}$ we showed that $\left\|\vec{T}(t)\right\| = 1$ implies that $\vec{T}(t)$ is orthogonal to $\vec{T}'(t)$ for every value of t. Explain what this

tells us about $\left\|\vec{T}(t) \times \vec{T}'(t)\right\|$ and conclude that

$$\|\vec{r}'(t) \times \vec{r}''(t)\| = \left(\frac{ds}{dt}\right)^2 \left\|\vec{T}'(t)\right\|.$$

(e) Finally, use the fact that $\kappa = \frac{\|\vec{T}'(t)\|}{\|\vec{r}'(t)\|}$ to verify that

$$\kappa = \frac{\|vr'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3}.$$

14. In this exercise we explore how to find the osculating circle for a given curve. As an example, we will use the curve defined by $f(x) = x^2$. Recall that this curve can be parameterized by $x(t) = t$ and $y(t) = t^2$.

(a) Use Definition 10.5.3 to find $\vec{T}(t)$ for our function f.

(b) To find the center of the osculating circle, we will want to find a vector that points from a point on the curve to the center of the circle. Such a vector will be orthogonal to the tangent vector at that point. Recall that $\vec{T}(s) = \langle \cos(\phi(s)), \sin(\phi(s)) \rangle$, where ϕ is the angle the tangent vector to the curve makes with a horizontal vector. Use this fact to show that

$$\vec{T} \cdot \frac{dT}{ds} = 0.$$

Explain why this tells us that $\frac{dT}{ds}$ is orthogonal to $\vec{T}$. Let $\vec{N}$ be the unit vector in the direction of $\frac{dT}{ds}$. The vector $\vec{N}$ is called the *principal unit normal vector* and points in the direction toward which the curve is turning. The vector $\vec{N}$ also points toward the center of the osculating circle.

(c) Find $\vec{T}$ at the point $(1, 1)$ on the graph of f. Then find $\vec{N}$ at this same point. How do you know you have the correct direction for $\vec{N}$?

(d) Let P be a point on the curve. Recall that $\rho = \frac{1}{\kappa}$ at point P is the radius of the osculating circle at point P. We call ρ the *radius of curvature* at point P. Let C be the center of the osculating circle to the curve at point P, and let O be the origin. Let γ be the vector $\overrightarrow{OC}$. See Figure 10.6.11 for an illustration using an arbitrary function f.

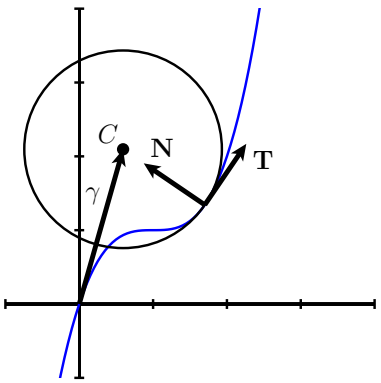


Figure 10.6.11 An osculating circle.

Which vector, in terms of ρ and $\vec{N}$ points from the point P to the point C? Use this vector to explain why

$$\gamma = \vec{r} + \rho\vec{N},$$

where $\vec{r} = \overrightarrow{OP}$.

(e) Finally, use the previous work to find the center of the osculating circle for f at the point $(1, 1)$. Draw pictures of the curve and the osculating circle to verify your work.

15. In Activity 10.6.2, you verified for a *specific* line that the curvature is 0. The generalization of this is that the unit speed parameterization of a line through (x_0, y_0, z_0) with direction vector $\langle a, b, c\rangle$ can be given by

$$\begin{aligned}\vec{r}(s) &= \langle x(s), y(s), z(s)\rangle \\ &= \left\langle x_0 + \frac{a}{\sqrt{a^2+b^2+c^2}}s, y_0 + \frac{b}{\sqrt{a^2+b^2}}s, z_0 + \frac{c}{\sqrt{a^2+b^2+c^2}}s \right\rangle.\end{aligned}$$

(a) Compute the speed of this parameterization to verify that this has unit speed. This will also show that $\vec{r}\,'(s) = \vec{v}(s) = \vec{T}(s)$ (which is not true in general).

(b) Calculate $\vec{T}(s)$, $\frac{d\vec{T}}{ds}$, and $\left\|\frac{d\vec{T}}{ds}\right\|$ for this unit speed parameterization of a line.

(c) Write a few sentences to explain why your result for the calculation of the curvature of a line makes sense in general.

16.

(a) If $\vec{u} \times \vec{w} = \vec{0}$, what does this say about the direction of $\vec{u}$ and $\vec{w}$?

(b) Remember that the vector $\vec{v}$ has magnitude and direction, so the derivative $\frac{d\vec{v}}{dt} = \vec{a}$ will measure the change in both the magnitude and direction of of $\vec{v}$. Write a few sentences to describe how if $\vec{v}$ and $\vec{a} = \frac{d\vec{v}}{dt}$ are in parallel directions, then the unit normal vector will be undefined at that point.

10.7 Splitting the Acceleration Vector

Motivating Questions

- How can we split the acceleration vector into parts in the direction of travel and the direction of turning?

10.7.1 Introduction

Position is the central measurement of the study of motion, which is known as kinematics. Velocity and acceleration are the first and second derivatives of position, respectively, and are frequently important properties in the study of physics and mechanics. To better understand velocity as a vector-valued function of time, we examined how the magnitude and direction of position change separately. We separated velocity into its critical parts: speed (magnitude of velocity) and direction of travel (unit tangent vector). Speed is a scalar measurement for how fast the position of an object is changing as a function of time, while $\vec{T}$ is a unit vector that measures the direction in which position is changing.

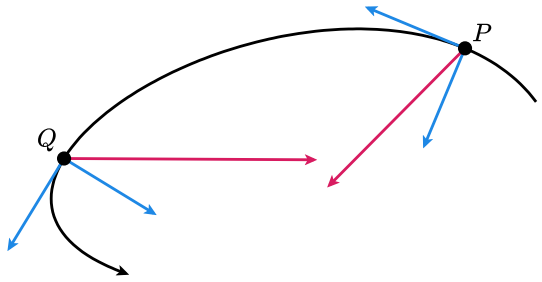
As a vector quantity, acceleration is related to many important physical laws and descriptions. For example, Newton's Second Law of Motion relates the forces acting on an object to its acceleration through the formula $\vec{F}_{net} = m\vec{a}$. This is often summarized as the sum of the forces acting on an object is the mass times the acceleration of the object. This description relates a combination of

forces on a object to its position and velocity through integration of $\vec{a}$. For an example of these calculations, refer back to Example 10.3.7.

To understand acceleration better as a vector quantity, we will examine acceleration as a change in the magnitude and direction of the velocity vector. In this section, we will use our knowledge of calculus and vector measurements to split the acceleration vector into parts related to the changes in the magnitude and direction of the velocity vector. We will also consider how these measurements relate to physical contexts.

Preview Activity 10.7.1 All your good work as CEO and lead engineer at *Steer Clear* has started to pay off, literally. When you showed your work on using the location tracking system (LTS) to develop navigation and telemetry tools to an investment group (your grandparents), they were impressed and decided to give you money to further develop your self-driving car. You decided to use this infusion of cash to buy a gyroscopic accelerometer which will measure acceleration as a vector with magnitude and direction. In order to use your new instrument for information on the "driving" part of your self-driving car, you need to understand how your accelerometer readings will relate to the operations needed to drive a car.

(a) The image below shows the path you drove on a test drive along with points P and Q and the direction of travel. At the points P and Q, we have drawn three vectors. The acceleration vector $\vec{a}$ provided by your new gyroscopic accelerometer is shown in magenta. The two blue vectors are the unit tangent vector $\vec{T}$ and unit normal vector $\vec{N}$ at that point. Label each of the blue vectors as being $\vec{T}$ or $\vec{N}$.



(b) Draw the projection of $\vec{a}$ on $\vec{T}$ and the projection of $\vec{a}$ on $\vec{N}$ at point P and at point Q in the graph above.

(c) Is $\vec{a} \cdot \vec{T}$ positive, negative, or zero at P? What about at Q? Write a sentence or two to justify your answers.

(d) Is $\vec{a} \cdot \vec{N}$ is positive, negative, or zero at P? What about at Q? Write a sentence or two to justify your answers.

The preview activity shows how you can graphically separate the acceleration vector into parts in the direction of travel and direction of turning. In this section, we will look at how to calculate these separate parts algebraically, understand what is being measured by these quantities, and show how these tools will give us important physical measurements as well as an efficient tool for calculating $\vec{N}$, the unit normal vector to a curve.

10.7.2 Splitting Acceleration

Recall that by the definition of $\vec{T}$, $\vec{v}$ is parallel to $\vec{T}$ when $\vec{v}$ exists and is nonzero. Additionally, Key Idea 10.5.12 states that $\vec{T} \cdot \vec{N} = 0$. In other words,

all of the velocity is in a direction parallel to the unit tangent vector and the unit normal vector is *always* orthogonal to the unit tangent vector to a curve, when they exist. In this subsection, we will look at splitting the acceleration vector into parts in the direction of travel and the direction of turning.

Suppose $\vec{a}(t)$ is the acceleration vector for the motion along a curve in $\mathbb{R}^3$. From Section 10.5, we have the TNB frame as a way of describing key directions related to motion along a curve in space. We would like to split the acceleration vector into a part in each of directions corresponding to $\vec{T}$, $\vec{N}$, and $\vec{B}$. Algebraically, we want to find scalars c_1, c_2, and c_3 such that $\vec{a} = c_1\vec{T} + c_2\vec{N} + c_3\vec{B}$. To do so, we will make use of several key facts from Section 10.5:

$$\vec{T}\cdot\vec{N} = 0 \qquad \vec{T}\cdot\vec{B} = 0 \qquad \vec{N}\cdot\vec{B} = 0 \qquad \left\|\vec{T}\right\| = \left\|\vec{N}\right\| = \left\|\vec{B}\right\| = 1.$$

If we compute the dot product of each of $\vec{T}$, $\vec{N}$, and $\vec{B}$ with the expression for $\vec{a}$, we find

$$\begin{aligned}
\vec{a}\cdot\vec{T} &= (c_1\vec{T} + c_2\vec{N} + c_3\vec{B})\cdot\vec{T} \\
&= c_1(\vec{T}\cdot\vec{T}) + c_2(\vec{N}\cdot\vec{T}) + c_3(\vec{B}\cdot\vec{T}) \\
&= c_1 \\
\vec{a}\cdot\vec{N} &= (c_1\vec{T} + c_2\vec{N} + c_3\vec{B})\cdot\vec{N} \\
&= c_1(\vec{T}\cdot\vec{N}) + c_2(\vec{N}\cdot\vec{N}) + c_3(\vec{B}\cdot\vec{N}) \\
&= c_2 \\
\vec{a}\cdot\vec{B} &= (c_1\vec{T} + c_2\vec{N} + c_3\vec{B})\cdot\vec{B} \\
&= c_1(\vec{T}\cdot\vec{B}) + c_2(\vec{N}\cdot\vec{B}) + c_3(\vec{B}\cdot\vec{B}) \\
&= c_3
\end{aligned}$$

These computations demonstrate that because the TNB frame directions are all unit vectors that are orthogonal to each other, we can compute the amount of acceleration in each of these directions with a dot product!

Definition 10.7.1 The **tangential component of acceleration** is $a_{\vec{T}} = \vec{a}\cdot\vec{T}$ and the **normal component of acceleration** is $a_{\vec{N}} = \vec{a}\cdot\vec{N}$. ◆

Our use of the word "component" in Definition 10.7.1 is consistent with our use in Section 9.4 because $\vec{T}$ and $\vec{N}$ are both unit vectors. Furthermore, because the unit tangent and unit normal vectors are also orthogonal to each other, $a_{\vec{T}}$ and $a_{\vec{N}}$ give the amount of acceleration in the direction of travel and the direction of turning, respectively. You may notice that we only gave names for acceleration in the tangential and normal directions above. The acceleration vector for motion along a curve can *only* have components in the tangential and normal directions to the curve. It is not possible for the acceleration vector of parameterized curve to have any component in a direction, such as $\vec{B}$, that does not lie in the plane defined by $\vec{T}$ and $\vec{N}$, which we state formally and justify below.

Key Idea 10.7.2 *If $\vec{a}(t)$ is the acceleration vector for a parameterization of a curve in $\mathbb{R}^n$, then*

$$\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$$

Proof. We claim that the scalar $a_{\vec{N}}$ measures the acceleration due to change in the direction of $\vec{v}$ and the scalar $a_{\vec{T}}$ measures the acceleration due to the change in magnitude of $\vec{v}$. This interpretation of $a_{\vec{T}}$ and $a_{\vec{N}}$ follows from a product rule argument based on how we split the velocity vector into speed

(magnitude) and $\vec{T}$ (direction). We start by splitting velocity into its direction and magnitude:

$$\vec{v}(t) = \text{speed}(t) \cdot \vec{T}(t)$$

Taking the derivative of the velocity using a product rule, we get the following formula for acceleration:

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d}{dt}(\text{speed}) \cdot \vec{T} + (\text{speed}) \cdot \frac{d}{dt}(\vec{T})$$

We can rearrange the terms in our definition for $\vec{N}$ to solve for $\vec{T}\,'(t)$, which gives

$$\vec{N} = \frac{\vec{T}\,'}{\left\|\vec{T}\,'\right\|} \Rightarrow \vec{T}\,' = \left\|\vec{T}'\right\| \vec{N}.$$

Applying this to our acceleration equation above, we get

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d}{dt}(\text{speed}) \cdot \vec{T} + (\text{speed}) \cdot \left\|\frac{d\vec{T}}{dt}\right\| \vec{N}. \tag{10.7.1}$$

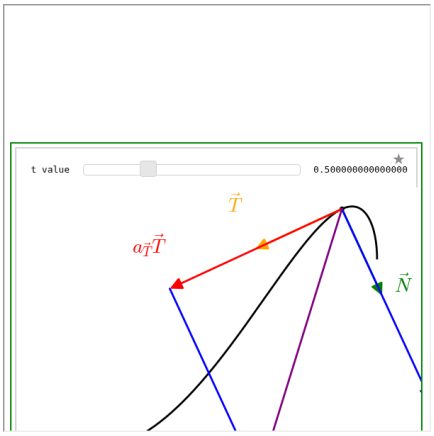
This algebraic splitting of acceleration gives a nice way to interpret $a_{\vec{T}}$. Specifically, $a_{\vec{T}}$ is the rate of change of speed!

The scalar $a_{\vec{N}}$ is harder to interpret right now because our algebraic splitting shows that $a_{\vec{N}} = (\text{speed}) \left\|\frac{d\vec{T}}{dt}\right\|$. Since $\frac{d\vec{T}}{dt}$ and $\left\|\frac{d\vec{T}}{dt}\right\|$ are as hard to calculate directly as $\vec{N}$ is, we do not yet have an intuitive interpretation for the meaning of $a_{\vec{N}}$. We will return to the meaning of $a_{\vec{N}}$ later in this section, but for now you should think of $a_{\vec{N}}$ as the acceleration due to a change in the direction of the velocity vector. In other words, $a_{\vec{N}}$ is the amount of acceleration due to turning. ■

Note that nothing in the argument for Key Idea 10.7.2 depended on being in three dimensions, and all of the calculus tools used in the argument work for any $\mathbb{R}^n$.

Definition 10.7.3 The **splitting of the acceleration (vector)** refers to finding scalars $a_{\vec{T}}$ and $a_{\vec{N}}$ such that $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$. This allows us to describe the amount of the acceleration that is in the unit tangent direction and the amount that is in the unit normal direction, relative to the motion on the curve. ◆

Figure 10.7.4 shows the splitting of the acceleration vector graphically. Move the slider to change the location on the curve where the splitting of the acceleration is shown. Examine different points along the curve to see how each of $\vec{T}$, $\vec{N}$, $a_{\vec{T}}$, and $a_{\vec{N}}$ change at different locations. The vector $a_{\vec{N}}\vec{N}$ is drawn twice to demonstrate its relationship to $\vec{N}$ and how $\vec{a}$ is a vector sum of $a_{\vec{T}}\vec{T}$ and $a_{\vec{N}}\vec{N}$.



Standalone
In Context
Embed

Figure 10.7.4 A two-dimensional curve with the splitting of the acceleration vector shown graphically

Activity 10.7.2 In this activity, we will work to understand the sign of $a_{\vec{T}}$ and $a_{\vec{N}}$.

(a) The tangential component of acceleration is defined as $a_{\vec{T}} = \vec{a} \cdot \vec{T}$, while equation (10.7.1) establishes that $a_{\vec{T}} = \frac{d(\text{speed})}{dt}$ is also true. Use these formulas to determine whether $a_{\vec{T}}$ can be zero. Either explain why $a_{\vec{T}}$ is always nonzero or describe scenarios in which $a_{\vec{T}} = 0$.

(b) Determine whether $a_{\vec{T}}$ can be negative. Either explain $a_{\vec{T}}$ is never negative or describe scenarios in which $a_{\vec{T}} < 0$

(c) The normal component of acceleration is defined as $a_{\vec{N}} = \vec{a} \cdot \vec{N}$, while equation (10.7.1) establishes that $a_{\vec{N}} = (\text{speed}) \left\| \frac{d\vec{T}}{dt} \right\|$ is also true. Use these formulas to determine whether $a_{\vec{N}}$ can be negative. Either explain why $a_{\vec{N}}$ is negative or describe scenarios in which $a_{\vec{N}} < 0$.

(d) Determine whether $a_{\vec{N}}$ can be zero. Either explain why $a_{\vec{N}}$ is always nonzero or describe scenarios in which $a_{\vec{N}} = 0$.

Exercise 7 and Exercise 8 relate the scalar and vector measurements involved in the splitting of acceleration to different modes of transportation and looks at how different actions when riding a bike or driving a car change these vector calculus quantities. Exercise 9 considers how to visually identify the different measurements involved the splitting of the acceleration vector.

10.7.3 Efficient Calculations for the Splitting of $\vec{a}$

You saw in Activity 10.5.4 that calculating $\vec{N}$ for even a curve with a very "nice" parameterization algebraically using the definition requires extensive algebraic manipulation. Before going into an exampmle of the computations involved in splitting the acceleration vector, we will identify a relationship of the vectors involved in splitting the acceleration vector that yields another way of approaching $\vec{N}$. To do this, we consider the right triangle of vectors in Figure 10.7.5. If we apply the Pythagorean Theorem to this right triangle, we can then use that the unit tangent vector and unit normal vector both have length one to conclude that $\|\vec{a}\|^2 = (a_{\vec{T}})^2 + (a_{\vec{N}})^2$.

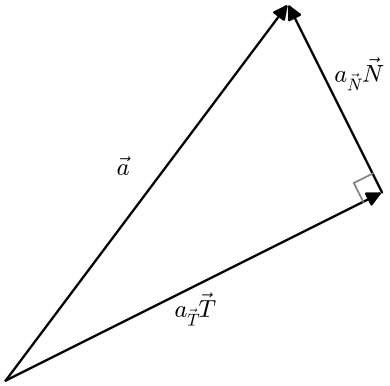


Figure 10.7.5 A plot of the right triangle of vectors made by $\vec{a}, a_{\vec{T}}\vec{T}, a_{\vec{N}}\vec{N}$

Now we can solve for $a_{\vec{N}}$ and find

$$a_{\vec{N}} = \sqrt{\|\vec{a}\|^2 - (a_{\vec{T}})^2}. \tag{10.7.2}$$

Notice that since $a_{\vec{N}}$ is *always* nonnegative, we know that we can take the positive square root when solving here. We will soon use this formula to help find $a_{\vec{N}}$. Since $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$, we can also solve for $\vec{N}$ and have

$$\vec{N} = \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}},$$

which we will use to help find $\vec{N}$.

Example 10.7.6 In this example we find the splitting of acceleration for a curve parameterized by $\vec{r}(t) = \langle t, t^2, t^3\rangle$ with $0 \le t \le 2$. While this curve's parameterization is simple algebraically, we will find that the formulas involved in the splitting of acceleration become more complex fairly quickly. To get started, we will find the splitting of the acceleration vector for a single parameter value. In the second step, we will look at finding general formulas for $a_{\vec{T}}$ and $a_{\vec{N}}$.

(a) In many circumstances, we need to compute the splitting of acceleration for only one parameter value. For this example, we will look at how clean the computations can be if we want to look only at $t = 1$. The first steps for calculating $\vec{v}$ and $\vec{a}$ are to use the definitions and substitute $t = 1$

$$\begin{aligned} \vec{r}(t) &= \langle t, t^2, t^3\rangle & \vec{r}(1) &= \langle 1, 1, 1\rangle \\ \vec{v}(t) &= \langle 1, 2t, 3t^2\rangle & \vec{v}(1) &= \langle 1, 2, 4\rangle \\ \vec{a}(t) &= \langle 0, 2, 6t\rangle & \vec{a}(1) &= \langle 0, 2, 6\rangle. \end{aligned}$$

The next quantity to calculate is $\vec{T}$. Note that

$$\text{speed}(1) = \|\vec{v}(1)\| = \sqrt{1 + 4 + 9} = \sqrt{14}$$

here. Thus,

$$\vec{T}(1) = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\langle 1, 2, 3\rangle}{\sqrt{14}}.$$

This allows us to calculate $a_{\vec{T}}(1)$ from the definition

$$a_{\vec{T}}(1) = \vec{a}(1) \cdot \vec{T}(1) = \langle 0, 2, 6\rangle \cdot \frac{\langle 1, 2, 3\rangle}{\sqrt{14}}$$

$$= \frac{1}{\sqrt{14}}(1(0) + 2(2) + 3(6)) = \frac{22}{\sqrt{14}}.$$

We now use $a_{\vec{N}}(1) = \sqrt{\|\vec{a}(1)\|^2 - (a_{\vec{T}}(1))^2}$ to efficiently calculate the normal component of the acceleration and have

$$a_{\vec{N}}(1) = \sqrt{\|\vec{a}(1)\|^2 - (a_{\vec{T}}(1))^2} = \sqrt{\frac{76}{14}} = \sqrt{\frac{38}{7}}.$$

We can now compute how much acceleration there is due to turning and use that to explicitly compute $\vec{N}(1)$. We have

$$\begin{aligned} a_{\vec{N}}(1)\vec{N}(1) &= \vec{a}(1) - a_{\vec{T}}(1)\vec{T}(1) \\ &= \langle 0, 2, 6\rangle - \frac{22}{\sqrt{14}}\frac{\langle 1, 2, 3\rangle}{\sqrt{14}} \\ &= \left\langle -\frac{11}{7}, -\frac{16}{7}, \frac{9}{7}\right\rangle. \end{aligned}$$

Thus,

$$\vec{N} = \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}} = \sqrt{\frac{7}{38}}\left\langle -\frac{11}{7}, -\frac{16}{7}, \frac{9}{7}\right\rangle.$$

(b) We now consider splitting the acceleration vector for the same curve, but we do so for an arbitrary parameter value t. That is, we will calculate $\vec{v}$, $\vec{a}$, $\vec{T}$, $a_{\vec{T}}$, $a_{\vec{N}}$, and $\vec{N}$ as functions of t. Along the way, we will omit a number of straightforward steps in simplifying. Since we already calculated $\vec{r}(t)$, $\vec{v}(t)$, and $\vec{a}(t)$ in the previous part, we can move on to calculate $\vec{T}$. Since $\text{speed}(t) = \|\vec{v}(t)\|$, we have $\text{speed}(t) = \sqrt{1 + 4t^2 + 9t^4}$ here. Thus,

$$\vec{T}(t) = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\langle 1, 2t, 3t^2\rangle}{\sqrt{1 + 4t^2 + 9t^4}}.$$

The definition $a_{\vec{T}} = \vec{a} \cdot \vec{T}$ allows us to calculate $a_{\vec{T}}$:

$$\begin{aligned} a_{\vec{T}} = \vec{a} \cdot \vec{T} &= \langle 0, 2, 6t\rangle \cdot \frac{\langle 1, 2t, 3t^2\rangle}{\sqrt{1 + 4t^2 + 9t^4}} \\ &= \frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}}. \end{aligned}$$

For this parameterization,

$$\begin{aligned} a_{\vec{N}} &= \sqrt{\|\vec{a}\|^2 - (a_{\vec{T}})^2} \\ &= \sqrt{(4 + 36t^2) - \left(\frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}}\right)^2} \\ &= \sqrt{\frac{4(9t^4 + 9t^2 + 1)}{1 + 4t^2 + 9t^4}}. \end{aligned}$$

We can now also calculate $\vec{N}$ as a function of t without having to go through the complex algebra required to compute $\frac{d\vec{T}}{dt}$. We have

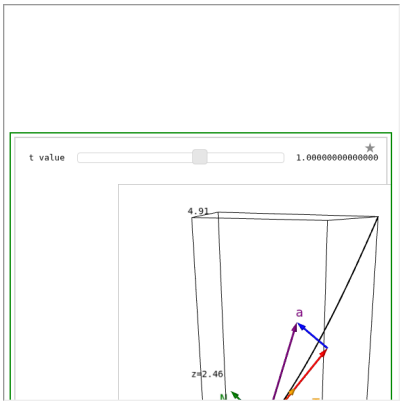
$$a_{\vec{N}}\vec{N} = \vec{a} - a_{\vec{T}}\vec{T}$$

$$= \langle 0, 2, 6t \rangle - \frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}} \frac{\langle 1, 2t, 3t^2 \rangle}{\sqrt{1 + 4t^2 + 9t^4}},$$

so we can write $\vec{N}$ as

$$\begin{aligned} \vec{N} &= \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}} \\ &= \sqrt{\frac{1 + 4t^2 + 9t^4}{4(9t^4 + 9t^2 + 1)}} \left(\langle 0, 2, 6t \rangle - \frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}} \frac{\langle 1, 2t, 3t^2 \rangle}{\sqrt{1 + 4t^2 + 9t^4}} \right). \end{aligned}$$

While this process is still algebraically tedious, these steps are less difficult than finding $\vec{N}$ directly (as in Activity 10.5.4). We also now have insight into how the acceleration is split in the direction of travel and direction of turning. Graphically, we can represent this splitting with Figure 10.7.7.



Standalone
In Context
Embed

Figure 10.7.7 The curve given by $\vec{r}(t) = \langle t, t^2, t^3 \rangle$ with the splitting of acceleration shown graphically

▲

The method used in the previous example can be applied broadly and shows how to calculate all parts of Definition 10.7.3 in terms of $\vec{v}$ and $\vec{a}$ (for both an interval or a single value of t).

Efficient Calculation of $a_{\vec{T}}, a_{\vec{N}}, \vec{T}, \vec{N}$ in terms of $\vec{v}$ and $\vec{a}$.

$$\begin{aligned} \vec{T} &= \frac{\vec{v}}{\|\vec{v}\|} \\ a_{\vec{T}} &= \frac{\vec{a} \cdot \vec{v}}{\|\vec{v}\|} \Rightarrow a_{\vec{T}}\vec{T} = \frac{\vec{a} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\vec{v} \\ a_{\vec{N}} &= \left\| \vec{a} - a_{\vec{T}}\vec{T} \right\| = \sqrt{\|\vec{a}\|^2 - (a_{\vec{T}})^2} = \sqrt{(\vec{a} \cdot \vec{a}) - \frac{(\vec{a} \cdot \vec{v})^2}{\|\vec{v}\|}} \\ a_{\vec{N}}\vec{N} &= \vec{a} - a_{\vec{T}}\vec{T} \\ \vec{N} &= \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}} \end{aligned}$$

Activity 10.7.3 Compute the splitting of the acceleration for the curve given by $\vec{r}(t) = \langle t, \frac{t^2}{2}, \frac{t^3}{6} \rangle$ at $t = 1$. You should calculate $a_{\vec{T}}$, $a_{\vec{N}}$, $\vec{T}$, and $\vec{N}$. Ad-

ditionally, you should verify $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$ and that $\vec{T}$ is orthogonal to $\vec{N}$.

Exercise 7, Exercise 8, and Exercise 9 offer conceptual and physical meaning to the different measurements involved in the splitting of acceleration. We will take a moment now to relate $a_{\vec{N}}$, the acceleration due to a change in the direction of travel, and κ, the curvature, which measures rate of change for the direction of travel on a curve.

The previous statement may seem a bit like we said the same thing twice, but $a_{\vec{N}}$ and κ are *not* the same. Theorem 10.6.7 shows that

$$\kappa = \frac{\left\| \frac{d\vec{T}}{dt} \right\|}{\|\vec{r}\,'(t)\|}.$$

We can solve for $\left\| \frac{d\vec{T}}{dt} \right\|$ in terms of κ to get $\left\| \frac{d\vec{T}}{dt} \right\| = \kappa \cdot \text{speed}$. We saw earlier in this section that

$$a_{\vec{N}} = \text{speed} \cdot \left\| \frac{d\vec{T}}{dt} \right\|$$

Thus, we can write $a_{\vec{N}}$ in terms of speed and curvature as

$$a_{\vec{N}} = \kappa \cdot (\text{speed})^2. \tag{10.7.3}$$

Equation (10.7.3) helps demonstrate the difference between $a_{\vec{N}}$ and κ. Specifically, κ measures how fast the path is curving and $a_{\vec{N}}$, the normal component of acceleration, measures the acceleration felt due to turning. Equation (10.7.3) shows that if two drivers are driving along the same road but the first driver is going twice as fast as the second driver is going, then the first driver feels *four times* as much acceleration due to turning as the second driver does, even though both drivers are following the exact same road. Curvature is how fast the *road* is turning and $a_{\vec{N}}$ is how fast it feels like the driver is turning based on speed, which is determined by the parameterization.

Activity 10.7.4 The Driver or the Road? In this activity, you will determine if the measurements related to the splitting of acceleration are properties of the driver or properties of the road. This is a continuation of Activity 10.4.3. As a reminder, a measurement is a property of the driver if the value of that measurement ***can*** be different for different drivers when measured at the same location on the racetrack. A measurement is a property of the road when different drivers ***must*** have the same value(s) when measured at the same location on the racetrack.

(a) Pedal Usage: We measure the pedal usage as $+$, $-$, or 0 in the following way:

- $+$ if the gas pedal is being used
- $-$ if the brake pedal is being used
- 0 if no pedal is being used

Since our drivers are safety minded, they do not use more than one pedal at a time. Note that this measurement is a sign rather than a numerical value. Which vector calculus quantity's sign corresponds to the pedal usage? Write a sentence to explain your answer.

Hint. The vector calculus quantities related to the motion along a curve we have covered are time, position, velocity, acceleration, speed, arc

length, unit tangent vector, unit normal vector, binormal vector, tangential acceleration component, normal acceleration component, curvature, and osculating circle.

(b) Describe how the value of $a_{\vec{N}}$ would be felt by the driver in our analogy.

(c) In terms of the race car analogy, explain why $a_{\vec{N}}$ can't be negative.

(d) Write a few sentences about whether $a_{\vec{T}}$ is a property of the driver or a property of the road.

Hint. $a_{\vec{T}} = \frac{d}{dt}(\text{speed})$

(e) Write a few sentences about whether $a_{\vec{N}}$ is a property of the driver or a property of the road.

Hint. $a_{\vec{N}} = \kappa(\text{speed})^2$

Summary

- The acceleration vector for a parameterization of a curve, C, can be split into parts in the direction of travel and direction of turning

$$\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$$

where $a_{\vec{T}} = \vec{a} \cdot \vec{T}$ and $a_{\vec{N}} = \vec{a} \cdot \vec{N}$.

- The scalar $a_{\vec{T}}$ measures the rate of change of the speed. The scalar $a_{\vec{N}}$ measure the acceleration due to change in direction and can be calculated by

$$a_{\vec{N}} = \kappa(speed)^2$$

where κ is the curvature.

10.7.4 Exercises

The WeBWorK problems are written by many different authors. Some authors use parentheses when writing vectors, e.g., $(x(t), y(t), z(t))$ instead of angle brackets $\langle x(t), y(t), z(t)\rangle$. Please keep this in mind when working WeBWorK exercises.

1. Calculate the velocity and acceleration vectors, and speed for

$$\mathbf{r}(t) = \langle \cos(t), \sin(3t), \cos(3t) \rangle$$

when $t = \frac{5\pi}{3}$.
Velocity: _____
Acceleration: _____
Speed: _____
Usage:To enter a vector, for example $\langle x, y, z\rangle$, type "< x, y, z >"

2. If $\mathbf{r}(t) = -8t\mathbf{i} + 8t^2\mathbf{j} + 4t\mathbf{k}$, compute the tangential and normal components of the acceleration vector.
Tangential component $a_T(t) =$ _____
Normal component $a_N(t) =$ _____

3. If $\mathbf{r}(t) = \cos(1t)\mathbf{i} + \sin(1t)\mathbf{j} - 7t\mathbf{k}$, compute the tangential and normal components of the acceleration vector.
Tangential component $a_T(t) =$ _____
Normal component $a_N(t) =$ _____

4. In this exercise, we will compute and compare elements of the splitting of acceleration for three different people (Nick, Jane, and Tony) as they drive around the same racetrack.

(a) Suppose that at time $t = b$, Jane is at a location which we will denote P_1 in her race car. The velocity and acceleration vectors for Jane are $\langle 2, 0, -2\rangle$ and $\langle 3, 2, -3\rangle$. Calculate the following:

(a) $\vec{T}(b)$

(b) $a_{\vec{T}}(b)$

(c) $a_{\vec{N}}(b)$

(d) $\vec{N}(b)$

(b) Suppose that at time $t = c$, Nick is at a location P_1 in his race car. Without any calculation at all, answer the following:

(a) $\vec{T}(c)$

(b) $\vec{N}(c)$

(c) The velocity and acceleration vectors for Nick (at time c) are $\langle 1, 0, -1\rangle$ and $\langle -1, 2, 1\rangle$. Calculate the following:

(a) $a_{\vec{T}}(c)$

(b) $a_{\vec{N}}(c)$

(d) Is Jane using the gas pedal or the brake pedal when she is at P_1? Explain how you know this based on your calculations above.

(e) Is Nick using the gas pedal or the brake pedal when he is at P_1? Explain how you know this based on your calculations above.

(f) Nick and Jane talk to their friend Tony who claims that his velocity and acceleration vectors at position P_1 are $\langle 4, 0, 6\rangle$ and $\langle 8, 9, 10\rangle$. Explain how Jane and Nick know he is lying.

5.

(a) Compute the splitting of the acceleration for the curve given by $\vec{r}(t) = \langle t, \frac{t^2}{2}, \frac{t^3}{6}\rangle$ at $t = -1$. You should calculate $a_{\vec{T}}$, $a_{\vec{N}}$, $\vec{T}$, and $\vec{N}$, then verify $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$ and that $\vec{T}$ is orthogonal to $\vec{N}$.

(b) Compute the splitting of the acceleration for the curve given by $\vec{r}(t) = \langle t, e^t, te^t\rangle$ at $t = 0$. You should calculate $a_{\vec{T}}$, $a_{\vec{N}}$, $\vec{T}$, and $\vec{N}$, then verify $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$ and that $\vec{T}$ is orthogonal to $\vec{N}$.

6. Given a curve with a parameterization such that $\vec{v}(11) = \langle 3, 4\rangle$ and $\vec{a}(11) = \langle -2, 1\rangle$.

(a) Find $\vec{T}(11)$

(b) Find $a_{\vec{T}}(11)$

(c) Find $a_{\vec{T}}(11)\vec{T}(11)$

(d) Find $a_{\vec{N}}(11)$

(e) Find $a_{\vec{N}}(11)\vec{N}(11)$

(f) Find $\vec{N}(11)$

7. In this exercise we want to relate the scalar and vector measurements involved in the splitting of acceleration to different modes of transportation. While everyone has different experiences with piloting different modes of transportation, we will phrase this question in terms of driving a car. Feel free to substitute another mode of transportation like riding a bike, riding a scooter, driving a boat, rowing a boat, flying a plane, or flying a spaceship.

(a) In the context of driving a car (or substitute your mode of transportation), write a couple sentences about what vector calculus quantity is controlled by the brake pedal. Be specific which vector calculus quantity is involved and how the brake pedal changes this measurement.

Hint. The vector calculus quantities we have covered are time, position, velocity, acceleration, speed, arc length, unit tangent, unit normal, binormal, tangential acceleration component, and normal acceleration component.

(b) In the context of driving a car (or substitute your mode of transportation), write a couple sentences about what vector calculus quantity is controlled by the accelerator pedal. Be specific which vector calculus quantity is involved and how the accelerator pedal changes this measurement.

(c) In the context of driving a car (or substitute your mode of transportation), write a couple sentences about what vector calculus quantity is controlled by the direction of the steering wheel (consider this to be either left, right, or straight). Be specific which vector calculus quantity is involved and how the the direction of the steering wheel changes this measurement.

8. In this exercise we want to relate the scalar and vector measurements involved in the splitting of acceleration to different modes of transportation. While everyone has different experiences with piloting different modes of transportation, we will phrase this question in terms of riding a bike. Feel free to substitute another mode of transportation like driving a car, riding a scooter, driving a boat, rowing a boat, flying a plane, or flying a spaceship.

(a) In the context of riding a bike, what does the unit tangent vector, $\vec{T}$, measure?

(b) In the context of riding a bike, what does the unit normal vector, $\vec{N}$, measure?

(c) In the context of riding a bike, how would you increase $a_{\vec{T}}$?

(d) In the context of riding a bike, how would you decrease $a_{\vec{T}}$?

(e) In the context of riding a bike, how would you decrease $a_{\vec{N}}$?

(f) In the context of riding a bike, how would you increase $a_{\vec{N}}$?

9. In this exercise, we will look at how to visually identify the different measurements involved the splitting of the acceleration vector. While each of $\vec{a}$, $\vec{T}$, $\vec{N}$, $a_{\vec{T}}\vec{T}$, and $a_{\vec{N}}\vec{N}$ are drawn, the colors have been randomized so you will need to look at how each element is related to the motion along the curve.

(a) The vectors $\vec{a}$, $\vec{T}$, $\vec{N}$, $a_{\vec{T}}\vec{T}$, and $a_{\vec{N}}\vec{N}$ are drawn in different colors in Figure 10.7.8. Match each color with proper vector.

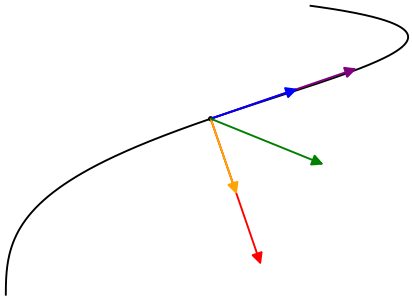


Figure 10.7.8 A 2D curve with splitting of acceleration shown using randomized colors

(b) The vectors $\vec{a}$, $\vec{T}$, $\vec{N}$, $a_{\vec{T}}\vec{T}$, and $a_{\vec{N}}\vec{N}$ are drawn in different colors in Figure 10.7.9. Match each color with proper vector.

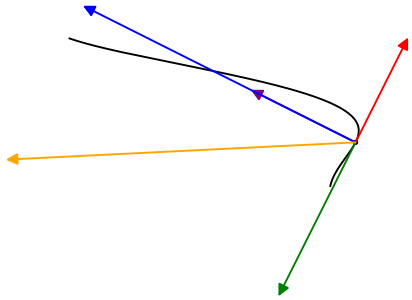


Figure 10.7.9 A 2D curve with splitting of acceleration shown using randomized colors

(c) The vectors $\vec{a}$, $\vec{T}$, $\vec{N}$, $a_{\vec{T}}\vec{T}$, and $a_{\vec{N}}\vec{N}$ are drawn in different colors in Figure 10.7.10. Match each color with proper vector.

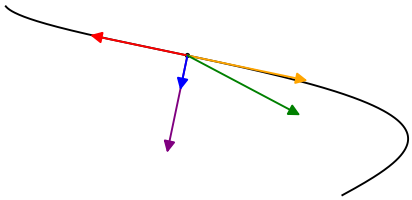


Figure 10.7.10 A 2D curve with splitting of acceleration shown using randomized colors

Chapter 11

Derivatives of Multivariable Functions

11.1 Calculus of Several Variables

For much of your study of algebra, precalculus, and calculus you have been focused on working with functions that have a scalar input and a scalar output, like $y = f(x)$. In Chapter 10, we used vector-valued functions of one variable as our first case of multivariable functions. Vector-valued functions of one variable have a single scalar input and an output corresponding to multiple variables that we either organized as a vector output such as $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ or as parametric functions of the form $(x(t), y(t), \text{and } z(t))$. Our work in Chapter 10 focused on paths in space and the motion of an object on these paths. This narrow focus was because we had just one direction to move while staying on the path. We worked with vector-valued functions of one variable as our first new class of functions because the calculus of these object was relatively easy. We applied limits, derivatives, and integrals to these functions componentwise, but we saw how useful a combination of vector tools and calculus measurements was for describing many features of vector-valued functions of one variable and their graphs as paths in space.

A wide range of theoretical and applied problems involve a larger space of inputs and outputs. For instance, when studying weather patterns and behavior it is useful to measure temperature or atmospheric pressure. Both temperature and pressure are scalar measurements because they are measured by a single number. However, these measurements vary over three dimensions (location in terms of north/south, east/west, and elevation). Temperature can be given by a function with a location in three dimensions as the input and the temperature at that location, a scalar, as an output.

Wind direction and strength are also very important when working with weather patterns. Because it has both magnitude and direction, wind is measured with a vector that varies by location in a three-dimensional space. Therefore, wind would be given by a function that takes a location in space as its input and outputs a vector. We would call both the temperature and the wind functions **functions of several variables** or **multivariable functions** because each of these functions has multiple scalar inputs. We will look at the calculus of multivariable functions with scalar outputs such as temperature and pressure in this chapter and the next. The final chapter 13 of the book studies functions with multivariable inputs *and* outputs.

In Chapter 10, all of the applications and analogies were centered around

an object moving along a curve in space. When working with functions of several variables, we will want to use several different types of functions for our applications and examples because of the possible conceptual relationships between inputs and outputs. For instance, the first type of example we will consider is thinking of the elevation or height of land as being a function of map coordinates. In this case, the input and output have the same units and are connected to make the surface. In other words the input and output are each the same type of measurement and would be easily understood by the plot of a surface.

The example of temperature as a multivariable function as described above gives a different kind of interpretation, one in which the output is a different kind of measurement than the inputs. For instance, think of a function whose input is the (x, y) location on a heated metal plate and the output is the temperature of the metal plate at this location. In this case, the input is thought of as a point or location and the output is a measure of thermal energy at that location. The inputs and outputs are different kinds of measurements, but a plot of the different temperature values across the different (x, y) locations would still visually make sense.

We can also consider a function that describes the horizontal distance traveled by an arrow before the arrow hits the ground as determined by the angle at which the arrow is shot and the initial speed of the arrow. In this case, the inputs to our function are an angle and a speed, while the output is a distance representing horizontal displacement. The units and types of measurements for the inputs and outputs are all different, and a plot of the corresponding surface of displacement, angle, and speed values would be a more abstract visual representation than in the temperature and elevation examples.

In the next few chapters, we will use these types of examples, as well as applications to economics, to motivate our study of functions of several variables.

11.2 Functions of Several Variables

Motivating Questions

- How can we use the idea of a function to model relationships that depend on multiple inputs?
- How does thinking about fixing one of the inputs in a multivariable relationship help us understand that relationship?
- How does thinking about the collection of input values that result in the same output from a multivariable relationship help us understand that relationship?

11.2.1 Introduction

In this section, we will start our work with (scalar valued) multivariable functions by looking at several basic elements of these type of functions, such as domain, range, different presentations, and how to plot these functions. Our preview activity will use a finance example to get you used to the notation and descriptions of different aspects of functions of several variables.

Preview Activity 11.2.1 Suppose you invest money in an account that pays 5% interest compounded continuously. If you have an initial investment of P dollars in the account, then A, the amount of money in the account after t

years is given by

$$A = Pe^{0.05t}.$$

The variables P and t are independent of each other, so using functional notation we write

$$A(P, t) = Pe^{0.05t}.$$

(a) Find the amount of money in the account after 7 years if you originally invest 1000 dollars.

(b) Evaluate $A(5000, 8)$. Write a sentence to explain what this calculation represents.

(c) Now consider only the situation where the amount invested is fixed at 1000 dollars. Calculate the amount of money in the account after t years as indicated in the table below. Round payments to the nearest penny.

Duration (in years)	2	3	4	5	6
Amount (dollars)					

(d) Now consider the situation where we want to know the amount of money in the account after 10 years given various initial investments. Calculate the amount of money in the account as indicated in the table below. Round payments to the nearest penny.

Initial investment (dollars)	500	1000	5000	7500	10000
Amount (dollars)					

(e) Describe as best you can what combinations of initial investments and time will result in an account containing $10,000.

11.2.2 Functions of Several Variables

Up to this chapter, we have primarily been concerned with functions of a single variable. Remember that a function is a rule that assigns exactly one output for each allowed input. For instance, the rule that assigns a student ID number to each student at your school is a function because each student (input) gets assigned one and only one ID number. The rule that assigns the classrooms for your courses this semester is *not* a function because you likely have more than one classroom (multiple outputs) for each student (a single input).

We saw the behavior of a function in Preview Activity 11.2.1, where each pair of inputs, (P, t), produces a single output $A(P, t)$. Additionally, the values of the two variables P and t did not depend on one another. That is, we could choose any value of P without limiting what value t might have, and we could select any value of t to use without regard to what value P might have. For that reason we say that the variables t and P are **independent** of each other. Thus, we call $A = A(P, t)$ a function of the two independent variables P and t. This is the key idea in defining a function of two independent variables.

Definition 11.2.1 A **function f of two independent variables** is a rule that assigns to each ordered pair (x, y) in some set D exactly one real number $f(x, y)$. ◆

There is no reason to restrict ourselves to functions of only two variables; we can use any number of variables we like. For example,

$$f(x, y, z) = x^2 - 2xz + \cos(y)$$

defines f as a function of the three variables x, y, and z. In general, a function of n independent variables is a rule that assigns to an ordered n-tuple $(x_1, x_2, \ldots, x_n)$ in some set D exactly one real number. [1]

As with functions of a single variable, it is important to understand the set of inputs for which the function is defined.

Definition 11.2.2 The **domain** of a function f is the set of all inputs for which the function is defined.

The **range** of a function f is the set of all actual outputs for the function. ◆

Recall that common ways for scalar functions to *not* be defined for a particular input include:

- dividing by zero,
- taking the square root of negative numbers, and
- restricted domains from other functions like arcsin or logarithms.

For multivariable functions, we have similar phenomena to consider.

Example 11.2.3 In this example, we will look at the domain and range of a couple of multivariable functions.

(a) Let $f(x, y) = \sqrt{xy}$. For the domain of f we can only use (x, y)-inputs where $\sqrt{xy}$ is defined. This requires that xy be nonnegative. We will use **set-builder notation** to describe this domain. The domain of f is $\{(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$. This notation is read as "the set of (x, y) coordinates in the cartesian plane such that x times y is greater than or equal to zero". Set builder notation is flexible because this follows the form {kind of objects | conditions on objects}. The symbol | is read as "such that". Note that the domain of f does not say that x and y are both nonnegative (quadrant I), but rather that the *product* of x and y needs to be nonnegative (quadrants I *and* III).

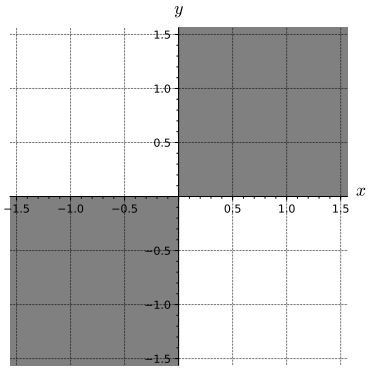


Figure 11.2.4 A plot of the domain of $f(x, y) = \sqrt{xy}$

When we consider the range of $f(x, y) = \sqrt{xy}$ we need to think about what values can be outputs of f. You may be tempted to say that outputs are scalars, so the range should be all real numbers, $\mathbb{R}$. However, there is no way to get a negative output for f. Thus, not all real numbers are *actual* outputs of f. Instead, the range of f will be the interval $[0, \infty)$

[1] You may notice that we use variable names like x, y, z when using three variables but when there are more than three variables, it will be convenient to index the variables like x_i so that we can refer to the variables by index. For example the fifth variable in $(x_1, x_2, \ldots, x_n)$ will be x_5 and the last two variables will be x_{n-1} and x_n.

because those are the values that f actually outputs. (Specifically, for any nonnegative real number c, $f(c, c) = \sqrt{cc} = c$.)

(b) We now consider the domain and range of $g(x, y) = \frac{\arcsin(y)}{2x}$. First we need to recall the domain and range of the arcsine function are $[-1, 1]$ and $[-\frac{\pi}{2}, \frac{\pi}{2}]$, respectively. Thus, the only valid inputs to g must have y-values in the interval $[-1, 1]$ and we must also have $x \neq 0$ because we cannot divide by zero. Thus the domain of g is $\{(x, y) \in \mathbb{R}^2 \mid x \neq 0 \text{ and } -1 \leq y \leq 1\}$.

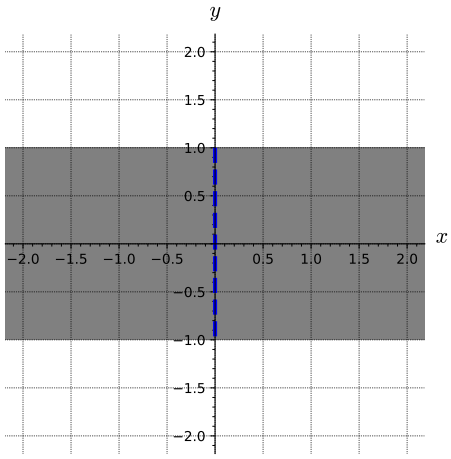


Figure 11.2.5 A plot of the domain of $g(x, y) = \frac{\arcsin(y)}{2x}$

The range of arcsine is $[-\frac{\pi}{2}, \frac{\pi}{2}]$ but since the output of g will be $\frac{\arcsin(y)}{2x}$, the range of g will be $\mathbb{R}$, all real numbers. For example, the input $(\frac{\pi}{4a}, 1)$ will be sent to a for any value of $a \neq 0$ and $g(0, 1) = 0$.

▲

Activity 11.2.2 Identify the domain and range of each of the following functions. Sketch each domain in the xy-plane.

(a) $f(x, y) = x^2 + y^2$

(b) $f(x, y) = \sqrt{x^2 + y^2}$

(c) $Q(x, y) = \dfrac{x + y}{x^2 - y^2}$

(d) $s(x, y) = \dfrac{1}{\sqrt{1 - xy^2}}$

11.2.3 Representing Functions of Two Variables

Algebraic rules. You have already seen one representation of a function of several variables: the algebraic notation $f(x, y) = \sqrt{xy}$, which shows an algebraic rule to find the output of the function f for given input values x and y. This kind of representation is convenient to use with algebraic rules for derivatives or for computing the output of a function given particular inputs. The drawback to this representation is that you need a lot of intuition about the type of function being used to understand how the output changes over a range of values.

Tables. One of the techniques we use to study functions of one variable is to create a table of values. We can do the same for functions of two variables, except that our tables must allow us to keep track of both input variables. We can do this with a two-dimensional table, where we list the x-values down the first column and the y-values across the first row.

As an example, suppose we launch a projectile (perhaps by hitting a golf ball with a golf club) from ground level. Under ideal conditions, by which we mean ignoring wind resistance, spin, or any other forces except the force of gravity, the horizontal distance the object travels before hitting the ground depends on the initial velocity x the object is given, and the angle y at which it is launched. If we let f represent the horizontal distance the object travels, then f is a function of the two variables x and y, and we represent f in functional notation by

$$f(x, y) = \frac{x^2 \sin(2y)}{g},$$

where g is the acceleration due to gravity.[2] To create a table of values for f, we list the x-values down the first column and the y-values across the first row. The value $f(x, y)$ is displayed in the location where the x row intersects the y column, as shown in Table 11.2.6, where we measure x in feet per second and y in radians. For example, $f(75, 0.8) = 175.7$ is shown in the table by looking at the value in the row with $x = 75$ and column corresponding to $y = 0.8$.

Activity 11.2.3 Complete Table 11.2.6 by filling in the missing values of the function f. Round entries to the nearest tenth.

Table 11.2.6 Values of $f(x, y) = \dfrac{x^2 \sin(2y)}{g}$ with $g = 32\text{ft/s}^2$

$x \Downarrow \backslash y \Rightarrow$	0.4	0.6	0.8	1.0	1.2
50	56.0	72.8	78.1	71.0	
75		163.8	175.7	159.8	118.7
100	224.2	291.3	312.4	284.2	211.1
125	350.3	455.1		444.0	329.8
150	504.4	655.3	702.8	639.3	474.9
175	686.5	892.0	956.6	870.2	

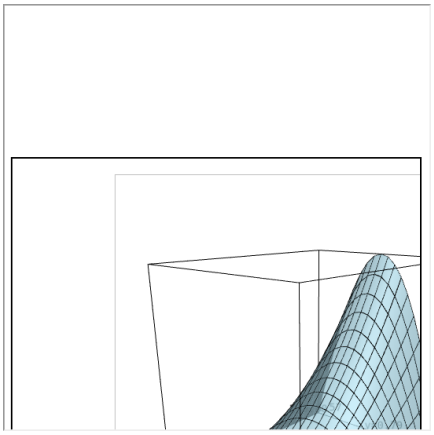
Representing a function with a table should reinforce the idea that each of the two independent values for the input variables will have a corresponding output value. However, when looking at a large table of values, it is normal to feel that this representation provides little insight into the relationship described by the function. Additionally, the table is only useful for the input values selected in the rows and columns. While this may seem like a very limited use case, many applied problems involve tables like this because an algebraic function can be extremely difficult to find for every case and it can be very expensive to test many input values to obtain the corresponding output values. You likely have already used spreadsheets to organize information in a multivariable setting like this.

Graphs. If f is a function of a single variable x, then we define the graph of f to be the set of points of the form $(x, f(x))$, where x is in the domain of f. We then plot these points using the coordinate axes in order to visualize the graph. We can do a similar thing with functions of several variables. Table 11.2.6 identifies points of the form $(x, y, f(x, y))$, and we define the graph of f to be the set of these points in $\mathbb{R}^3$.

[2] Note that g is constant, 9.8 meters per second squared or 32 feet per second squared.

Definition 11.2.7 The **graph** of a function $f = f(x, y)$ is the set of points of the form $(x, y, f(x, y))$, where the point (x, y) is in the domain of f. ◆

We also often refer to the graph of a function f of two variables as the **surface** generated by f. Points of the form $(x, y, f(x, y))$ are in three dimensions. If we consider the graph of the projectile distance function defined by $f(x, y) = \frac{x^2 \sin(2y)}{g}$, then the function f is continuous in both variables. When these points are plotted in the rectangular coordinate system, the points form a surface in three-dimensional space. The graph of the distance function f is shown in Figure 11.2.8.

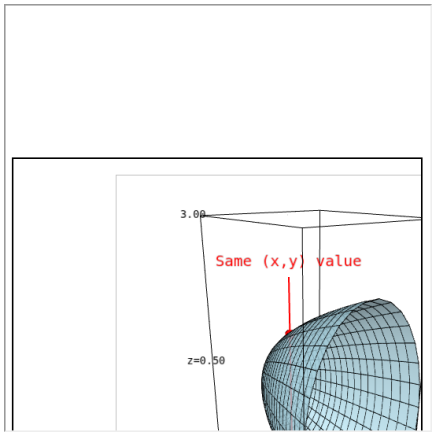


Standalone
In Context
Embed

Figure 11.2.8 The projectile distance surface given by $z = f(x, y) = \frac{x^2 \sin(2y)}{g}$

There are many graphing tools available for drawing three-dimensional surfaces as indicated in the Preface and most of the in-text three-dimensional graphics are generated by SageMath. Since we will be able to visualize graphs of functions of two independent variables using plots in three dimensions, but not functions of more than two variables, we will primarily deal with functions of two variables in this chapter. It is important to note, however, that the techniques we develop apply to functions of any number of variables and we will highlight these extensions to higher dimensions in Section 11.9.

You may have noticed that we used the notation $z = f(x, y)$ in the caption to Figure 11.2.8, which is explicitly saying that we are viewing the z-coordinate in our plot as the output of the function f. In other words, only (x, y, z) points of the form $(x, y, f(x, y))$ are part of the graph plotted. In some situations, it may be advantageous to express a graph of interest using a function with some input and output roles in a different configuration of coordinates. For instance, the surface shown in Figure 11.2.9 cannot be expressed with the z-coordinate as a function of the x- and y-coordinates because there would be more than one z-coordinate associated to inputs such as $(x, y) = (0, 1)$, as highlighted by the red points in the figure. The points on this surface can be expressed with y as a function of x and z, however! Specifically $y = g(x, z) = \frac{x^2}{5} + z^2$. We will typically look at examples where z is expressed as a function of x and y, but you should think carefully if it is advantageous to consider a different coordinate view.



Standalone
In Context
Embed

Figure 11.2.9 The surface given by $y = g(x, z) = \frac{x^2}{5} + z^2$

11.2.4 Traces

When studying functions of several variables, we are often interested in how each individual variable affects the function while the other variable is fixed. In Preview Activity 11.2.1, we saw that the amount of money in an account depends on the amount initially invested and the duration of the investment. However, if we fix the initial investment, the amount of money in the account depends only on the duration of the investment, and if we fix the duration of the investment, then the amount of money in the account depends only on the initial investment. This idea of keeping one variable constant while we allow the other to change will be an important tool for us when studying functions of several variables. [3]

As another example, consider again the projectile distance function f defined by

$$f(x, y) = \frac{x^2 \sin(2y)}{g},$$

where x is the initial velocity of an object in feet per second, y is the launch angle in radians, and g is the acceleration due to gravity (32 feet per second squared). If we hold the launch angle constant at $y = 0.6$ radians, we can

[3] This will be the first of many times we will employ the following approach

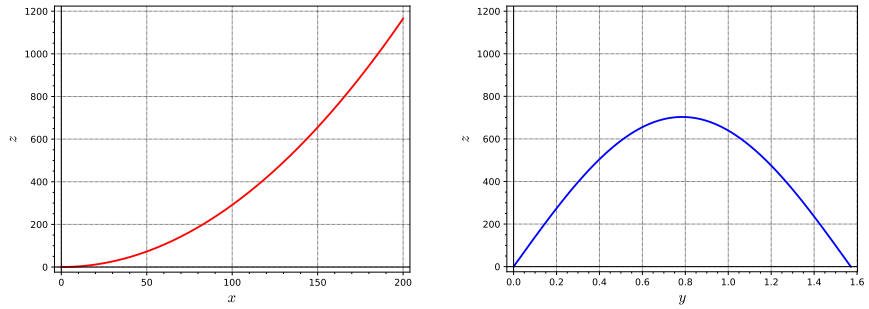
consider f a function of the initial velocity alone. In this case we have

$$f(x) = \frac{x^2}{32}\sin(2 \cdot 0.6).$$

Similarly, if we fix the initial velocity at 150 feet per second, we can consider the projectile distance as a function of the launch angle only. In this case we have

$$f(y) = \frac{150^2 \sin(2y)}{32}.$$

In Figure 11.2.10, we show two plots. Figure 11.2.10(a) shows what happens when fixing $y = 0.6$, while Figure 11.2.10(b) shows the two-dimensional graph obtained by fixing $x = 150$.



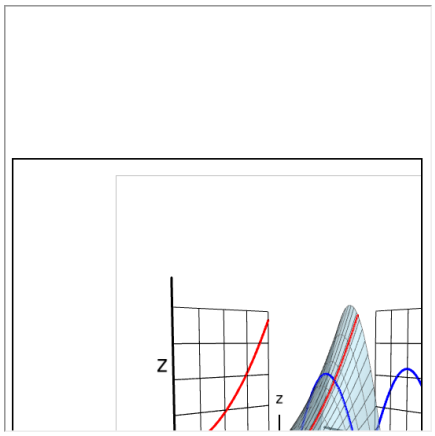
(a) A plot of $f(x, 0.6) = \frac{x^2}{32}\sin(2 \cdot 0.6)$

(b) A plot of $f(150, y) = \frac{150^2 \sin(2y)}{32}$

Figure 11.2.10 Plots illustrating values of f when one variable is fixed

We can plot the curve from Figure 11.2.10(a) on the surface by tracing out the points on the surface when $y = 0.6$, as shown in red in Figure 11.2.11. The formula for $f(x, 0.6)$ shows that f is quadratic in the x-direction. More descriptively, as we increase the launch velocity while keeping the launch angle constant, the horizontal distance the object travels increases proportional to the square of the initial velocity.

We can plot the curve from Figure 11.2.10(b) on the surface by tracing out the points on the surface when $x = 150$, as shown in blue in Figure 11.2.11. The formula for $f(150, y)$ shows that f is sinusoidal in the y-direction. More descriptively, as we increase the launch angle while keeping the initial velocity constant, the horizontal distance traveled by the object is proportional to the sine of twice the launch angle.



Standalone
In Context
Embed

Figure 11.2.11 The surface given by $z = f(x, y) = \frac{x^2 \sin(2y)}{g}$ with traces $y = 0.6$ (in red) and $x = 150$ (in blue)

The curves we define when we fix one of the independent variables in our two-variable function are called **traces**. In Figure 11.2.11, we have made a copy of the trace off to the side of the figure so you can see each trace as a curve on the proper two-dimensional slice. Figure 11.2.10 illustrates these traces as two-dimensional plots.

Definition 11.2.12 A **trace** in the x-direction of a function f of the independent variables x and y is a curve of the form $z = f(x, c)$, where c is a constant. Similarly, a **trace** in the y-direction of a function f of the two independent variables x and y is a curve of the form $z = f(c, y)$, where c is a constant. ◆

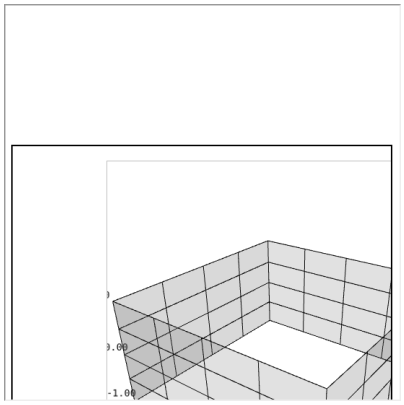
Understanding trends in the behavior of functions of two variables can be challenging, as can sketching their graphs; traces help us with both of these tasks.

Activity 11.2.4 In this activity, we investigate the use of traces to better understand a function through both tables and graphs.

(a) Identify the $y = 0.6$ trace for the distance function f defined by $f(x, y) = \frac{x^2 \sin(2y)}{g}$ by highlighting or circling the appropriate cells in Table 11.2.6. Write a sentence to describe the behavior of the function along this trace.

(b) Identify the $x = 150$ trace for the distance function by highlighting or circling the appropriate cells in Table 11.2.6. Write a sentence to describe the behavior of the function along this trace.

(c) In the next several parts, we will be looking at using traces to help us draw an accurate plot of the surface given by $z = g(x, y) = yx^2$. As a first step, find the equation for the $y = 1$ trace of $z = g(x, y) = yx^2$ and draw a graph of the $y = 1$ trace on the corresponding face in Figure 11.2.13.



Standalone
In Context
Embed

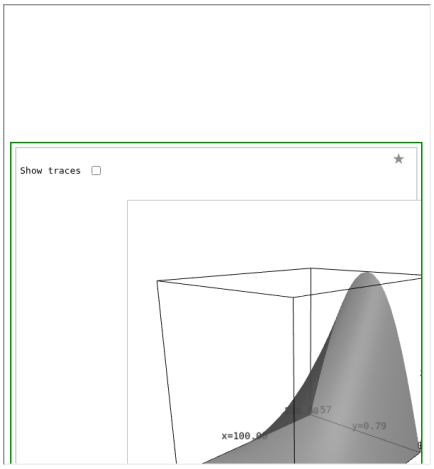
Figure 11.2.13 A bounding box for the region with $-1 \leq x, y, z \leq 1$

(d) Find the equation for each of the following traces and draw a plot of the trace on the corresponding face in Figure 11.2.13.

- $y = -1$
- $x = 1$
- $x = -1$

(e) Draw a few traces (at least three more) that correspond to values in the middle of the plot to fill in a plot of the surface given by $z = x^2 y$.

Activity 11.2.4 shows how drawing traces on a plot of a surface can give the viewer more information about the orientation of the graph. In Figure 11.2.14 you can see a plot of the surface generated in our projectile motion example. Without any of the traces drawn on the surface, it is difficult to understand the shape of the surface, and it would be even harder if not for the shading on the computer-generated graphics. You can click on the check box at the top of the figure to add traces to the plot. Notice how much easier it is to distinguish the shape of the surface at different points!



Standalone
In Context
Embed

Figure 11.2.14 A plot of the distance surface with or without traces for perspective

When drawing plots by hand, it is important to use parallel structure to make it easier for the viewer to orient the traces as being in a direction parallel to one of the coordinate axes. It may take time for you to feel confident drawing these kinds of features by hand. However, visual representations of surfaces and their features will be very valuable in our work for the next several chapters. We will continue to provide tips and suggestions to help you with

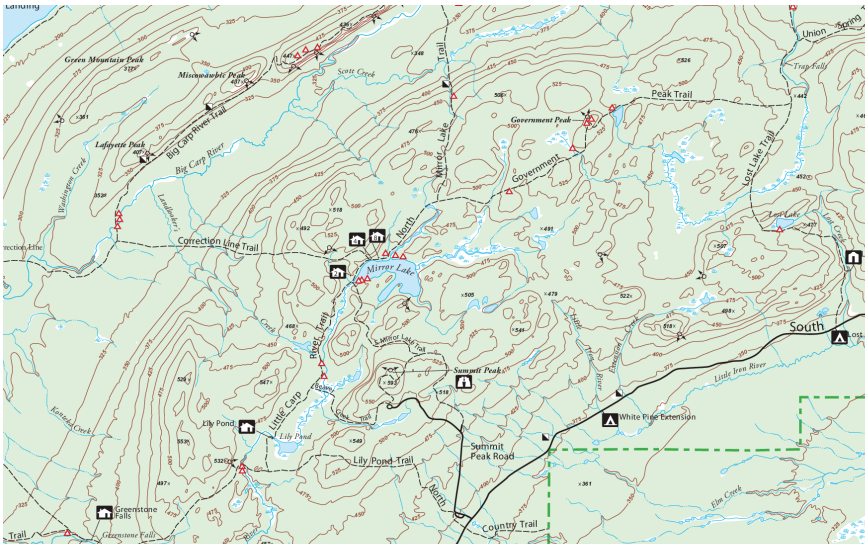
your visualizations along the way.

11.2.5 Contour Maps and Level Curves

As you saw earlier in this section, traces give important information about what a slice of a surface looks like when holding one of the inputs constant. This corresponds to looking at the intersection of a fundamental plane of the form $x = a$ or $y = b$ with the surface given by $z = f(x, y)$. These fundamental planes are oriented vertically when we consider a conventional right-handed coordinate system. In this section, we will explore the intersection of a surface given by $z = f(x, y)$ with a fundamental plane of the form $z = c$. This will correspond to looking at the points on the surface with a fixed height.

You may have seen topographic maps such as the one of the Porcupine Mountains in the upper peninsula of Michigan shown in Activity 11.2.5.[4] The curves on these maps show the locations with a particular elevation (as labeled on the curve). The amount of space between these curves also depicts the rate of change in elevation: curves on the topographic map that are close together signify steep ascents or descents, while curves that are far apart indicate slower changes in elevation. Thus, these topographic maps with curves of constant elevation can tell us a lot about three-dimensional surfaces. Mathematically, if $f(x, y)$ represents the elevation at the point (x, y), then each of the curves with constant elevation is the graph of an equation of the form $f(x, y) = k$, for some constant k.

Activity 11.2.5 Use the topographical map of the Porcupine Mountains below to answer the following questions. Note that points of interest are sometimes marked with an X and have their elevation listed.



(a) Identify the highest and lowest points you can find.

(b) Describe how your elevation would change if you walked in a straight line from the lowest to the highest elevation points. You may want to sketch a plot of the elevation along your path.

Hint. Draw a straight line path on the map from your lowest to highest point and look at the elevation as labeled at points along this path.

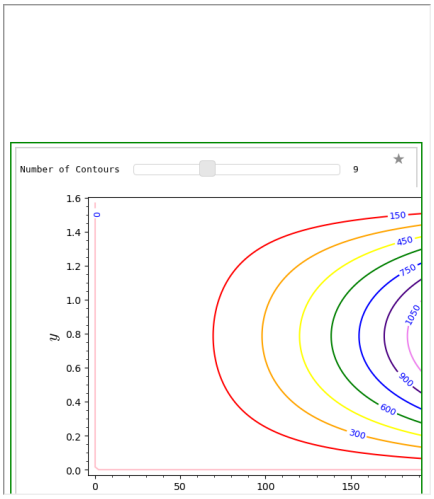
[4] Map source: Michigan Department of Natural Resources, with permission of the Michigan DNR and Bob Wild.

(c) If you walk along the Big Carp River Trail (in the top left part of the image) from the left to the right as shown on the map, describe which parts of the trail will have steep increases in elevation and which parts you think will be the most like level ground.

Curves on a surface that describe points at the same height or level are called **level curves** and as you saw in the topographic map above, a plot with level curves can be useful for representing information on a surface using just a two dimensional plot.

Definition 11.2.15 A **level curve** (or **contour**) of a function f of two independent variables x and y is a curve of the form $k = f(x, y)$, where k is a constant. ◆

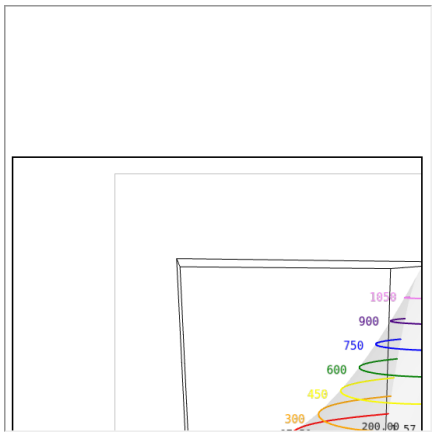
Topographical maps can be used to create a three-dimensional surface from the two-dimensional contours or level curves. For example, level curves of the distance function defined by $f(x, y) = \frac{x^2 \sin(2y)}{32}$ plotted in the xy-plane are shown in Figure 11.2.16. You can change the number of contours in the plot by moving the slider at the top of Figure 11.2.16. Additionally, you will sometimes see a legend for contours shown when different colors are used or a contour plot with different colors and shading filling in the region between contours.



Standalone
In Context
Embed

Figure 11.2.16 A contour plot of $f(x, y) = \frac{x^2 \sin(2y)}{32}$

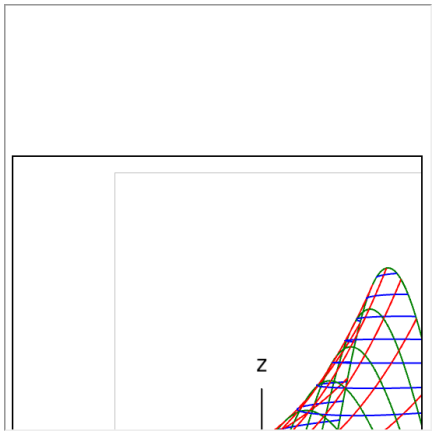
If we lift these contours and plot them at their respective heights, then we get a picture of the surface itself, as illustrated at right in Figure 11.2.17.



Standalone
In Context
Embed

Figure 11.2.17 A plot of the contours of $f(x,y) = \frac{x^2 \sin(2y)}{32}$ at the appropriate heights

The use of level curves and traces can help us construct the graph of a function of two variables. For the surface defined by $z = f(x,y) = \frac{x^2 \sin(2y)}{32}$, we can plot a grid of x-traces, y-traces, and contours to get a good idea of what the surface looks like as shown in Figure 11.2.18. If you compare the paths plotted in Figure 11.2.18 to Figure 11.2.8, you will see that the mesh shown on the surfaces in most plots *is* a grid of traces and contours and that is how the orientation of surfaces is typically drawn by computer-generated plots.



Standalone
In Context
Embed

Figure 11.2.18 A plot of the contours and traces of $f(x,y) = \frac{x^2 \sin(2y)}{32}$ at the appropriate values

Isotherms are another example of where you are likely to encounter level curves. Isotherms are a plot in which the curves on the map connect locations that have the same temperature. A plot with lots of isotherms will allow you to see how the temperature changes over a region.

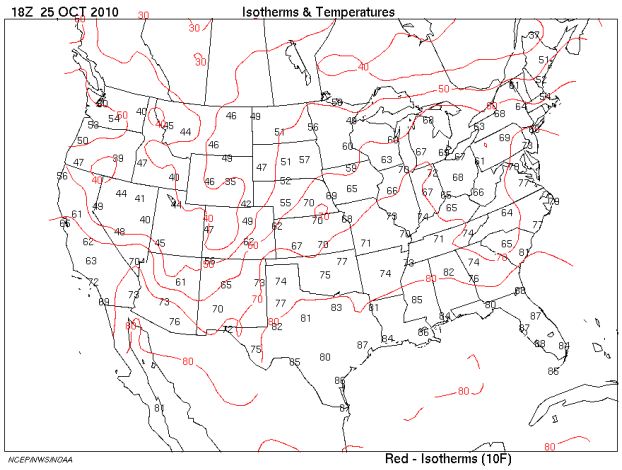
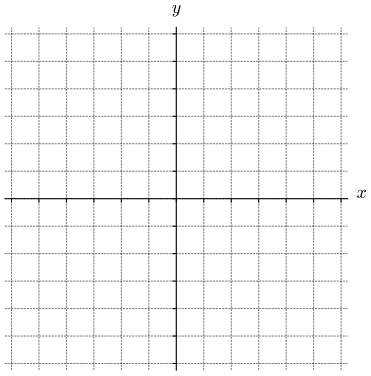


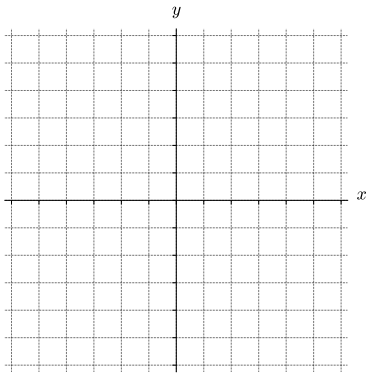
Figure 11.2.19 An isotherm map with temperatures at major cities (from NCEP/NWS/NOAA)

Activity 11.2.6 In this activity, you will make contour plots by hand and look at how the spacing of contours in your plot should give an idea about how different surface shapes can be distinguished.

(a) Let $f(x, y) = x^2 + y^2$. Draw the level curves $f(x, y) = k$ for $k = 1$, $k = 2$, $k = 3$, and $k = 4$ on the axes below. Be sure to label the scale of the axes. Explain what the surface defined by f looks like.



(b) Let $g(x, y) = \sqrt{x^2 + y^2}$. Draw the level curves $g(x, y) = k$ for $k = 1$, $k = 2$, $k = 3$, and $k = 4$ on the axes below. Use the same scale on these axes as in the previous part. Explain what the surface defined by g looks like.

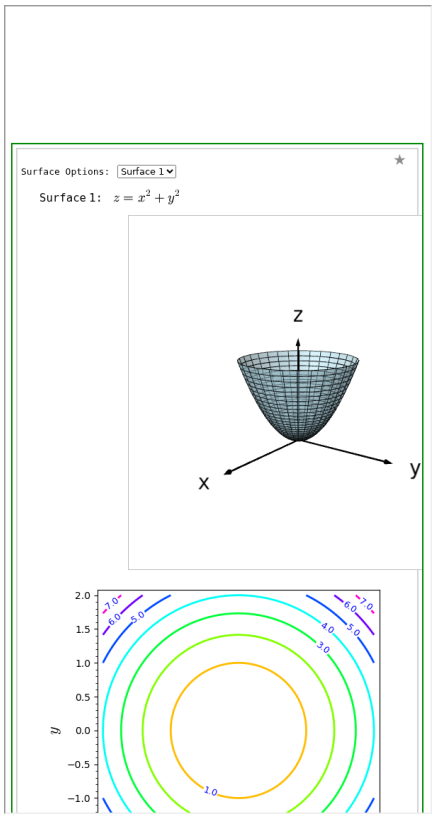


(c) Compare and contrast the graphs of f and g. How are they alike? How are they different? Use traces for each function to help answer these questions.

The traces and level curves of a function of two variables are curves in space. Traces are easier to parameterize as a curve in space because one input variable is fixed and the other input variable can act as the parameter. For example, the trace with $y = b$ of the surface given by $z = f(x, y)$ will be parameterized by $\vec{r}(t) = \langle t, b, f(t, b) \rangle$ because the z-coordinate can be expressed in terms of the fixed input value and the other input variable (as the parameter). Parameterizing contours, however, can be difficult and often relies on knowing the particular shape of the contours. Remember that it may not be possible to express a curve of the form $k = f(x, y)$ with one coordinate as a function of the other.

11.2.6 A gallery of functions

We end this section by considering a collection of functions and illustrating their surface graphs and contour plots. Many (but not all) of these surfaces will be familiar to you from Section 9.8, so now you will have an opportunity to see how the contour plots relate to the analyses you did in that section.



Standalone
In Context
Embed

Figure 11.2.20 A plot of various functions and their contour plots

Summary

- A function f of several variables is a rule that assigns a single number to an ordered collection of independent inputs. The domain of a function of several variables is the set of all inputs for which the function is defined.
- A trace of a function f of two independent variables x and y is a curve

of the form $z = f(x, c)$ (for an x-trace) or $z = f(c, y)$ (for a y-trace), where c is a constant. A trace tells us how the function depends on a single independent variable if we treat the other independent variable as a constant.

- A level curve of a function f of two independent variables x and y is a curve of the form $k = f(x, y)$, where k is a constant. A level curve describes the set of inputs that lead to a specific output of the function. A level curve is also called a contour.

11.2.7 Exercises

1. Evaluate the function at the specified points.
$f(x, y) = x + yx^2, (1, -2), (1, 2), (-1, 3)$
At $(1, -2)$: _____
At $(1, 2)$: _____
At $(-1, 3)$: _____

2. Suppose $f(x, y) = \left(x^2 - y^2, 2xy, x^3\right)$. Compute the following values:

$f(-2, 5) =$ _____
$f(5, -2) =$ _____
$f(0, 6) =$ _____
$f(t^2, -2t) =$ _____
$f(s + t, s - t) =$ _____

3. Suppose $f(x, y) = xy^2 - 1$. Compute the following values:

$f(0, -1) =$ _____
$f(-1, 0) =$ _____
$f(0, 0) =$ _____
$f(-2, -1) =$ _____
$f(t, 6t) =$ _____
$f(uv, u - v) =$ _____

4. Suppose $f(x, y) = xy + 1$. Compute the following values:

$f(x + 2y, x - 2y) =$ _____
$f(xy, 6x^2y^3) =$ _____

5. Evaluate the function at the specified points.
$f(x, y) = y + xy^4, (-3, -1), (-2, 1), (-5, -5)$
At $(-3, -1)$: _____
At $(-2, 1)$: _____
At $(-5, -5)$: _____

6. The domain of the function $f(x, y) = \sqrt{x} + \sqrt{y}$ is _____

7. Match the functions with the graphs of their domains.

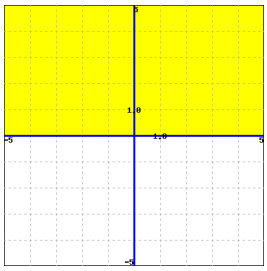
(a) $f(x, y) = \sqrt{x^2y^5}$

(b) $f(x, y) = \ln(3x - 3y)$

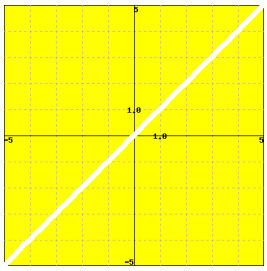
(c) $f(x, y) = 3x - 3y$

(d) $f(x, y) = e^{\frac{1}{3x-3y}}$

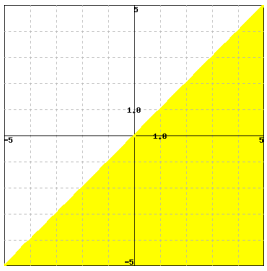
A.



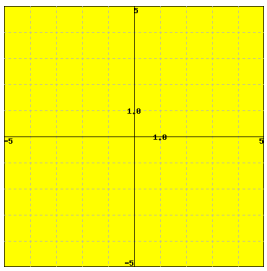
B.



C.



D.



8. Match the functions with the graphs of their domains.

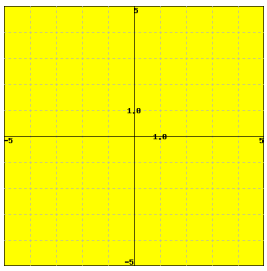
(a) $f(x, y) = 2y - 3x$

(b) $f(x, y) = \sqrt{x^3 y^4}$

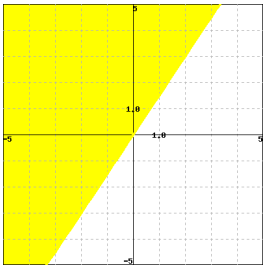
(c) $f(x, y) = e^{\frac{1}{2y-3x}}$

(d) $f(x, y) = \ln(2y - 3x)$

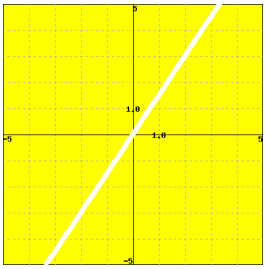
A.



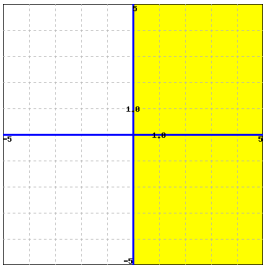
B.



C.



D.



9. Let $g(x, y) = \cos(2x + 4y)$.

1. Evaluate $g(1, -2)$.

Answer: $g(1, -2) =$ _____

2. What is the range of $g(x, y)$?

Answer (in interval notation): _____

10. Let $F(x, y) = 1 + \sqrt{4 - y^2}$.

1. Evaluate $F(3, 1)$.

Answer: $F(3, 1) =$ _____

2. What is the range of $F(x, y)$?

Answer (in interval notation): _____

11. The balance, B, in dollars, in a bank account depends on the amount deposited, A dollars, the annual interest rate, r %, and the time, t, in months since the deposit, so $B = f(A, r, t)$.

(a) Is f an increasing or decreasing function of:

A? (☐ increasing ☐ decreasing ☐ neither increasing nor decreasing ☐ somtimes increasing and sometimes decreasing)

r? (☐ increasing ☐ decreasing ☐ neither increasing nor decreasing ☐ somtimes increasing and sometimes decreasing)

t? (☐ increasing ☐ decreasing ☐ neither increasing nor decreasing ☐ somtimes increasing and sometimes decreasing)

(b) Interpret the statement $f(1250, 3.25, 32) \approx 2974$ by writing a sentence, including units. Then use your sentence to complete the following:

The units of 1250 are: (☐ percent ☐ months ☐ dollars ☐ dollars/month ☐ percent/month)

The units of 3.25 are: (☐ percent ☐ months ☐ dollars ☐ dollars/month ☐ percent/month)

The units of 32 are: (☐ percent ☐ months ☐ dollars ☐ dollars/

month ☐ percent/month)

The units of 2974 are: (☐ percent ☐ months ☐ dollars ☐ dollars/month ☐ percent/month) .

(Note because this is all multiple choice, you will not see which parts of the problem are correct.)

12. Sketch a contour diagram of each function. Then, decide whether its contours are predominantly lines, parabolas, ellipses, or hyperbolas.

(a) $z = x^2 + 2y^2$

(b) $z = y - 3x^2$

(c) $z = x^2 - 3y^2$

(d) $z = -3x^2$

13. Match the surfaces with the verbal description of the level curves by placing the letter of the verbal description to the left of the number of the surface.

(a) $z = \sqrt{(x^2 + y^2)}$

(b) $z = xy$

(c) $z = \sqrt{(25 - x^2 - y^2)}$

(d) $z = x^2 + y^2$

(e) $z = 2x + 3y$

(f) $z = 2x^2 + 3y^2$

(g) $z = \dfrac{1}{x - 1}$

A. two straight lines and a collection of hyperbolas

B. a collection of equally spaced concentric circles

C. a collection of unequally spaced parallel lines

D. a collection of equally spaced parallel lines

E. a collection of concentric ellipses

F. a collection of unequally spaced concentric circles

14. For each surface, decide whether it could be a bowl, a plate, or neither. Consider a plate to be any fairly flat surface and a bowl to be anything that could hold water, assuming the positive z-axis is up.

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

(b) $z = x^2 + y^2$

(c) $z = -\sqrt{4 - x^2 - y^2}$

(d) $z = 5$

(e) $x + y + z = 4$

15. Suppose $z = f(x, y) = (x^2 + 4)\sin(y) + xy^2$. By setting one variable equal to a constant (either $x = c, y = c$, or $z = c$), answer the following questions.

The intersection of the graph of f with the plane (☐ y = c ☐ z = c ☐ x = c) is a sine curve if $c =$ _____.

The intersection of the graph of f with the plane (☐ y = c ☐ x = c ☐ z = c) is a line if $c =$ _____.

The intersection of the graph of f with the plane (☐ x = c ☐ z = c ☐ y = c) is a parabola if $c =$ _____.

16. The temperature adjusted for wind-chill, w, is a temperature which tells you how cold it feels, as a result of the combination of wind and temperature [see wind-chill]. See the table below, which gives temperature adjusted for wind-chill, w, as a function of temperature T and wind speed s.

	$T = 35$	$T = 30$	$T = 25$	$T = 20$	$T = 15$	$T = 10$	$T = 5$	$T = 0$
$s = 5$	31	25	19	13	7	1	-5	-11
$s = 10$	27	21	15	9	3	-4	-10	-16
$s = 15$	25	19	13	6	0	-7	-13	-19
$s = 20$	24	17	11	4	-2	-9	-15	-22
$s = 25$	23	16	9	3	-4	-11	-17	-24

Use this table to make tables of the temperature adjusted for wind-chill (w) as a function of temperature for wind speeds 15 and 25 mph:

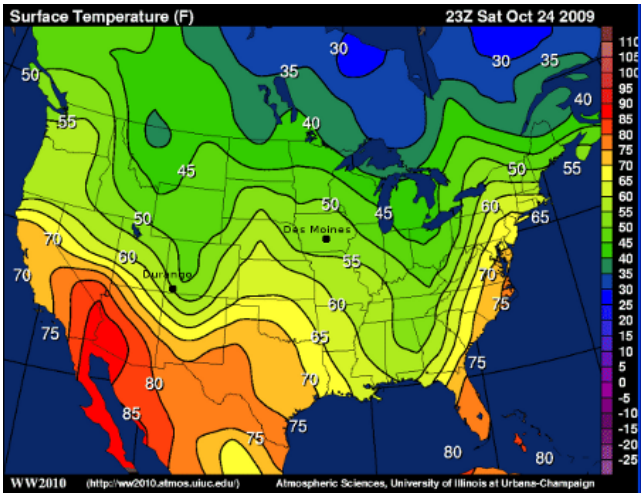
$s = 15$:

independent variable = ___ =	_____	_____	_____	_____	_____	_____	_____	_____
dependent variable = ___ =	_____	_____	_____	_____	_____	_____	_____	_____

$s = 25$:

independent variable = ___ =	_____	_____	_____	_____	_____	_____	_____	_____
dependent variable = ___ =	_____	_____	_____	_____	_____	_____	_____	_____

17.



(a) Describe a possible graph of the temperature, T, on an east-west line through Durango, Colorado, if the origin is at Durango, the positive

x-axis corresponds to east of Durango, and the vertical z-axis is the temperature. For negative x within a few hundred miles of the origin, the sign of $T'(x)$ is (☐ Positive ☐ Negative ☐ Zero) , while for positive x a few hundred miles from the origin, the sign of $T'(x)$ is (☐ Zero ☐ Negative ☐ Positive) . We estimate that $T(0) =$ (☐ 51 ☐ 57 ☐ 58 ☐ 59 ☐ 50 ☐ 53 ☐ 52 ☐ 60 ☐ 55 ☐ 56 ☐ 54) degrees Fahrenheit.

(b) Describe a possible graph of the temperature, T, on a north-south line through Des Moines, Iowa, if the origin is at Des Moines, the positive y-axis corresponds to north of Des Moines, and the vertical z-axis is the temperature. For negative y within a few hundred miles of the origin, the sign of $T'(y)$ is (☐ Positive ☐ Negative ☐ Zero) , while for positive y a few hundred miles from the origin, the sign of $T'(y)$ is (☐ Zero ☐ Positive ☐ Negative) . We estimate that $T(0) =$ (☐ 56 ☐ 55 ☐ 60 ☐ 54 ☐ 50 ☐ 59 ☐ 58 ☐ 57 ☐ 51 ☐ 52 ☐ 53) degrees Fahrenheit.

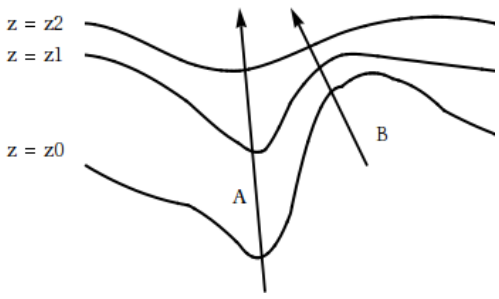
18. Find an equation for the contour of $f(x, y) = 2x^2y + 6x + 15$ that goes through the point $(4, 2)$.

Equation: _____

19. Write the level surface $2 = \frac{2x^2+5y}{z}$ as the graph of a function $f(x, y)$.

$f(x, y) =$ _____

20. The figure below shows a contour map of a hill with two paths, A and B. The heights on the contours are $z = z_0 = 100$, $z = z_1 = 175$, and $z = z_2 = 250$.



(a) On which path, A or B, will you have to climb more steeply? (☐ A ☐ B)

(b) On which path, A or B, will you probably have a better view of the surrounding countryside? (Assuming trees do not block your view.) (☐ A ☐ B)

(c) Alongside which path is there more likely to be a stream? (☐ A ☐ B)

21. On a piece of paper, sketch each of the following surfaces:

(i) $4 = x + 2y + z$

(ii) $z = 4$

Use your graphs to fill in the following descriptions of cross-sections of the surfaces.

(a) For ***(i)*** $(4 = x + 2y + z)$:

Cross sections with x fixed give (☐ a horizontal line in a plane parallel to the yz-plane ☐ an empty set, or one or two vertical lines in a plane parallel to the yz-plane ☐ a vertical line in a plane parallel to the yz-plane ☐ two vertical lines in a plane parallel to the yz-plane ☐ a line in a plane parallel to the yz-plane ☐ an upward opening parabola in a plane parallel to the yz-plane ☐ a downward opening parabola in a plane parallel to

the yz-plane ☐ a circle centered on the origin in a plane parallel to the yz-plane ☐ a point in a plane parallel to the yz-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the yz-plane)

Cross sections with y fixed give (☐ a horizontal line in a plane parallel to the xz-plane ☐ an empty set, or one or two vertical lines in a plane parallel to the xz-plane ☐ a vertical line in a plane parallel to the xz-plane ☐ two vertical lines in a plane parallel to the xz-plane ☐ a line in a plane parallel to the xz-plane ☐ an upward opening parabola in a plane parallel to the xz-plane ☐ a downward opening parabola in a plane parallel to the xz-plane ☐ a circle centered on the origin in a plane parallel the xz-plane ☐ a point in a plane parallel to the xz-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the xz-plane)

Cross sections with z fixed give (☐ a horizontal line in a plane parallel to the xy-plane ☐ two horizontal lines in a plane parallel to the xy-plane ☐ an empty set, or one or two horizontal lines in a plane parallel to the xy-plane ☐ a vertical line in a plane parallel to the xy-plane ☐ two vertical lines in a plane parallel to the xy-plane ☐ a line in a plane parallel to the xy-plane ☐ an upward opening parabola in a plane parallel to the xy-plane ☐ a downward opening parabola in a plane parallel to the xy-plane ☐ a circle centered on the origin in a plane parallel to the xy-plane ☐ a point in a plane parallel to the xy-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the xy-plane ☐ an empty set, or a plane parallel to the xy-plane)

(b) For ***(ii)*** ($z = 4$):

Cross sections with x fixed give (☐ a horizontal line in a plane parallel to the yz-plane ☐ an empty set, or one or two vertical lines in a plane parallel to the yz-plane ☐ a vertical line in a plane parallel to the yz-plane ☐ two vertical lines in a plane parallel to the yz-plane ☐ a line in a plane parallel to the yz-plane ☐ an upward opening parabola in a plane parallel to the yz-plane ☐ a downward opening parabola in a plane parallel to the yz-plane ☐ a circle centered on the origin in a plane parallel to the yz-plane ☐ a point in a plane parallel to the yz-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the yz-plane)

Cross sections with y fixed give (☐ a horizontal line in a plane parallel to the xz-plane ☐ an empty set, or one or two vertical lines in a plane parallel to the xz-plane ☐ a vertical line in a plane parallel to the xz-plane ☐ two vertical lines in a plane parallel to the xz-plane ☐ a line in a plane parallel to the xz-plane ☐ an upward opening parabola in a plane parallel to the xz-plane ☐ a downward opening parabola in a plane parallel to the xz-plane ☐ a circle centered on the origin in a plane parallel the xz-plane ☐ a point in a plane parallel to the xz-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the xz-plane)

Cross sections with z fixed give (☐ a horizontal line in a plane parallel to the xy-plane ☐ two horizontal lines in a plane parallel to the xy-plane ☐ an empty set, or one or two horizontal lines in a plane parallel to the xy-plane ☐ a vertical line in a plane parallel to the xy-plane ☐ two vertical lines in a plane parallel to the xy-plane ☐ a line in a plane parallel to the xy-plane ☐ an upward opening parabola in a plane parallel to the xy-plane ☐ a downward opening parabola in a plane parallel to the xy-plane ☐ a circle centered on the origin in a plane parallel to the xy-plane ☐ a point in a plane parallel to the xy-plane ☐ an empty set, or a circle centered on the origin in a plane parallel to the xy-plane ☐ an empty set, or a plane parallel to the xy-plane)

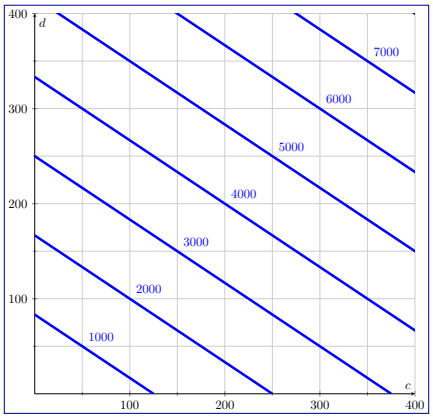
22. A car rental company charges a one-time application fee of 30 dollars, 60 dollars per day, and 11 cents per mile for its cars.

(a) Write a formula for the cost, C, of renting a car as a function of the number of days, d, and the number of miles driven, m.

$C =$ _____

(b) If $C = f(d, m)$, then $f(5, 530) =$ _____

23. A store sells CDs at one price and DVDs at another price. The figure below shows the revenue (in dollars) of the music store as a function of the number, c, of CDs and the number, d, of DVDs that it sells. The values of the revenue are shown on each line.



(Hint: for this problem there are many possible ways to estimate the requisite values; you should be able to find information from the figure that allows you to give an answer that is essentially exact.)

a. What is the price of a CD? _____ dollars

b. What is the price of a DVD? _____ dollars

24. Consider the concentration, C, (in mg/liter) of a drug in the blood as a function of the amount of drug given, x, and the time since injection, t. For $0 \le x \le 6$ mg and $t \ge 0$ hours, we have

$$C = f(x, t) = 20te^{-(6-x)t}$$

$f(1, 2) =$ _____

Give a practical interpretation of your answer: $f(1, 2)$ is

◯ the amount of a 2 mg dose in the blood 1 hours after injection.

◯ the concentration of a 1 mg dose in the blood 2 hours after injection.

◯ the change in concentration of a 1 mg dose in the blood 2 hours after injection.

◯ the amount of a 1 mg dose in the blood 2 hours after injection.

◯ the change in concentration of a 2 mg dose in the blood 1 hours after injection.

◯ the concentration of a 2 mg dose in the blood 1 hours after injection.

25. A manufacturer sells aardvark masks at a price of \$270 per mask and butterfly masks at a price of \$520 per mask. A quantity of a aardvark masks and b butterfly masks is sold at a total cost of \$600 to the manufacturer.

(a) Express the manufacturer's profit, P, as a function of a and b.
$P(a, b) =$ _____ dollars.
(b) The curves of constant profit in the ab-plane are

◯ hyperbolas

◯ ellipses

◯ circles

◯ lines

◯ parabolas

26. Consider the concentration, C, in mg per liter (L), of a drug in the blood as a function of x, the amount, in mg, of the drug given and t, the time in hours since the injection. For $0 \le x \le 4$ and $t \ge 0$, we have $C = f(x, t) = te^{-t(5-x)}$.

Graph the following two single variable functions on a separate page, being sure that you can explain their significance in terms of drug concentration.

(a) $f(4, t)$
(b) $f(x, 2.5)$
Using your graph in ***(a)***, where is $f(4, t)$
a maximum? $t =$ _____
a minimum? $t =$ _____
Using your graph in ***(b)***, where is $f(x, 2.5)$
a maximum? $x =$ _____
a minimum? $x =$ _____

27. By setting one variable constant, find a plane that intersects the graph of $z = 2y^2 - 4x^2 + 2$ in a:

(a) Parabola opening upward: the plane _____ = _____

(Give your answer by specifying the variable in the first answer blank and a value for it in the second.)

(b) Parabola opening downward: the plane _____ = _____

(Give your answer by specifying the variable in the first answer blank and a value for it in the second.)

(c) Pair of intersecting straight lines: the plane _____ = _____

(Give your answer by specifying the variable in the first answer blank and a value for it in the second.)

28. The Ideal Gas Law, $PV = RT$, relates the pressure (P, in pascals), temperature (T, in Kelvin), and volume (V, in cubic meters) of 1 mole of a gas ($R = 8.314 \frac{\text{J}}{\text{mol K}}$ is the universal gas constant), and describes the behavior of gases that do not liquefy easily, such as oxygen and hydrogen. We can solve the ideal gas law for the volume and hence treat the volume as a function of the pressure and temperature:

$$V(P, T) = \frac{8.314T}{P}.$$

(a) Explain in detail what the trace of V with $P = 1000$ tells us about a key relationship between two quantities.

(b) Explain in detail what the trace of V with $T = 5$ tells us.

(c) Explain in detail what the level curve $V = 0.5$ tells us.

(d) Use 2 or three additional traces in each direction to make a rough sketch of the surface over the domain of V where P and T are each nonnegative. Write at least one sentence that describes the way the surface looks.

(e) Based on all your work above, write a couple of sentences that describe the effects that temperature and pressure have on volume.

29. When people buy a large ticket item like a car or a house, they often take out a loan to make the purchase. The loan is paid back in monthly installments until the entire amount of the loan, plus interest, is paid. The monthly payment that the borrower has to make depends on the amount P of money borrowed (called the principal), the duration t of the loan in years, and the interest rate r. For example, if we borrow \$18,000 to buy a car, the monthly payment M that we need to make to pay off the loan is given by the formula

$$M(r,t) = \frac{1500r}{1 - \frac{1}{\left(1+\frac{r}{12}\right)^{12t}}}.$$

(a) Find the monthly payments on this loan if the interest rate is 6% and the duration of the loan is 5 years.

(b) Create a table of values that illustrates the trace of M with r fixed at 5%. Use yearly values of t from 2 to 6. Round payments to the nearest penny. Explain in detail in words what this trace tells us about M.

(c) Create a table of values that illustrates the trace of M with t fixed at 3 years. Use rates from 3% to 11% in increments of 2%. Round payments to the nearest penny. Explain in detail what this trace tells us about M.

(d) Consider the combinations of interest rates and durations of loans that result in a monthly payment of \$200. Solve the equation $M(r,t) = 200$ for t to write the duration of the loan in terms of the interest rate. Graph this level curve and explain as best you can the relationship between t and r.

30. Consider the function h defined by $h(x,y) = 8 - \sqrt{4 - x^2 - y^2}$.

(a) What is the domain of h? (Hint: describe a set of ordered pairs in the plane by explaining their relationship relative to a key circle.)

(b) The *range* of a function is the set of all outputs the function generates. Given that the range of the square root function $g(t) = \sqrt{t}$ is the set of all nonnegative real numbers, what do you think is the range of h? Why?

(c) Choose 4 different values from the range of h and plot the corresponding level curves in the plane. What is the shape of a typical level curve?

(d) Choose 5 different values of x (including at least one negative value and zero), and sketch the corresponding traces of the function h.

(e) Choose 5 different values of y (including at least one negative value and zero), and sketch the corresponding traces of the function h.

(f) Sketch an overall picture of the surface generated by h and write at least one sentence to describe how the surface appears visually. Does the surface remind you of a familiar physical structure in nature?

11.3 Limits

Motivating Questions

- What do we mean by the limit of a function f of two variables at a point (a, b)?
- What techniques can we use to show that a function of two variables does not have a limit at a point (a, b)?
- What does it mean for a function f of two variables to be continuous at a point (a, b)?

11.3.1 Introduction

In this section, we want to study limits of functions of several variables and will primarily look at limits of functions of two variables. In single variable calculus, we studied the idea of a limit, which turned out to be a critical concept which served as the critical third step in the Classic Calculus Approach to understanding the derivative and the definite integral. In this section, we will study how the concept of limit for functions of two variables is similar to what we encountered for functions of a single variable. We will use the notion of the limit of a function of several variables as the last step in understanding and defining important concepts related to differentiability later in this chapter.

We did not need to generalize our ideas related to limits when we looked at vector-valued functions of one variable in Chapter 10 because we applied our one-variable limits componentwise. That is, the limit of the function $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ is evaluated as three separate one-variable limits.

We begin by reviewing the idea behind a limit of a function of one variable. We say that a function f has a limit L as x approaches a provided that we can make the values $f(x)$ as close to L as we like by taking x sufficiently close (but not equal) to a. We denote this behavior by writing

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

Remember that you may need to look at the outputs of the function f as you approach the input $x = a$ from the right and left separately in order to ensure that the limit exists.

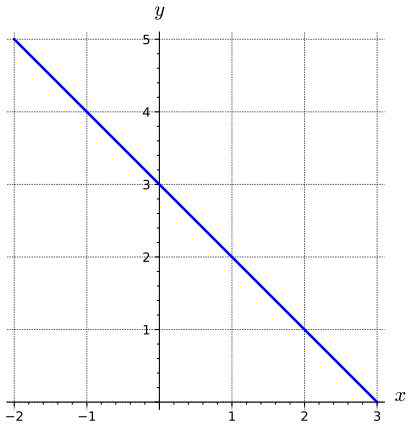
Preview Activity 11.3.1 In this Preview Activity, we will recall and work with several ideas related to limits of single-variable functions by working with tables, graphs, and the algebraic forms of these functions.

(a) Let f defined by $f(x) = 3 - x$. Complete the table below.

x	-0.2	-0.1	-0.01	0.01	0.1	0.2
$f(x)$						

(b) Write a sentence about what the table in part (a) suggests regarding $\lim_{x \to 0} f(x)$.

(c) Write a couple of sentences to explain how your values in the table in part (a) and $\lim_{x \to 0} f(x)$ are demonstrated in the graph below.



(d) Let $g(x) = \dfrac{x}{|x|}$. Fill in the blanks of the table below with the appropriate outputs of g for values near $x = 0$. Note that g is not defined at $x = 0$.

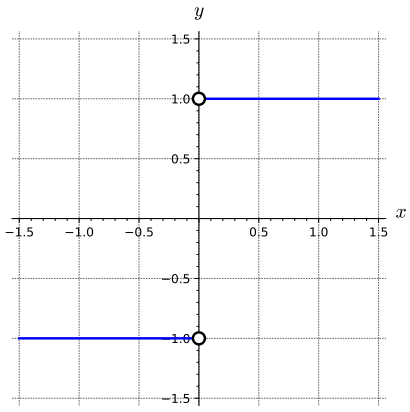
x	-0.1	-0.01	-0.001	0.001	0.01	0.1
$g(x)$						

(e) Write a few sentences about what the values in the table in part (d) suggests about $\lim_{x \to 0} g(x)$.

(f) Write a couple of sentences to explain how your values in the table in part (d) and

$$\lim_{x \to 0} g(x)$$

are demonstrated in the plot below.



(g) Let $h(x) = \dfrac{1 - x^2}{x + 1}$. At what point(s) is h undefined?

(h) Show that the limit as x goes to -1 of $h(x)$ exists. Write a sentence or two about why $h(-1)$ does not exist but the limit as x goes to -1 does exist.

(i) Graph $h(x)$ for a region around $x = -1$ and explain how this graph relates to your answer to the previous two parts.

In Preview Activity 11.3.1, we used the notion of limit from single variable calculus and saw several different outcomes for limits in terms of tables and

graphs. Specifically, when the limit $\lim_{x \to a} f(x)$ exists, its value will be the value that the output of f approaches as the x-value inputs get closer to a. We must look at whether the output of f approaches the same value as the input values approach $x = a$ from different directions. Tables of values for the function can give evidence that a limit will either exist or not but will not have enough evidence to *show* a limit exists. Similarly, we can use graphs to visually represent the behavior of the output values as we approach a particular input from different directions.

11.3.2 Limits of Functions of Two Variables

In this section, we will look at how the concept of limits on single variable functions can be expanded to work with functions of two variables. We will focus on two variables here because tools like tables and graphs readily make sense for a function of the form $f(x, y)$ but become more difficult to use for functions of more than two variables. All of the concepts of limits and continuity can be expanded to functions of any number of variables but discussion of the details of this type of generalization will be deferred to a later mathematics course.

The biggest difference between limits of single-variable functions and multivariable functions is the number of ways we can approach a particular input. For a single-variable function, there are only two directions in which we can approach an input like $x = a$, which we call approaching from the left and approaching from the right. We illustrate this in Figure 11.3.1. When considering a limit of a single-variable function, we can evaluate and compare the output of a function in these two directions. If the output of the function does not agree as you approach from the left and right, then the limit does not exist, as we saw in part 11.3.1.e.

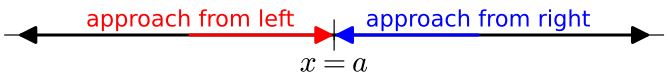


Figure 11.3.1 The two ways to approach an input in 1D

In contrast to the single variable case, there are ***many*** different ways to approach a point (a, b), when considering this point as the input for a function $f(x, y)$. In Figure 11.3.2, we show seven different ways to approach the point (a, b). When we look at how to get close to an input in two dimensions, we must think about more than just the straight-line directions. For example, we have to consider curved paths. In the next example, we will explore how we can think about the various representations of a multivariable function and what it would mean to examine the output of a multivariable function as the inputs get close to a particular value.

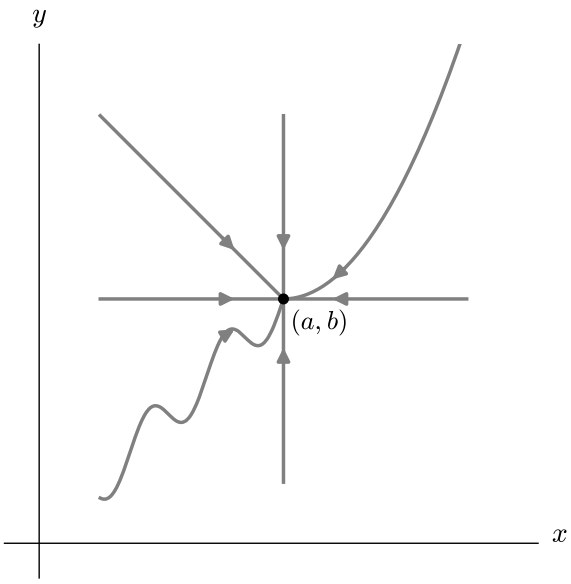


Figure 11.3.2 Seven different ways to approach a point in two dimensions

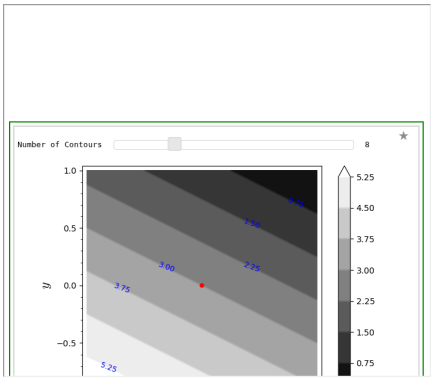
Example 11.3.3 In this example, we will use tables and graphs to examine a function of two variables just as you did in Preview Activity 11.3.1 for functions of one variable. Here we consider the output of the function $f(x, y) = 3-x-2y$ approaches as the inputs get close to $(0, 0)$. Table 11.3.4 shows a spreadsheet of output values for f near the input $(0, 0)$. We have purposefully omitted the value of $f(0, 0)$ because we need to look at what is happening to f *near* $(0, 0)$, not what is happening at $(0, 0)$.

Table 11.3.4 Values of $f(x, y) = 3 - x - 2y$

$x\backslash y$	-1.0	-0.1	0.0	0.1	1.0
-1.0	6	4.2	4	3.8	2
-0.1	5.1	3.3	3.1	2.9	1.1
0.0	5	3.2		2.8	1
0.1	4.9	3.1	2.9	2.7	0.9
1.0	4	2.2	2.0	1.8	0

Table 11.3.4 shows that as the inputs (x, y) get closer to $(0, 0)$, the output of $f(x, y)$ gets closer to 3. We can look at some aspect of directionality using a table, like how the row corresponding to $x = 0$ shows values of f for inputs on the y-axis and the values on the diagonal (from top left to bottom right) shows values of f for inputs with equal x- and y-coordinates.

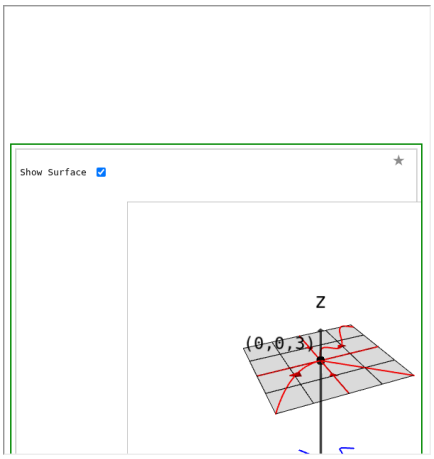
Visual representations like contour graphs and surface plots will give more information with respect to the different directions we can approach $(0, 0)$. Figure 11.3.5 shows a contour plot of $f(x, y) = 3 - x - 2y$ where you can use the slider to change the number of contours used in the plot. The input point $(0, 0)$ is highlighted in red and you can see that regardless of how many contours are displayed, the output value of f seems to approach 3 near the red point. Just as with the table, you can get some idea on the directionality of how the outputs will approach 3, but because the contour plot does not give information about what is happening between contours, you are left to mentally fill in this information.



Standalone
In Context
Embed

Figure 11.3.5 A contour plot of $f(x, y) = 3 - x - 2y$ with the point $(0, 0)$ highlighted in red

A plot of the surface $z = f(x, y) = 3 - x - 2y$ will give more information about what is happening in different directions as the inputs approach $(0, 0)$. The drawback to using surface plots is that they are difficult to create by hand. Even computer-generated plots of surfaces can be difficult to interpret and are prone to errors because of imprecise numerical calculations. In our case, the plot of $z = f(x, y) = 3 - x - 2y$ is a plane, which we have experience with from earlier in the course. We can see in Figure 11.3.6 that if we approach the input point $(0, 0)$ in any direction or along any path, then the output of f (as shown by the z-coordinate on the plot) gets closer to 3.



Standalone
In Context
Embed

Figure 11.3.6 A plot of $z = f(x, y) = 3 - x - 2y$ with several paths approaching the input $(0, 0)$

If you were not convinced that the output of f approaches 3 as (x, y) approaches $(0, 0)$, you could look at approaching the origin along a particular path. For instance, we could look to simplify our limit to look at approaching the origin along the path given by $y = 2x$. This means that we don't really need to examine our function's output at all points around the origin. Instead, we consider only points of the form $(x, 2x)$. The limit expression becomes a one-dimensional problem in this case because along the path given by $y = 2x$, $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ becomes $\lim\limits_{(x,2x)\to(0,0)} f(x, 2x)$, which is the same as $\lim\limits_{x\to 0} 3 - x - 2(2x)$. Note that the limit along this path is now of the type you evaluated in your single-variable calculus course. In fact, you can evaluate $\lim\limits_{x\to 0} 3 - x - 2(2x)$ to get 3. It would be possible to use other paths such as $x = \sin(y)$ or even a parameterized path of the form $\vec{r}(t) = \langle t\cos(t), t\sin(t)\rangle$

to consider a two-variable limit. However, we can *never* test *all* possible paths that approach this limit point. Thus, while the idea of reducing a two-variable limit to a one-variable limit along a particular path can be helpful, we will not be able to exhaust all possible ways to approach our limit point and show a limit exists.

In all of our representations, we see evidence that if the input of our function, (x, y), approaches $(0, 0)$, then the output of our function goes to 3, regardless of the direction or way we approach $(0, 0)$. We generalize this idea with the following definition. ▲

Definition 11.3.7 Given a function $f = f(x, y)$, we say that f **has limit** L **as** (x, y) **approaches** (a, b) provided that we can make $f(x, y)$ as close to L as we like by taking (x, y) sufficiently close (but not equal) to (a, b). We write

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

◆

Our work in Example 11.3.3 showed that the limit of $f(x, y) = 3 - x - 2y$ as (x, y) approaches $(0, 0)$ is 3. Symbolically, this looks like

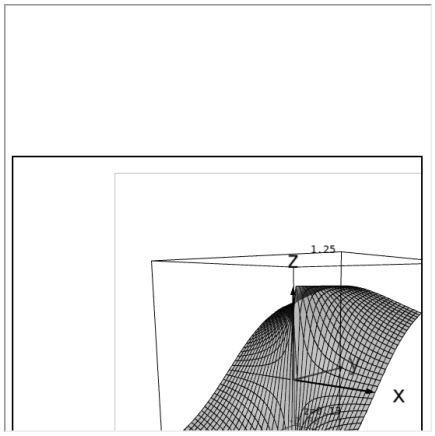
$$\lim_{(x,y)\to(0,0)} 3 - x - 2y = 3$$

To investigate the limit of a single variable function, $\lim_{x\to a} f(x)$, we often consider the behavior of f as x approaches a from the right and from the left. Similarly, we may investigate limits of two-variable functions, $\lim_{(x,y)\to(a,b)} f(x, y)$ by considering the behavior of f as (x, y) approaches (a, b) along a particular path. Remember that the multivariable function situation is more complicated because there are infinitely many ways in which (x, y) may approach (a, b), so we will not be able to check every possible path. In the next activity, we see how it is important to consider a variety of those paths in investigating whether or not a limit exists.

Activity 11.3.2 Consider the function f, defined by

$$f(x, y) = \frac{y}{\sqrt{x^2 + y^2}}.$$

The graph of this function is shown below.



Standalone
In Context
Embed

(a) Is f defined at the point $(0, 0)$? What, if anything, does this say about whether f has a limit at the point $(0, 0)$?

(b) Values of f (to three decimal places) at several points close to $(0, 0)$ are shown below.

$x\backslash y$	-1.000	-0.100	0.000	0.100	1.000
-1.000	-0.707	—	0.000	—	0.707
-0.100	—	-0.707	0.000	0.707	—
0.000	-1.000	-1.000	—	1.000	1.000
0.100	—	-0.707	0.000	0.707	—
1.000	-0.707	—	0.000	—	0.707

Based on these calculations, state whether you think f has a limit at $(0,0)$ and give an argument supporting your statement.

Hint. The blank spaces in the table are there to help you see the patterns.

(c) Now we formalize your conjecture from the previous part by considering what happens if we restrict our attention to different paths. First, we look at f for points in the domain along the x-axis. That is, we consider what happens when $y = 0$. Write out what the output of f will when $y = 0$. In other words, what is $f(x, 0)$ equal to?

What is the behavior of $f(x, 0)$ as $x \to 0$? If we approach $(0, 0)$ by moving along the x-axis, what value do we find as the limit?

(d) What is the behavior of f along the line $y = x$ when $x > 0$. That is, what is the value of $f(x, x)$ when $x > 0$? If we approach $(0, 0)$ by moving along the line $y = x$ in the first quadrant, what value do we find as the limit?

Hint. Here you are considering the limit of $f(x, x)$ as $x \to 0^+$.

(e) In general, if $\lim\limits_{(x,y)\to(0,0)} f(x, y) = L$, then $f(x, y)$ approaches L as (x, y) approaches $(0, 0)$, regardless of the path we take in letting $(x, y) \to (0, 0)$. Based on the last two parts of this activity, can the limit

$$\lim_{(x,y)\to(0,0)} f(x, y)$$

exist? Write a sentence or two to explain your reasoning.

(f) Substitute $y = mx$ into $f(x, y)$ and simplify your result for $f(x, mx)$. Use your expression and limit rules to evaluate

$$\lim_{x\to 0} f(x, mx).$$

Use this limit to explain the the behavior of $f(x, y)$ as (x, y) approaches $(0, 0)$ along lines of the form $y = mx$ and how this observation reinforces your conclusion about the existence of $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ from the previous part of this activity.

As we have seen in Activity 11.3.2, if $f(x, y)$ has two different limits along two different paths as (x, y) approaches (a, b), then we can conclude that $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ does *not* exist. This is similar to the one-variable example $g(x) = x/|x|$ as shown in Figure 11.3.8; $\lim\limits_{x\to 0} g(x)$ does not exist because we see different limits as x approaches 0 from the left and the right.

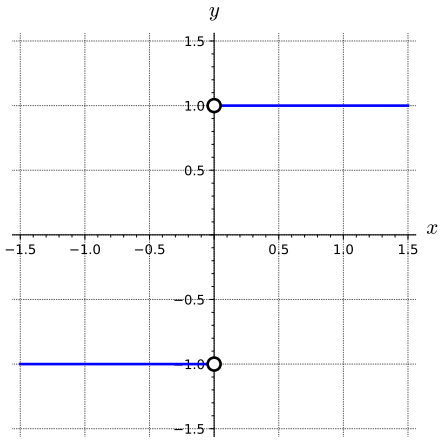


Figure 11.3.8 A plot of $g(x) = \frac{x}{|x|}$ on right

As a general rule, we have the following fact:

Limits along different paths.

If $f(x,y)$ has two different limits as (x,y) approaches (a,b) along two different paths, then $\lim_{(x,y)\to(a,b)} f(x,y)$ does not exist.

As the next activity shows, studying the limit of a two-variable function f by considering the behavior of f along various paths can require subtle insights.

Activity 11.3.3 Define the function g by

$$g(x,y) = \frac{x^2 y}{x^4 + y^2}.$$

We will investigate the limit $\lim_{(x,y)\to(0,0)} g(x,y)$.

(a) What is the behavior of g on the x-axis? That is, what is $g(x,0)$ and what is the limit of g as (x,y) approaches $(0,0)$ along the x-axis?

(b) What is the behavior of g on the y-axis? That is, what is $g(0,y)$ and what is the limit of g as (x,y) approaches $(0,0)$ along the y-axis?

(c) What is the behavior of g on the line $y = mx$? That is, what is $g(x,mx)$ and what is the limit of g as (x,y) approaches $(0,0)$ along the line $y = mx$?

(d) Based on what you have seen so far, do you think $\lim_{(x,y)\to(0,0)} g(x,y)$ exists? If so, what do you think its value is?

(e) Now consider the behavior of g on the parabola $y = x^2$? What is $g(x,x^2)$ and what is the limit of g as (x,y) approaches $(0,0)$ along this parabola?

(f) State whether you think the limit $\lim_{(x,y)\to(0,0)} g(x,y)$ exists or not and provide a justification of your statement.

This activity shows that we need to be careful when studying the limit of a two-variable functions by considering its behavior along different paths. If we find two different paths that result in two different limits, then we conclude that the limit does not exist. However, we can *never* conclude that the limit

of a function *does* exist *only* by considering its behavior along different paths.

Generally speaking, demonstrating that the limit $\lim_{(x,y)\to(a,b)} f(x,y)$ exists requires a more careful argument.

Example 11.3.9 We consider the function f defined by

$$f(x,y) = \frac{x^2y^2}{x^2+y^2}.$$

We will try to determine whether $\lim_{(x,y)\to(0,0)} f(x,y)$ exists or not.

Note that if either x or y is 0, then $f(x,y) = 0$. Therefore, if f has a limit at $(0,0)$, it must be 0. We will therefore argue that

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0$$

by showing that we can make $f(x,y)$ as close to 0 as we wish by taking (x,y) sufficiently close (but not equal) to $(0,0)$. In the following arguments, we view x and y as being real numbers that are close, but not equal, to 0.

Since $0 \le x^2$, we have

$$y^2 \le x^2 + y^2,$$

which implies that

$$\frac{y^2}{x^2+y^2} \le 1.$$

Multiplying both sides by x^2 and observing that $f(x,y) \ge 0$ for all (x,y) gives

$$x^2 \ge x^2\left(\frac{y^2}{x^2+y^2}\right) = \frac{x^2y^2}{x^2+y^2} = f(x,y) \ge 0.$$

Thus, $0 \le f(x,y) \le x^2$. Since $x^2 \to 0$ as $x \to 0$, we can make $f(x,y)$ as close to 0 as we like by taking x sufficiently close to 0. For this example, it turns out that we don't even need to worry about making y close to 0. Therefore,

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2} = 0.$$

▲

In spite of the fact that these two most recent examples illustrate some of the complications that arise when studying limits of two-variable functions, many of the properties that are familiar from our study of single variable functions hold in precisely the same way.

Properties of Limits.

Let $f = f(x,y)$ and $g = g(x,y)$ be functions so that $\lim_{(x,y)\to(a,b)} f(x,y)$ and $\lim_{(x,y)\to(a,b)} g(x,y)$ both exist. Then

1. $\lim_{(x,y)\to(a,b)} x = a \quad$ and $\quad \lim_{(x,y)\to(a,b)} y = b$
2. $\lim_{(x,y)\to(a,b)} cf(x,y) = c\left(\lim_{(x,y)\to(a,b)} f(x,y)\right)$ for any scalar c
3. $\lim_{(x,y)\to(a,b)} [f(x,y) \pm g(x,y)] = \lim_{(x,y)\to(a,b)} f(x,y) \pm \lim_{(x,y)\to(a,b)} g(x,y)$

4. $\lim_{(x,y)\to(a,b)} [f(x,y)g(x,y)] = \left(\lim_{(x,y)\to(a,b)} f(x,y)\right)\left(\lim_{(x,y)\to(a,b)} g(x,y)\right)$

5. $\lim_{(x,y)\to(a,b)} \frac{f(x,y)}{g(x,y)} = \frac{\lim_{(x,y)\to(a,b)} f(x,y)}{\lim_{(x,y)\to(a,b)} g(x,y)}$ if $\lim_{(x,y)\to(a,b)} g(x,y) \neq 0$

We can use these properties and results from single variable calculus to verify that many limits exist. For example, these properties show that the function f defined by

$$f(x,y) = 3x^2y^3 + 2xy^2 - 3x + 1$$

has a limit at every point (a,b) and, moreover,

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b).$$

The reason for this is that polynomial functions of a single variable have limits at every point.

11.3.3 Using Polar Coordinates in Limits

In Activity 11.3.3, we saw how looking at all different types of paths that approach the input of interest is necessary to determine whether the limit exists. Additionally, Example 11.3.9 shows how a careful set of algebra and inequalities can be used to prove a limit does exist. When working with limits in single variable calculus, you often either used properties like those listed above or had to use other tools related to measuring the rates of change in different functions.

One tool that will work well for functions with two independent input variables is polar coordinates. If we convert our problem (both the function and the limit) to polar coordinates with the point of interest at the origin, then we can evaluate our limit in terms of r going to zero. Remember from Subsection 9.9.2 that polar coordinates allow us to separate the distance from the origin and the rotation around the origin. In other words, the r-coordinate measures how far we are from the origin (our limit point) and θ will describe the direction with which we are approaching the origin (our limit point). In Subsection 9.9.2, there are several examples of converting points and equations into polar coordinates so will refer you there to review as necessary.

Example 11.3.10 We want to understand the following limit, which we first saw in Activity 11.3.2, by using polar coordinates:

$$\lim_{(x,y)\to(0,0)} \frac{y}{\sqrt{x^2+y^2}}$$

First, we note that this function is well-suited to convert to polar coordinates by writing x and y in terms of r and θ. Substituting in $y = r\sin(\theta)$ and $r^2 = x^2 + y^2$ into our function we have

$$\frac{y}{\sqrt{x^2+y^2}} = \frac{r\sin(\theta)}{\sqrt{r^2}} = \sin(\theta).$$

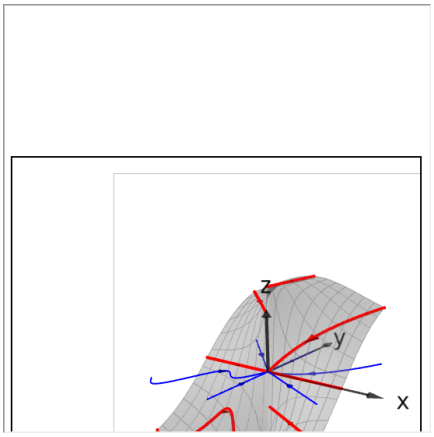
In polar coordinates, our expression simplifies to $\sin(\theta)$.

We now convert the limit expression from $(x,y) \to (0,0)$ to $r \to 0$. The

whole conversion of the limit becomes

$$\lim_{(x,y)\to(0,0)} \frac{y}{\sqrt{x^2+y^2}} = \lim_{r\to 0} \sin(\theta).$$

Evaluating the limit of $\sin(\theta)$ as r goes to 0 gives $\sin(\theta)$. The result of our algebraic limit calculation may seem surprising at first, because your prior experience with limits suggests that the limit of a function should either be a scalar or not exist. The conversion from rectangular to polar coordinates means we must interpret this result in terms of rectangular coordinates. Another way of saying what it means for the limit to be $\sin(\theta)$ is that our conclusion is that the limit has a dependence on the *direction* in which we approached the origin. Specifically, when approaching the origin along a line that makes an angle of θ with the positive x-axis, the output of the function approaches $\sin(\theta)$. In terms of rectangular coordinates, we say that the limit of $\frac{y}{\sqrt{x^2+y^2}}$ as $(x, y) \to (0, 0)$ does not exist because the limits along different paths going to the origin do not agree with each other.



Standalone
In Context
Embed

Figure 11.3.11 A plot of $z = f(x, y) = \frac{y}{\sqrt{x^2+y^2}}$ with a variety of paths approaching the origin

Figure 11.3.11 shows in blue a collection of paths in the xy-plane that approach the origin. Each of these paths is shown in red on the surface $z = f(x, y) = \frac{y}{\sqrt{x^2+y^2}}$ to demonstrate how the output of f, which is represented as the height of the surface, approaches different values as these paths each approach $(0, 0)$. Note that the limit along each of these paths exists, but the limit as $(x, y) \to (0, 0)$ does not exist because there are paths that yield different limits. This should give you a greater appreciation for a surface plot like we saw in Activity 11.3.2. Computer-generated plots will create a surface plot but it is important for you to identify features on that plot that will have behavior quite different than the smooth, continuous appearance. ▲

Activity 11.3.4 In this activity, we again consider the limit as (x, y) goes to $(0, 0)$ of $g(x, y) = \frac{x^2y^2}{x^2+y^2}$, which we previously saw in Example 11.3.9.

(a) Convert g to polar coordinates.

(b) Find the limit as $r \to 0$ of your expression in polar coordinates.

(c) Write a couple of sentences explaining how your answer to the previous part shows that the original limit is 0 and contrast with the previous example.

These examples may tempt you to conclude that polar coordinates will solve all problems related to finding limits of functions of two variables. However, that simplification misses the importance of how convenient it was to convert to polar coordinates in these functions. If we were examining a function that did not have a convenient conversion to polar coordinates or the algebra to simplify and evaluate the limit in polar coordinates was difficult, then polar coordinates will not be a great tool for evaluating limits. (Consider, for example, what would happen if you tried polar coordinates in Exercise 17.) Polar coordinates as a tool is presented here because it highlights how to geometrically separate position in two dimensions into a distance and a measure of directionality. These ideas recur in many places throughout this text and have already been used in our work on vectors. For instance, we often analyze the magnitude of vector measurements differently than we analyze direction aspects of vectors.

11.3.4 Continuity

Recall that a function f of a single variable x is said to be continuous at $x = a$ provided that the following three conditions are satisfied:

1. $f(a)$ exists,
2. $\lim_{x \to a} f(x)$ exists, and
3. $\lim_{x \to a} f(x) = f(a)$

Using our understanding of limits of multivariable functions, we can define continuity in the same way.

Definition 11.3.12 A function $f = f(x, y)$ is **continuous** at the point (a, b) provided that

1. f is defined at the point (a, b),
2. $\lim_{(x,y) \to (a,b)} f(x, y)$ exists, and
3. $\lim_{(x,y) \to (a,b)} f(x, y) = f(a, b)$.

◆

For instance, we have seen that the function f defined by $f(x, y) = 3x^2y^3 + 2xy^2 - 3x + 1$ is continuous at every point. And just as with single variable functions, continuity has certain properties that are based on the properties of limits.

Properties of continuity.

Let f and g be functions of two variables that are continuous at the point (a, b). Then

1. cf is continuous at (a, b) for any scalar c
2. $f + g$ is continuous at (a, b)
3. $f - g$ is continuous at (a, b)
4. fg is continuous at (a, b)
5. $\frac{f}{g}$ is continuous at (a, b) if $g(a, b) \neq 0$

Using these properties, we can apply results from single variable calculus to decide about continuity of multivariable functions. For example, the coordinate functions f and g defined by $f(x,y) = x$ and $g(x,y) = y$ are continuous at every point. We can then use properties of continuity listed to conclude that every polynomial function in x and y is continuous at every point. For example, $g(x,y) = x^2$ and $h(x,y) = y^3$ are continuous functions, so their product $f(x,y) = x^2y^3$ is a continuous multivariable function.

Summary

- A function $f = f(x,y)$ has a limit L at a point (a,b) provided that we can make $f(x,y)$ as close to L as we like by taking (x,y) sufficiently close (but not equal) to (a,b).
- If (x,y) has two different limits as (x,y) approaches (a,b) along two different paths, we can conclude that $\lim_{(x,y)\to(a,b)} f(x,y)$ does not exist.
- Properties similar to those for one-variable functions allow us to conclude that many limits exist and to evaluate them.
- A function $f = f(x,y)$ is continuous at a point (a,b) in its domain if f has a limit at (a,b) and

$$f(a,b) = \lim_{(x,y)\to(a,b)} f(x,y).$$

11.3.5 Exercises

1. Find the limits, if they exist, or type *DNE* for any which do not exist.

$$\lim_{(x,y)\to(0,0)} \frac{2x^2}{5x^2+5y^2}$$

1) Along the x-axis: _____
2) Along the y-axis: _____
3) Along the line $y = mx$: _____
4) The limit is: _____

2. In this problem we show that the function

$$f(x,y) = \frac{7x-y}{x+y}$$

does not have a limit as $(x,y) \to (0,0)$.

(a) Suppose that we consider $(x,y) \to (0,0)$ along the curve $y = 4x$. Find the limit in this case:

$\lim_{(x,4x)\to(0,0)} \frac{7x-y}{x+y} =$ _____

(b) Now consider $(x,y) \to (0,0)$ along the curve $y = 5x$. Find the limit in this case:

$\lim_{(x,5x)\to(0,0)} \frac{7x-y}{x+y} =$ _____

(c) Note that the results from ***(a)*** and ***(b)*** indicate that f has no limit as $(x,y) \to (0,0)$ *(be sure you can explain why!)*.

To show this more generally, consider $(x,y) \to (0,0)$ along the curve $y = mx$, for arbitrary m. Find the limit in this case:

$\lim_{(x,mx)\to(0,0)} \frac{7x-y}{x+y} =$ _____

(Be sure that you can explain how this result also indicates that f has no limit as $(x,y) \to (0,0)$.

3. Show that the function

$$f(x,y) = \frac{x^3 y}{x^6 + y^3}.$$

does not have a limit at $(0,0)$ by examining the following limits.

(a) Find the limit of f as $(x,y) \to (0,0)$ along the line $y = x$.

$\lim\limits_{\substack{(x,y)\to(0,0)\\ y=x}} f(x,y) =$ _____

(b) Find the limit of f as $(x,y) \to (0,0)$ along the curve $y = x^3$.

$\lim\limits_{\substack{(x,y)\to(0,0)\\ y=x^3}} f(x,y) =$ _____

(Be sure that you are able to explain why the results in ***(a)*** *and* ***(b)*** *indicate that f does not have a limit at (0,0)!*

4. Find the limit, if it exists, or type N if it does not exist.

$$\lim_{(x,y)\to(0,0)} \frac{3x^2}{4x^2 + 1y^2} = \text{_____}$$

5. Find the limit, if it exists, or type N if it does not exist.

$$\lim_{(x,y)\to(0,0)} \frac{(x+17y)^2}{x^2 + 17^2 y^2} = \text{_____}$$

6. Find the limit, if it exists, or type 'DNE' if it does not exist.

$$\lim_{(x,y)\to(3,-5)} e^{\sqrt{5x^2+4y^2}} = \text{_____}$$

7. Find the limit, if it exists. If a limit *does not exist,* type "DNE".

$$\lim_{(x,y)\to(3,3)} \frac{5 - xy}{x^2 + 3y^2}$$

Limit: _____

8. Find the limit, if it exists. If a limit *does not exist,* type "DNE".

$$\lim_{(x,y)\to(4,2)} \left(8x^4 y - 2x^2 + y^2\right)$$

Limit: _____

9. Find the limit, if it exists, or type N if it does not exist.
(Hint: use polar coordinates.)

$$\lim_{(x,y)\to(0,0)} \frac{2x^3 + 1y^3}{x^2 + y^2} = \text{_____}$$

10. Find the limit (enter 'DNE' if the limit does not exist)
Hint: rationalize the denominator.

$$\lim_{(x,y)\to(0,0)} \frac{(3x^2 + 6y^2)}{\sqrt{(3x^2 + 6y^2 + 1)} - 1}$$

_____

11. The largest set on which the function $f(x,y) = 1/(8 - x^2 - y^2)$ is continuous is _____

A. All of the xy-plane

B. All of the xy-plane except the circle $x^2 + y^2 = 8$

C. The interior of the circle $x^2 + y^2 = 8$

D. The interior of the circle $x^2 + y^2 = 8$, plus the circle

E. The exterior of the circle $x^2 + y^2 = 8$

12. The largest set on which the function $f(x, y) = \sqrt{x + y} - \sqrt{x - y}$ is continuous is _____

A. the whole xy-plane

B. $(x, y)| - x < y < x$

C. $(x, y)| - x \leq y \leq x$

D. $(x, y)| - x < y \leq x$

E. $(x, y)|x \geq y$

13. Find the limit of the function

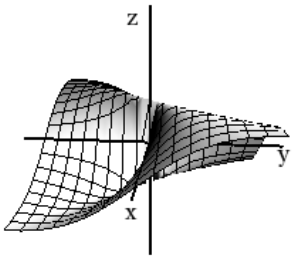
$$f(x, y) = \frac{\sin(5(x^2 + y^2))}{5(x^2 + y^2)}$$

as $(x, y) \to (0, 0)$. Assume that polynomials, exponentials, logarithmic, and trigonometric functions are continuous. *[Hint:* $\lim_{t\to 0} \frac{\sin t}{t} = 1$*.]*

$\lim\limits_{(x,y)\to(0,0)} \frac{\sin(5(x^2+y^2))}{5(x^2+y^2)} =$ _____

*(Enter **DNE** if the limit does not exist.)*

14. Consider the function f whose graph is shown below.



This function is given by

$$f(x, y) = \begin{cases} \frac{6xy}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

(a) Find a formula for the single variable function $f(0, y)$.

$f(0, y) =$ _____

What is $f(0, 0)$ for this function?

$f(0, 0) =$ _____

Find its limit as $y \to 0$:

$\lim\limits_{y\to 0} f(0, y) =$ _____

(b) Based on your work in ***(a)***, is the single variable function $f(0, y)$ continuous? (☐ yes ☐ no)

(c) Next, similarly consider $f(x, 0)$.

$f(x, 0) =$ _____

$f(0, 0) =$ _____

$\lim_{x \to 0} f(x, 0) =$ _____

(d) Based on this work in ***(a)***, is the single variable function $f(x, 0)$ continuous? (☐ yes ☐ no)

(e) Finally, consider f along rays emanating from the origin. Note that these are given by $y = mx$, for some (constant) value of m.

Find and simplify f on the ray $y = x$:

$f(x, x) =$ _____

(Notice that this means that $y = x$ is a contour of f. Be sure you can explain why this is.)

Find and simplify f on any ray $y = mx$.

$f(x, mx) =$ _____

(Again, notice that this means that any ray $y = mx$ is a contour of f; be sure you can explain why.)

(f) Is $f(x, y)$ continuous at $(0, 0)$? (☐ yes ☐ no)

15. Consider the function f defined by $f(x, y) = \frac{xy}{x^2+y^2+1}$.

(a) What is the domain of f?

(b) Evaluate limit of f at $(0, 0)$ along the following paths: $x = 0$, $y = 0$, $y = x$, and $y = x^2$.

(c) What do you conjecture is the value of $\lim_{(x,y)\to(0,0)} f(x, y)$?

(d) Is f continuous at $(0, 0)$? Why or why not?

(e) Use appropriate technology to sketch both surface and contour plots of f near $(0, 0)$. Write several sentences to say how your plots affirm your findings in (a) - (d).

16. Consider the function g defined by $g(x, y) = \frac{xy}{x^2+y^2}$.

(a) What is the domain of g?

(b) Evaluate limit of g at $(0, 0)$ along the following paths: $x = 0$, $y = x$, and $y = 2x$.

(c) What can you now say about the value of $\lim_{(x,y)\to(0,0)} g(x, y)$?

(d) Is g continuous at $(0, 0)$? Why or why not?

(e) Use appropriate technology to sketch both surface and contour plots of g near $(0, 0)$. Write several sentences to say how your plots affirm your findings in (a) - (d).

17. Consider the function h defined by $h(x, y) = \frac{2x^2y}{x^4+y^2}$.

(a) What is the domain of h?

(b) Evaluate the limit of h at $(0, 0)$ along all linear paths the contain the origin. What does this tell us about $\lim_{(x,y)\to(0,0)} h(x, y)$?

Hint. A non-vertical line through the origin has the form $y = mx$ for some constant m.

(c) Does $\lim_{(x,y)\to(0,0)} h(x,y)$ exist? Verify your answer. Check by using appropriate technology to sketch both surface and contour plots of h near $(0,0)$. Write several sentences to say how your plots affirm your findings about $\lim_{(x,y)\to(0,0)} h(x,y)$.

18. For each of the following prompts, provide an example of a function of two variables with the desired properties (with justification), or explain why such a function does not exist.

(a) A function p that is defined at $(0,0)$, but $\lim_{(x,y)\to(0,0)} p(x,y)$ does not exist.

(b) A function q that does not have a limit at $(0,0)$, but that has the same limiting value along any line $y = mx$ as $x \to 0$.

(c) A function r that is continuous at $(0,0)$, but $\lim_{(x,y)\to(0,0)} r(x,y)$ does not exist.

(d) A function s such that

$$\lim_{(x,x)\to(0,0)} s(x,x) = 3 \quad \text{and} \quad \lim_{(x,2x)\to(0,0)} s(x,2x) = 6,$$

for which $\lim_{(x,y)\to(0,0)} s(x,y)$ exists.

(e) A function t that is not defined at $(1,1)$ but $\lim_{(x,y)\to(1,1)} t(x,y)$ does exist.

19. Use the properties of continuity to determine the set of points at which each of the following functions is continuous. Justify your answers.

(a) The function f defined by $f(x,y) = \frac{x+2y}{x-y}$

(b) The function g defined by $g(x,y) = \frac{\sin(x)}{1+e^y}$

(c) The function h defined by

$$h(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

(d) The function k defined by

$$k(x,y) = \begin{cases} \frac{x^2y^4}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

11.4 First-Order Partial Derivatives

Motivating Questions

- Given a function of two variables, $f(x,y)$, how can we measure the rate of change with respect to one of the input variables?
- Given a function f of the variables x and y, what does the rate of change with respect to either x or y tell us about the graph of $z = f(x,y)$?

11.4.1 Introduction

The derivative plays a central role in single-variable calculus because it provides important information about a function in terms of each of the representations of the function. Thinking graphically, the derivative at a point tells us the slope of the line tangent to the graph at the given point. Numerically, the derivative at a point also provides the instantaneous rate of change of the function with respect to a change in the input variable. Algebraically, we can calculate the derivative as a function of the same input variable, which is useful as a tool to examine where the function is increasing or decreasing, as well as the concavity of the function's graph.

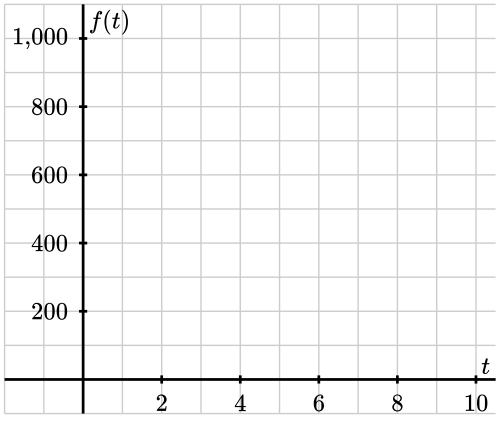
In this chapter, we are investigating functions of two or more variables and we can still ask how fast a function of several variables is changing, but we have to be careful about what we mean. Thinking graphically again, we can try to measure how steep the graph of the function is in a particular direction. As we saw in Section 11.3, the idea of measuring the behavior of a two variable function in a particular direction and generalizing to other directions is a *much* deeper idea when considering multiple input variables. Alternatively, we may want to know how quickly a function's output is changing in response to varying only one of the inputs. Over the next few sections, we will develop tools for addressing these questions. We begin with a Preview Activity that considers how we can apply our knowledge of single-variable calculus to measure the rate of change when we hold all but one input variable constant.

Preview Activity 11.4.1 Suppose we take out an \$18,000 car loan at interest rate r and we agree to pay off the loan in t years. The monthly payment, in dollars, is

$$M(r,t) = \frac{1500r}{1-\left(1+\frac{r}{12}\right)^{-12t}}.$$

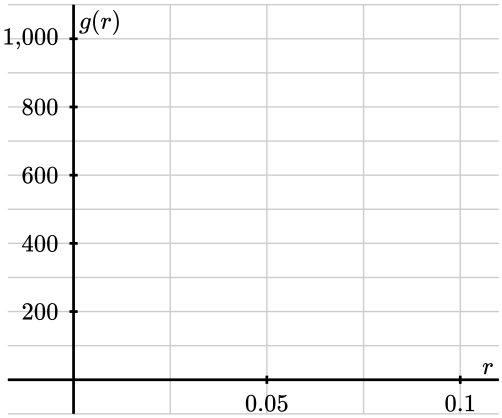
(a) If you take a loan with an interest rate of 3% (so $r = 0.03$) and will pay it off in $t = 4$ years, what will your monthly payment be?

(b) Suppose the interest rate on loans is fixed at 3%. Express M as a function f of t alone using $r = 0.03$. That is, let $f(t) = M(0.03, t)$. Sketch the graph of f on the axes below and write a couple of sentences to explain the meaning of the function f.



(c) Find the instantaneous rate of change $f'(4)$ and state the units on this quantity. What information does $f'(4)$ tell us about our car loan? What information does $f'(4)$ tell us about the graph you sketched in (b)?

(d) Suppose that we want to examine how the monthly payment will change if we plan to take 4 years to pay off the loan. Express M as a function of the interest rate r with a fixed loan term of $t = 4$. That is, let $g(r) = M(r, 4)$. Give the function $g(r)$ and sketch a graph of g on the axes provided. Write a couple of sentences to explain the meaning of the function g.



(e) Find the instantaneous rate of change $g'(0.03)$ and state the units on this quantity. Write a couple of sentences to explain what $g'(0.03)$ tell us about the car loan. What information does $g'(0.03)$ tell us about the graph you sketched in (d)?

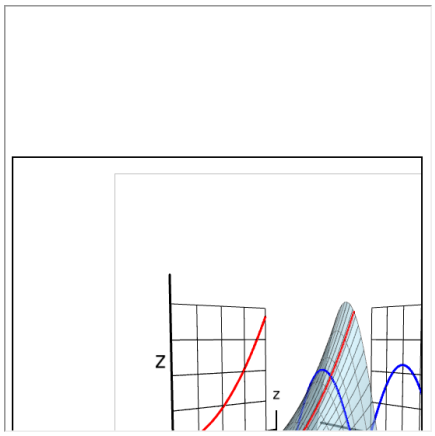
11.4.2 First-Order Partial Derivatives

In Section 11.2, we studied the behavior of a function of two or more variables by considering the *traces* of the function. By holding one of our input variables constant, we restricted our multivariable function to have only one input variable, which means we can use all of our calculus tools from single-variable calculus on these traces.

Recall that in an earlier example, we considered the function f defined by

$$f(x, y) = \frac{x^2 \sin(2y)}{32}.$$

This function measures horizontal distance (in feet) traveled by a projectile launched with an initial speed of x feet per second at an angle of y radians to the horizontal. The graph of this function is shown in Figure 11.4.1 with traces plotted on the surface for $x = 150$ (in blue) and $y = 0.6$ (in red).



Standalone
In Context
Embed

Figure 11.4.1 The projectile distance surface given by $z = f(x, y) = \frac{x^2 \sin(2y)}{g}$

If we fix the angle $y = 0.6$, we can view the trace $f(x, 0.6)$ as a function of x alone, as seen in Figure 11.4.2.

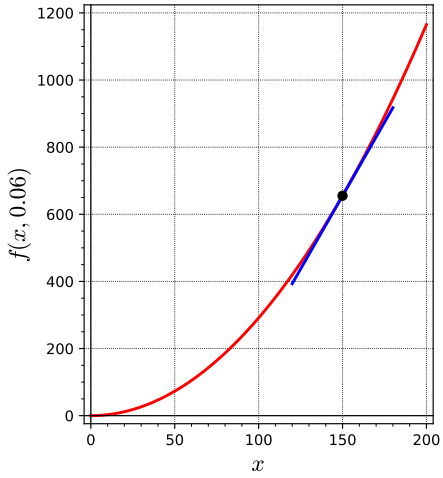


Figure 11.4.2 A plot of $f(x, 0.6)$

Since the trace is a function of only one variable, we can consider its derivative exactly as we did in single-variable calculus. With $y = 0.6$, we have

$$f(x, 0.6) = \frac{\sin(1.2)}{32} x^2,$$

and therefore

$$\frac{d}{dx}\Big[f(x, 0.6)\Big] = \frac{\sin(1.2)}{16} x.$$

When $x = 150$, this is

$$\frac{d}{dx}\Big[f(x, 0.6)\Big]\Big|_{x=150} = \frac{\sin(1.2)}{16} 150 \approx 8.74 \text{ feet per (feet per second)},$$

which gives the slope of the tangent line shown in Figure 11.4.2.

Thinking of this derivative as an instantaneous rate of change implies that for a fixed launch angle of 0.6 radians, increasing the initial speed of the projectile by one foot per second leads to an expected increase in the horizontal distance traveled of approximately 8.74 feet. Since the sign of the derivative of the trace is positive there would be an increase to the range of the projectile

if we increased the initial speed from 150 ft/s while holding the launch angle constant at 0.6 radians.

By holding y fixed and differentiating with respect to x, we obtain the first-order **partial derivative of f with respect to x**. Denoting this partial derivative as f_x, we have seen that

$$f_x(150, 0.6) = \frac{d}{dx}\Big[f(x, 0.6)\Big]\Big|_{x=150} = \lim_{h\to 0} \frac{f(150+h, 0.6) - f(150, 0.6)}{h}.$$

More generally, we have

$$f_x(a, b) = \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h},$$

provided this limit exists.

In the same way, we may obtain a trace by setting, say, $x = 150$ as shown in Figure 11.4.3.

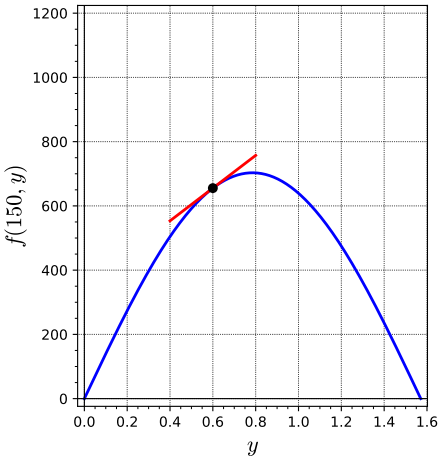


Figure 11.4.3 A plot of $f(150, y)$

This gives

$$f(150, y) = \frac{150^2}{32}\sin(2y),$$

and therefore

$$\frac{d}{dy}[f(150, y)] = \frac{150^2}{16}\cos(2y).$$

If we evaluate this quantity at $y = 0.6$, we have

$$\frac{d}{dy}\Big[f(150, y)\Big]\Big|_{y=0.6} = \frac{150^2}{16}\cos(1.2) \approx 509.5 \text{ feet per radian.}$$

Once again, the derivative gives the slope of the tangent line shown on the right in Figure 11.4.3. Thinking of this derivative as an instantaneous rate of change, we expect that, for a fixed initial speed of the projectile of 150 feet per second, the range of the projectile increases by approximately 509.5 feet for every radian we increase the launch angle y.

By holding x fixed and differentiating with respect to y, we obtain the first-order **partial derivative of f with respect to y**. As before, we denote this partial derivative as f_y and write

$$f_y(150, 0.6) = \frac{d}{dy}\Big[f(150, y)\Big]\Big|_{y=0.6} = \lim_{h\to 0} \frac{f(150, 0.6+h) - f(150, 0.6)}{h}.$$

As with the partial derivative with respect to x, we may express this quantity more generally at an arbitrary point (a, b). To recap, we have now arrived at the formal definition of the first-order partial derivatives of a function of two variables.

Definition 11.4.4 The first-order **partial derivatives of f with respect to x and y** at a point (a, b) are, respectively,

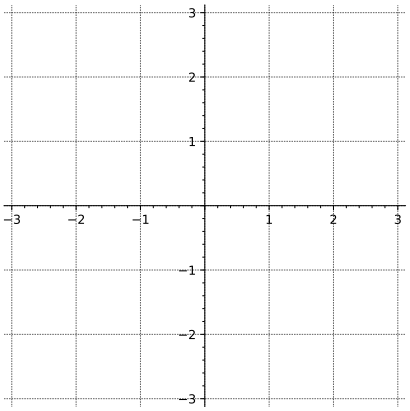
$$f_x(a,b) = \frac{\partial f}{\partial x}(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h}$$
$$f_y(a,b) = \frac{\partial f}{\partial y}(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h},$$

provided the limits exist. ◆

We did not use the classic calculus approach in our development here because we defined the partial derivatives along a trace of our function, which restricts our function to one variable. As you will see in the next several activities, this allows us to continue using ideas from single variable calculus rather than needing to develop new multivariable tools.

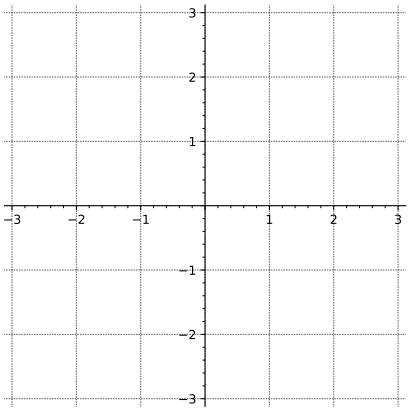
Activity 11.4.2 In this activity, we will be examining the function $f(x, y) = \dfrac{xy^2}{x+1}$ at the point $(1, 2)$. Specifically, we will be looking at how the traces of this function will be related to the values of the partial derivatives.

(a) Write a formula for the trace $f(x, 2)$ as a function of x. On the axes provided, draw the graph of the trace with $y = 2$ around the point where $x = 1$. Then sketch the tangent line at the point with $x = 1$ on your graph.



(b) Compute the partial derivative $f_x(1, 2)$ and relate its value to the sketch you just made.

(c) Write a formula for the trace $f(1, y)$ as a function of y. On the axes provided, draw the graph of the trace with $x = 1$. Then sketch the tangent line at the point with $y = 2$ on your graph.



(d) Compute the partial derivative $f_y(1,2)$ and relate its value to the sketch you just made.

What we have seen so far demonstrates that each partial derivative at a point arises as the derivative of a function of one variable defined by fixing all other coordinates.[1] In addition, we may consider each partial derivative as defining a new function with input given by the point (x,y), just as the derivative $f'(x)$ defines a new function of x in single-variable calculus. Due to the connection between single-variable derivatives and partial derivatives, we often use Leibniz-style notation to denote partial derivatives by writing

$$\frac{\partial f}{\partial x}(a,b) = f_x(a,b), \text{ and } \frac{\partial f}{\partial y}(a,b) = f_y(a,b).$$

We read the notation $\partial f/\partial x$ as "partial f-partial x".

To calculate the partial derivative f_x, we hold y fixed. In other words, we treat y as a constant. In Leibniz notation, observe that

$$\frac{\partial}{\partial x}(x) = 1 \text{ and } \frac{\partial}{\partial x}(y) = 0.$$

To see the contrast between how we calculate single-variable derivatives and partial derivatives and the difference between the notations $\frac{d}{dx}[\ ]$ and $\frac{\partial}{\partial x}[\ ]$, observe that

$$\frac{d}{dx}\Big[3x^2 - 2x + 3\Big] = 3\frac{d}{dx}\Big[x^2\Big] - 2\frac{d}{dx}\Big[x\Big] + \frac{d}{dx}\Big[3\Big] = 3\cdot 2x - 2$$
$$\frac{\partial}{\partial x}\Big[x^2y - xy + 2y\Big] = y\frac{\partial}{\partial x}\Big[x^2\Big] - y\frac{\partial}{\partial x}\Big[x\Big] + \frac{\partial}{\partial x}\Big[2y\Big] = y\cdot 2x - y.$$

Because there is a lot going on in the computations above, let us highlight some terms for emphasis. Note that when we are interested in computing the partial derivative with respect to x of a term of the form x^2y, we can factor out the y as if it were a coefficient because y is *constant with respect to changes in* x. Algebraically, this can be expressed as

$$\frac{\partial}{\partial x}\Big[x^2y\Big] = y\frac{\partial}{\partial x}\Big[x^2\Big].$$

Similarly, when we are computing the partial derivative with respect to x of a term of the form $2y$, the entire term is constant with respect to x. Thus,

$$\frac{\partial}{\partial x}\Big[2y\Big] = 0.$$

[1] At this stage, we have considered partial derivatives for functions of two variables, but the methods generalize to functions of three or more variables by fixing all but one variable.

To summarize, computing partial derivatives can be done directly by applying familiar differentiation rules from single-variable calculus, but we do so while holding all but one of the variables constant. The next activity has you practice computing partial derivatives using these ideas.

Activity 11.4.3

(a) If $f(x, y) = 3x^3 - 2x^2y^5$, find the partial derivatives f_x and f_y.

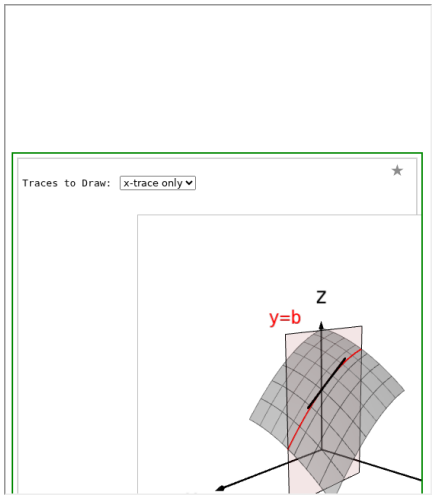
(b) If $f(x, y) = \dfrac{xy^2}{x+1}$, compute $\partial f/\partial x$ and $\partial f/\partial y$.

(c) If $g(r, s) = rs\cos(r)$, find the partial derivatives g_r and g_s.

(d) Find all possible first-order partial derivatives of $q(x, t, z) = \dfrac{x2^t z^3}{1+x^2}$.

11.4.3 Interpreting and Estimating First-Order Partial Derivatives

Throughout single-variable calculus, you frequently used the fact that the derivative of a single-variable function can be interpreted geometrically as the slope of the line tangent to the graph at a given point. Similarly, we have seen that the partial derivatives measure the slope of a line tangent to a *trace* of a function of two variables as shown in Figure 11.4.5. Use the drop-down to change the trace and tangent line shown or show both traces and tangent lines.



Standalone
In Context
Embed

Figure 11.4.5 A surface given by $z = f(x, y)$ with traces $y = b$ (in red) and $x = a$ (in blue)

Beyond representing the slope of the tangent line to the graph of a one-variable function or to a trace of a multivariable function, the values of derivatives and partial derivatives at a point give us the instantaneous rate of change of the function at that point. This perspective makes it possible for us to contemplate the units on partial derivatives, as the next activity asks you to investigate.

Activity 11.4.4 The speed of sound C traveling through ocean water is a function of temperature, salinity and depth. It may be modeled by the function

$$C = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3$$

$$+ (1.34 - 0.01T)(S - 35) + 0.016D.$$

Here C is the speed of sound in meters/second, T is the temperature in degrees Celsius, S is the salinity in grams/liter of water, and D is the depth below the ocean surface in meters.

(a) State the units for each of the partial derivatives, C_T, C_S and C_D. For each partial derivative, write a sentence that explains the physical meaning of the partial derivatives in terms of the rate of change for C.

(b) Use algebraic rules to find the partial derivatives C_T, C_S and C_D.

(c) Evaluate each of the three partial derivatives at the point where $T = 10$, $S = 35$ and $D = 100$. Write a few sentences to explain what the sign of each partial derivatives tell us about the behavior of the function C at the point $(10, 35, 100)$.

In single-variable calculus, you considered not only functions represented algebraically, but also those represented using tables of data or graphs. You used these representations and the difference quotient $\frac{f(a+h)-f(a)}{h}$ to estimate derivatives. This is possible because for a function f of a single variable x at a point where $x = a$, the difference quotient tells us the *average* rate of change of f over the interval $[a, a + h]$, while the derivative $f'(a)$ tells us the *instantaneous* rate of change of f at $x = a$. We can use a similar relationship between limits of difference quotients and partial derivatives to provide context and meaning to partial derivatives as a measurement. To do this, we need to make sure we are holding all but one input variable constant when we compute the difference quotient. The next two activities are focused on the meaning of partial derivatives using a table as a numerical representation of a multivariable function (Activity 11.4.5) and using a contour plot as a graphical representation of a multivariable function (Activity 11.4.6).

Activity 11.4.5 The wind chill, as frequently reported, is a measure of how cold it feels outside when the wind is blowing. In the table below, the wind chill w is a function of the wind speed v and the ambient air temperature T. Both w and T are measured in degrees Fahrenheit, while v is measured in miles per hour. This means that w can be expressed in the form $w = w(v, T)$.

$v\backslash T$	-20	-15	-10	-5	0	5	10
10	-41	-35	-28	-22	-16	-10	-4
15	-45	-39	-32	-26	-19	-13	-7
20	-48	-42	-35	-29	-22	-15	-9
25	-51	-44	-37	-31	-24	-17	-11
30	-53	-46	-39	-33	-26	-19	-12
35	-55	-48	-41	-34	-27	-21	-14

(a) When computing the partial derivative with respect to wind speed w_v, which variable is held constant and which variable changes? Do you think w_v is positive, negative, or zero at $(v, T) = (20, -10)$? Write a couple of sentences to explain your reasoning.

(b) When computing the partial derivative with respect to temperature w_T, which variable is held constant and which variable changes? Do you think w_T is positive, negative, or zero at $(v, T) = (20, -10)$? Write a couple of sentences to explain your reasoning.

(c) Recall that we can estimate a partial derivative of a single variable function f using the central difference $\frac{f(x+h)-f(x-h)}{2h}$ for small values of h. A

partial derivative is a derivative of an appropriate trace, so by considering a difference quotient for a trace of w, estimate the partial derivative $w_v(20, -10)$. Include units on your answer and write a sentence explaining the meaning of the value of $w_v(20, -10)$.

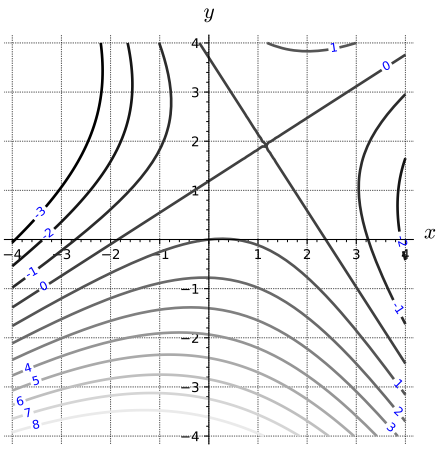
(d) Recall from single-variable calculus that for a function f, $f(a + h) \approx f(a) + hf'(a)$. Use this idea and your result above to estimate the wind chill $w(18, -10)$.

(e) Estimate the partial derivative $w_T(20, -10)$. Include units on your answer and write a sentence explaining the meaning of the value of $w_T(20, -10)$.

(f) Use your result above to estimate the wind chill $w(20, -12)$.

(g) Consider how you might combine your previous results to estimate the wind chill $w(18, -12)$. Explain your process.

Activity 11.4.6 Shown below is a contour plot of a function f. The output values of the function are indicated on the contours. This activity will ask you to use this contour plot to estimate some partial derivatives of f.



(a) We first wish to estimate the partial derivative $f_x(-2, -1)$.

(i) When computing f_x, which variable is treated as a constant? On the contour plot, sketch a line segment for which that variable is constant through the point $(-2, -1)$.

(ii) To estimate $f_x(-2, -1)$ using a difference quotient, you must identify two points in the xy-plane that are near $(-2, -1)$, on the line segment you sketched in the previous part (or an extension of that line segment), and for which you can use the contour plot to find the value of the function at those points. Mark these points, which we will call (a_1, b_1) and (a_2, b_2) on your contour plot and record the values of f at these points.

Hint. Since your points are on the line segment you sketched before, your points must have either $a_1 = a_2$ or $b_1 = b_2$.

(iii) Use a difference quotient to estimate $f_x(-2, -1)$.

Hint. If possible, you will use a central difference, but if your points are not equally spaced from $(-2, -1)$, you may have a different expression that is still a difference quotient.

(b) In this part and the next, you will need to follow a similar process to the previous part. However, the activity does not break down the steps each time. Estimate the partial derivative $f_y(-2,-1)$.

(c) Estimate the partial derivatives $f_x(1,-1)$ and $f_y(1,-1)$.

(d) Locate, if possible, one point (x,y) where $f_x(x,y)=0$.

(e) Locate, if possible, one point (x,y) where $f_x(x,y)<0$.

(f) Locate, if possible, one point (x,y) where $f_y(x,y)>0$.

Notice that the approximations, algebraic rules, and interpretations developed in this section for partial derivatives are the same as in single-variable calculus. Additionally, the form of the difference quotients and limit definition for partial derivatives maintains the same form as in single-variable calculus, which we also saw when studying vector-valued functions of one variable in the previous chapter:

$$\text{rate of change} = \lim \frac{\text{change in output of function}}{\text{change in input of function}}$$

Importantly and conveniently, all of the change we look at with a partial derivative is with respect to a single input variable as we consider holding all other input variables constant.

Notice that in this section, we did not need to use the classic calculus approach from scratch to define the partial derivative because we restricted our multivariable function along a trace to be a single-variable function. This was convenient, but we will soon see that we will need new tools to make similar arguments when studying the rate of change of a multivariable function in a direction that is not along a trace such as $x=a$ or $y=b$.

Summary

- If $f=f(x,y)$ is a function of two variables, there are two first-order partial derivatives of f: the partial derivative of f with respect to x,

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

  and the partial derivative of f with respect to y

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

  where each partial derivative exists only at those points (x,y) for which the limit exists.

- The partial derivative $f_x(a,b)$ tells us the instantaneous rate of change of f with respect to x at $(x,y)=(a,b)$ when y is fixed at b. Geometrically, the partial derivative $f_x(a,b)$ tells us the slope of the line tangent to the $y=b$ trace of the function f at the point $(a,b,f(a,b))$.

- The partial derivative $f_y(a,b)$ tells us the instantaneous rate of change of f with respect to y at $(x,y)=(a,b)$ when x is fixed at a. Geometrically, the partial derivative $f_y(a,b)$ tells us the slope of the line tangent to the $x=a$ trace of the function f at the point $(a,b,f(a,b))$.

11.4.4 Exercises

1. Let $z = \sqrt{4x + 5y}$. Then:

The rate of change in z at $(4, 1)$ as we change x but hold y fixed is _____, and

The rate of change in z at $(4, 1)$ as we change y but hold x fixed is _____.

2. Let $z = \cos\left(4y^2 + x\right)$. Then:

The rate of change in z at $(-5, 4)$ as we change x but hold y fixed is _____, and

The rate of change in z at $(-5, 4)$ as we change y but hold x fixed is _____.

3. Find the first partial derivatives of

$f(x, y) = \dfrac{4x - y}{4x + y}$ at the point $(x, y) = (3, 3)$.

$\dfrac{\partial f}{\partial x}(3, 3) =$ _____

$\dfrac{\partial f}{\partial y}(3, 3) =$ _____

4. Find the first partial derivatives of $f(x, y) = \sin(x - y)$ at the point (-8, -8).

A. $f_x(-8, -8) =$ _____

B. $f_y(-8, -8) =$ _____

5. Find the first partial derivatives of the function $z = (2x + 8y)^{11}$.

1. $\dfrac{\partial z}{\partial x} =$ _____

2. $\dfrac{\partial z}{\partial y} =$ _____

6. Find the first partial derivatives of the function $f(x, y) = \dfrac{3x}{(x + y)^2}$.

1. $f_x(x, y) =$ _____

2. $f_y(x, y) =$ _____

7. Find the first partial derivatives of the function $u(r, t) = 3\sin(r\cos t))$.

1. $u_r(r, t) =$ _____

2. $u_t(r, t) =$ _____

8. Find the first partial derivatives of the function $f(x, y) = \ln\left(x + \sqrt{x^2 + y^2}\right)$.

1. $f_x(3, 4) =$ _____

2. $f_x(12, 5) =$ _____

9. Let $z = 5e^{x^2y^2}$. Then:

$\frac{\partial z}{\partial x} =$ _____

$\frac{\partial z}{\partial y} =$ _____

10. Let $z = \frac{xy}{3x^2+3y^2}$. Then:

$\frac{\partial z}{\partial x} =$ _____

$\frac{\partial z}{\partial y} =$ _____

11. Find the partial derivatives indicated. Assume the variables are restricted to a domain on which the function is defined.

$$f(x, y) = x^5 + 7x^3y.$$

$f_x(2,4) =$ _____
$f_y(2,4) =$ _____

12. Find the partial derivatives indicated. Assume the variables are restricted to a domain on which the function is defined.

$$z = \left(x^5 + x - y\right)^3.$$

$\frac{\partial z}{\partial x} =$ _____
$\frac{\partial z}{\partial y} =$ _____

13. Find the partial derivative indicated. Assume the variables are restricted to a domain on which the function is defined.

$$z = \frac{5x^2y^6 - y^6}{16xy - 7}.$$

$z_y =$ _____

14. Find the partial derivative indicated. Assume the variables are restricted to a domain on which the function is defined.

$$z = \sin\left(4x^3y - 3xy^3\right).$$

$z_x =$ _____

15. Find the first partial derivatives of
$f(x,y) = \dfrac{2x-y}{2x+y}$ at the point $(x,y) = (3,2)$.
$\dfrac{\partial f}{\partial x}(3,2) =$ _____
$\dfrac{\partial f}{\partial y}(3,2) =$ _____

16. Find the first partial derivatives of $f(x,y) = \sin(x-y)$ at the point (-2, -2).
A. $f_x(-2,-2) =$ _____
B. $f_y(-2,-2) =$ _____

17. Find the partial derivatives of the function

$$w = \sqrt{5r^2 + 4s^2 + 4t^2}$$

$\frac{\partial w}{\partial r} =$ _____
$\frac{\partial w}{\partial s} =$ _____
$\frac{\partial w}{\partial t} =$ _____

18. Suppose that $f(x,y)$ is a smooth function and that its partial derivatives have the values, $f_x(3,-9) = 3$ and $f_y(3,-9) = 1$. Given that $f(3,-9) = -1$, use this information to estimate the value of $f(4,-8)$. Note this is analogous to finding the tangent line approximation to a function of one variable. In fancy terms, it is the first Taylor approximation.
Estimate of (integer value) $f(3,-8)$ _____
Estimate of (integer value) $f(4,-9)$ _____
Estimate of (integer value) $f(4,-8)$ _____

19. The gas law for a fixed mass m of an ideal gas at absolute temperature T, pressure P, and volume V is $PV = mRT$, where R is the gas constant.
Find the partial derivatives

$\dfrac{\partial P}{\partial V} =$_____

$\dfrac{\partial V}{\partial T} =$_____

$\dfrac{\partial T}{\partial P} =$_____

$\dfrac{\partial P}{\partial V}\dfrac{\partial V}{\partial T}\dfrac{\partial T}{\partial P} =$_____ (an integer)

20. Find the partial derivatives of the function

$$f(x,y) = \int_y^x \cos(7t^2 - 3t + 2)\,dt$$

$f_x(x,y) =$_____

$f_y(x,y) =$_____

21. Let $f(x,y) = e^{-2x}\sin(5y)$.

(a) Using difference quotients with $\Delta x = 0.1$ and $\Delta y = 0.1$, we estimate

$f_x(-3,1) \approx$ _____

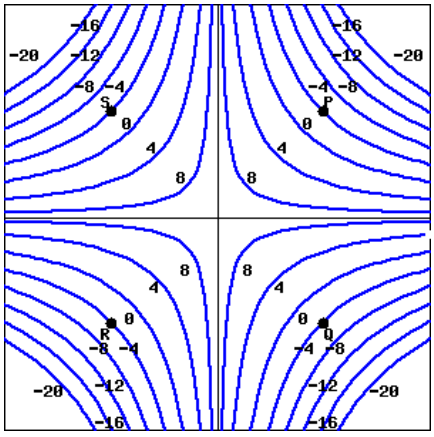
$f_y(-3,1) \approx$ _____

(b) Using difference quotients with $\Delta x = 0.01$ and $\Delta y = 0.01$, we find better estimates:

$f_x(-3,1) \approx$ _____

$f_y(-3,1) \approx$ _____

22. Determine the sign of f_x and f_y at each indicated point using the contour diagram of f shown below. (The point P is that in the first quadrant, at a positive x and y value; Q through T are located clockwise from P, so that Q is at a positive x value and negative y, etc.)



(a) At point P,

f_x is (☐ positive ☐ negative) and

f_y is (☐ positive ☐ negative) .

(b) At point Q,

f_x is (☐ positive ☐ negative) and

f_y is (☐ positive ☐ negative) .

(c) At point R,

f_x is (☐ positive ☐ negative) and

f_y is (☐ positive ☐ negative) .

23. An experiment to measure the toxicity of formaldehyde yielded the data in the table below. The values show the percent, $P = f(t,c)$, of rats

surviving an exposure to formaldehyde at a concentration of c (in parts per million, ppm) after t months.

	$t = 14$	$t = 16$	$t = 18$	$t = 20$	$t = 22$	$t = 24$
$c = 0$	100	100	100	99	97	95
$c = 2$	100	99	98	97	95	92
$c = 6$	96	95	93	90	86	80
$c = 15$	96	93	82	70	58	36

(a) Estimate $f_t(16, 2)$:

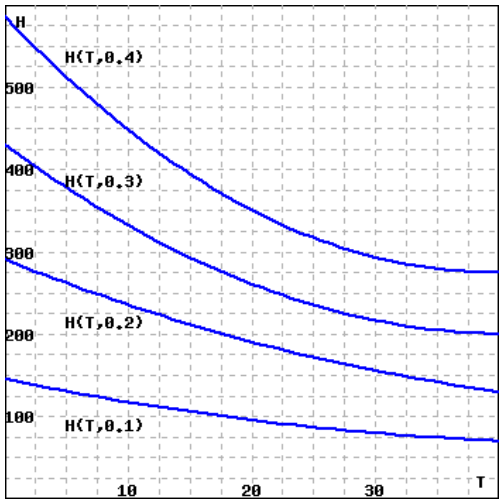
$f_t(16, 2) \approx$ _____

(b) Estimate $f_c(16, 2)$:

$f_c(16, 2) \approx$ _____

(Be sure that you can give the practical meaning of these two values in terms of formaldehyde toxicity.)

24. An airport can be cleared of fog by heating the air. The amount of heat required depends on the air temperature and the wetness of the fog. The figure below shows the heat $H(T, w)$ required (in calories per cubic meter of fog) as a function of the temperature T (in degrees Celsius) and the water content w (in grams per cubic meter of fog). Note that this figure is not a contour diagram, but shows cross-sections of H with w fixed at 0.1, 0.2, 0.3, and 0.4.



(a) Estimate $H_T(10, 0.2)$:

$H_T(10, 0.2) \approx$ _____

(Be sure you can interpret this partial derivative in practical terms.)

(b) Make a table of values for $H(T, w)$ from the figure, and use it to estimate $H_T(T, w)$ for each of the following:

$T = 20, w = 0.2 : H_T(T, w) \approx$ _____

$T = 30, w = 0.2 : H_T(T, w) \approx$ _____

$T = 20, w = 0.3 : H_T(T, w) \approx$ _____

$T = 30, w = 0.3 : H_T(T, w) \approx$ _____

(c) Repeat ***(b)*** to find $H_w(T, w)$ for each of the following:

$T = 20, w = 0.2 : H_w(T, w) \approx$ _____

$T = 30, w = 0.2 : H_w(T, w) \approx$ _____

$T = 20, w = 0.3 : H_w(T, w) \approx$ _____

$T = 30, w = 0.3 : H_w(T, w) \approx$ _____

(Be sure you can interpret this partial derivative in practical terms.)

25. The following questions refer to the table shown below, giving the temperature adjusted for wind-chill, C, in ° F, as a function $f(w,T)$ of the wind speed, w, in mph, and the temperature, T, in ° F. The temperature adjusted for wind-chill tells you how cold it feels, as a result of the combination of wind and temperature.

	$T=35$	$T=30$	$T=25$	$T=20$	$T=15$	$T=10$	$T=5$	$T=0$
$w=5$	31	25	19	13	7	1	-5	-11
$w=10$	27	21	15	9	3	-4	-10	-16
$w=15$	25	19	13	6	0	-7	-13	-19
$w=20$	24	17	11	4	-2	-9	-15	-22
$w=25$	23	16	9	3	-4	-11	-17	-24

(a) Estimate $f_w(20,20)$:

$f_w(20,20) \approx$ _____

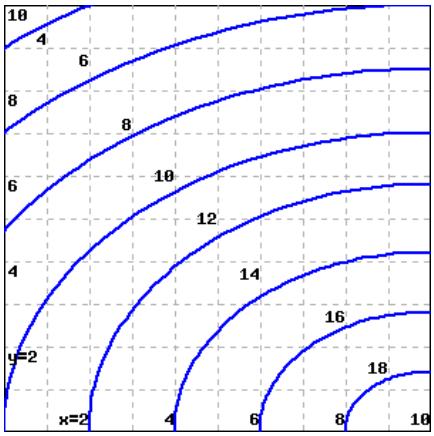
Be sure you can state in practical terms what your answer means.

(b) Estimate $f_T(15,10)$:

$f_T(15,10) \approx$ _____

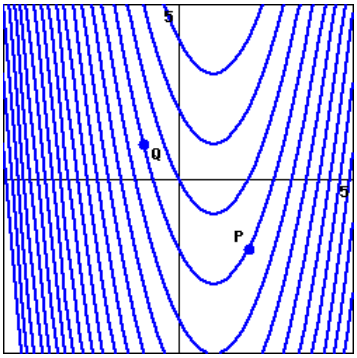
Be sure you can state in practical terms what your answer means.

26. Approximate $f_x(1,3)$ using the contour diagram of $f(x,y)$ shown below.



$f_x(1,3) \approx$ _____

27. The figure below shows contours of $f(x,y)$, with values of f on the contours omitted.



If $f_y(Q)$ is negative, answer the following questions about the other derivatives of f at P and Q.

(a) What is the sign of $f_x(Q)$? (☐ positive ☐ negative ☐ zero)

(b) What is the sign of $f_x(P)$? (☐ positive ☐ negative ☐ zero)
(c) What is the sign of $f_y(P)$? (☐ positive ☐ negative ☐ zero)

28. A one-meter long bar is heated unevenly, with temperature in °C at a distance x meters from one end at time t given by

$$H(x,t) = 140e^{-0.15t}\sin(\pi x)$$

$0 \le x \le 1$. On a sheet of paper, sketch a graph of H against x for $t = 0$ and $t = 1$. Think about why your graphs make sense.

(a) Calculate each of:
$H_x(0.2,t) =$ _____
$H_x(0.8,t) =$ _____.
(Be sure that you can say in words what the practical interpretation (in terms of temperature) of these two partial derivatives is, and why each has the sign that it does.)

(b) Calculate:
$H_t(x,t) =$ _____.
(Again, be sure that you understand why it has the sign that it does, and what its interpretation in terms of temperature is.)

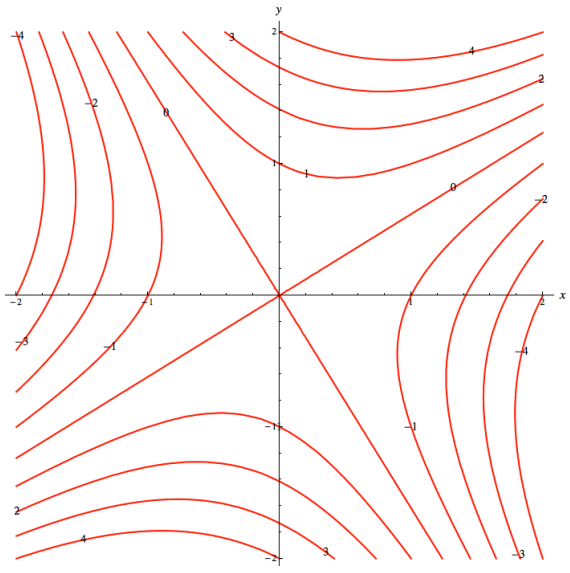
29. Consider the partial derivatives

$$f_x(x,y) = 5x^4y^6 - 15x^2y,$$

$$f_y(x,y) = 6x^5y^5 - 5x^3.$$

Is there a function f which has these partial derivatives?
(☐ Yes ☐ No)
If so, what is it?
$f =$ _____
*(Enter **none** if there is no such function.)*
Are there any others?
(☐ Yes ☐ No)

30. Determine the signs of the partial derivatives for the function $f(x,y)$ whose contour map is graphed below.



1. $f_x(1,0)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)
2. $f_y(1,0)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)

3. $f_x(-1,0)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)
4. $f_y(-1,0)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)
5. $f_x(0,-1)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)
6. $f_y(0,-1)$ (☐ choose one ☐ Positive ☐ Negative ☐ Zero)
Note: *You can click on the graph to enlarge the image.*
Note: *You only have two attempts at this problem.*

31. Use difference quotients with $\Delta x = 0.1$ and $\Delta y = 0.1$ to estimate $f_x(1,3)$ and $f_y(1,3)$ where

$$f(x,y) = e^{-x}\sin(y)\,.$$

$f_x(1,3) \approx$ _____
$f_y(1,3) \approx$ _____
Then give better estimates by using $\Delta x = 0.01$ and $\Delta y = 0.01$.
$f_x(1,3) \approx$ _____
$f_y(1,3) \approx$ _____

32. The Heat Index, I, (measured in *apparent degrees F*) is a function of the actual temperature T outside (in degrees F) and the relative humidity H (measured as a percentage). A portion of the table which gives values for this function, $I = I(T,H)$, is reproduced in Table 11.4.6.

Table 11.4.6 A portion of the heat index data

$T\downarrow \setminus H\rightarrow$	70	75	80	85
90	106	109	112	115
92	112	115	119	123
94	118	122	127	132
96	125	130	135	141

(a) State the limit definition of the value $I_T(94,75)$. Then, estimate $I_T(94,75)$, and write one complete sentence that carefully explains the meaning of this value, including its units.

(b) State the limit definition of the value $I_H(94,75)$. Then, estimate $I_H(94,75)$, and write one complete sentence that carefully explains the meaning of this value, including its units.

(c) Suppose you are given that $I_T(92,80) = 3.75$ and $I_H(92,80) = 0.8$. Estimate the values of $I(91,80)$ and $I(92,78)$. Explain how the partial derivatives are relevant to your thinking.

(d) On a certain day, at 1 p.m. the temperature is 92 degrees and the relative humidity is 85%. At 3 p.m., the temperature is 96 degrees and the relative humidity 75%. What is the average rate of change of the heat index over this time period, and what are the units on your answer? Write a sentence to explain your thinking.

33. Let $f(x,y) = \frac{1}{2}xy^2$ represent the kinetic energy in Joules of an object of mass x in kilograms with velocity y in meters per second. Let (a,b) be the point $(4,5)$ in the domain of f.

(a) Calculate $f_x(a,b)$.

(b) Explain as best you can in the context of kinetic energy what the partial derivative

$$f_x(a,b) = \lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}$$

tells us about kinetic energy.

(c) Calculate $f_y(a, b)$.

(d) Explain as best you can in the context of kinetic energy what the partial derivative

$$f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h}$$

tells us about kinetic energy.

(e) Often we are given certain graphical information about a function instead of a rule. We can use that information to approximate partial derivatives. For example, suppose that we are given a contour plot of the kinetic energy function (as in Figure 11.4.7) instead of a formula. Use this contour plot to approximate $f_x(4, 5)$ and $f_y(4, 5)$ as best you can. Compare to your calculations from earlier parts of this exercise.

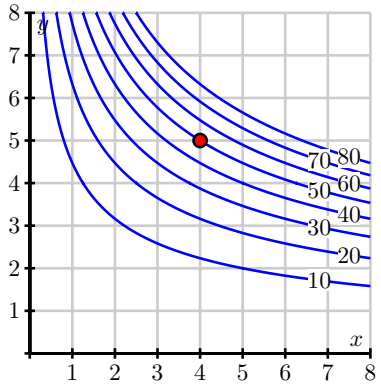


Figure 11.4.7 A contour plot of $f(x, y) = \frac{1}{2}xy^2$.

34. The temperature on an unevenly heated metal plate positioned in the first quadrant of the xy-plane is given by

$$C(x, y) = \frac{25xy + 25}{(x - 1)^2 + (y - 1)^2 + 1}.$$

Assume that temperature is measured in degrees Celsius and that x and y are each measured in inches. (*Note:* At no point in the following questions should you expand the denominator of $C(x, y)$.)

(a) Determine $\left.\frac{\partial C}{\partial x}\right|_{(x,y)}$ and $\left.\frac{\partial C}{\partial y}\right|_{(x,y)}$.

(b) If an ant is on the metal plate, standing at the point $(2, 3)$, and starts walking in the direction parallel to the positive y axis, at what rate will the temperature the ant is experiencing change? Explain, and include appropriate units.

(c) If an ant is walking along the line $y = 3$ in the positive x direction, at what instantaneous rate will the temperature the ant is experiencing change when the ant passes the point $(1, 3)$?

(d) Now suppose the ant is stationed at the point $(6, 3)$ and walks in a straight line towards the point $(2, 0)$. Determine the *average* rate of change in temperature (per unit distance traveled) the ant encounters in moving between these two points. Explain your reasoning carefully. What are the units on your answer?

35. Consider the function f defined by $f(x, y) = 8 - x^2 - 3y^2$.

(a) Determine $f_x(x, y)$ and $f_y(x, y)$.

(b) Find parametric equations in $\mathbb{R}^3$ for the tangent line to the trace $f(x, 1)$ at $x = 2$.

(c) Find parametric equations in $\mathbb{R}^3$ for the tangent line to the trace $f(2, y)$ at $y = 1$.

(d) State respective direction vectors for the two lines determined in (b) and (c).

(e) Determine the equation of the plane that passes through the point $(2, 1, f(2, 1))$ whose normal vector is orthogonal to the direction vectors of the two lines found in (b) and (c).

(f) Use a graphing utility to plot both the surface $z = 8 - x^2 - 3y^2$ and the plane from (e) near the point $(2, 1)$. What is the relationship between the surface and the plane?

36. Recall from single variable calculus that, given the derivative of a single variable function and an initial condition, we can integrate to find the original function. We can sometimes use the same process for functions of more than one variable. For example, suppose that a function f satisfies $f_x(x, y) = \cos(y)e^x + 2x + y^2$, $f_y(x, y) = -\sin(y)e^x + 2xy + 3$, and $f(0, 0) = 5$.

(a) Find all possible functions f of x and y such that $f_x(x, y) = \cos(y)e^x + 2x + y^2$. Your function will have both x and y as independent variables and may also contain summands that are functions of y alone.

(b) Use the fact that $f_y(x, y) = -\sin(y)e^x + 2xy + 3$ to determine any unknown non-constant summands in your result from part (a).

(c) Complete the problem by determining the specific function f that satisfies the given conditions.

11.5 Second-Order Partial Derivatives

Motivating Questions

- Given a function f of two independent variables x and y, how are the second-order partial derivatives of f defined?
- What do the second-order partial derivatives f_{xx}, f_{yy}, f_{xy}, and f_{yx} of a function f tell us about the function's behavior?

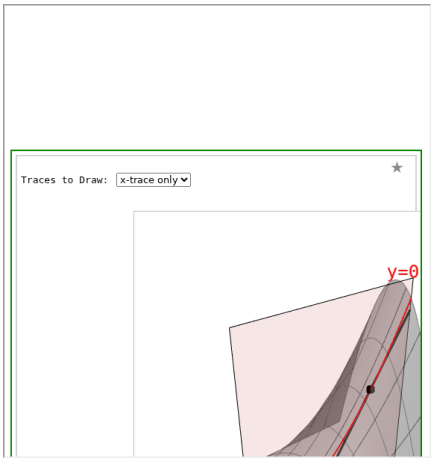
11.5.1 Introduction

Recall that for a single-variable function f, the second derivative of f is defined to be the derivative of the first derivative. That is, $f''(x) = \frac{d}{dx}\Big[f'(x)\Big] = \frac{d^2f}{dx^2}$, which can be stated in terms of the limit definition of the derivative by writing

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

In the following preview activity, we will explore how to calculate the four different second-order partial derivatives of a function of two variables as well as what these various derivatives tell us about the function's behavior.

Preview Activity 11.5.1 Once again, we consider the function f defined by $f(x,y) = \frac{x^2 \sin(2y)}{32}$ that measures a projectile's range as a function of its initial speed x and launch angle y. The graph of this function, including traces with $x = 150$ and $y = 0.6$, is shown in the interactive graph below.

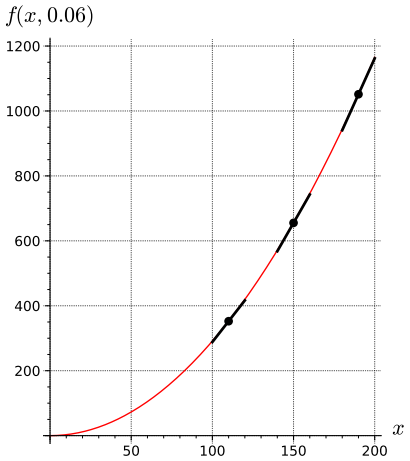


Standalone
In Context
Embed

(a) Compute the partial derivatives f_x and f_y as functions of x and y.

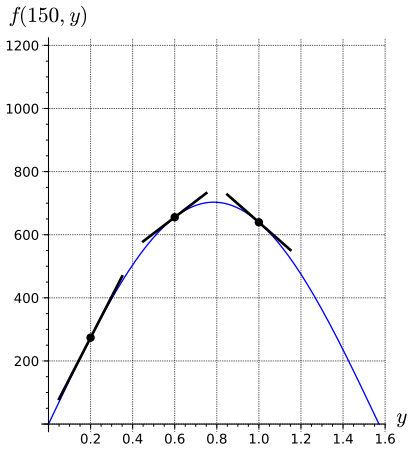
(b) Notice that f_x itself is a new function of x and y, so we may now compute the partial derivatives of f_x. Find the partial derivative $\frac{\partial}{\partial x}\left[f_x\right]$, which we will denote by f_{xx}. Verify that $f_{xx}(150, 0.6) \approx 0.058$.

(c) The graph below shows the trace of f with $y = 0.6$ with three tangent lines included. Write a few sentences to explain how $f_{xx}(150, 0.6) \approx 0.058$ is reflected in this figure.



(d) Find the partial derivative $\frac{\partial}{\partial y}\left[f_y\right]$, which we will denote by f_{yy} and compute the value of $f_{yy}(150, 0.6)$.

(e) The graph below shows the trace $f(150, y)$ and includes three tangent lines. Write a couple of sentences to explain how the value of $f_{yy}(150, 0.6)$ is reflected in this figure.



(f) Because f_x and f_y are each functions of *both* x and y, they each have *two* partial derivatives. You have already computed two of them. Now compute

$$\frac{\partial}{\partial y}\Big[f_x\Big] \qquad \text{and} \qquad \frac{\partial}{\partial x}\Big[f_y\Big]$$

for the range function $f(x, y) = \frac{x^2 \sin(2y)}{32}$ by using your earlier computations of f_x and f_y. Write a sentence to explain how you calculated these "mixed" partial derivatives.

As we saw in the preview activity, we can relate properties of second derivatives from single-variable calculus to the behavior of our trace functions. In this section, we will formalize our definitions of the second partial derivatives for a function of two variables and interpret how these give information about the functions' behavior. Later in this section, we will briefly generalize to functions of more than two variables and higher order derivatives.

11.5.2 Second-Order Partial Derivatives

A function f of two variables x and y has two first-order partial derivatives, $f_x = \frac{\partial f}{\partial x}$ and $f_y = \frac{\partial f}{\partial y}$. As we saw in Preview Activity 11.5.1, each of these first-order partial derivatives has two partial derivatives, giving a total of four *second-order* partial derivatives:

$$f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\Big[\frac{\partial f}{\partial x}\Big] = \frac{\partial^2 f}{\partial x^2}$$
$$f_{yy} = (f_y)_y = \frac{\partial}{\partial y}\Big[\frac{\partial f}{\partial y}\Big] = \frac{\partial^2 f}{\partial y^2}$$
$$f_{xy} = (f_x)_y = \frac{\partial}{\partial y}\Big[\frac{\partial f}{\partial x}\Big] = \frac{\partial^2 f}{\partial y \partial x}$$
$$f_{yx} = (f_y)_x = \frac{\partial}{\partial x}\Big[\frac{\partial f}{\partial y}\Big] = \frac{\partial^2 f}{\partial x \partial y}$$

We call f_{xx} and f_{yy} **unmixed second-order partial derivatives**. We refer to f_{xy} and f_{yx} as the **mixed second-order partial derivatives**.

You may need some practice to get accustomed to the notation for partial derivatives. By writing

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\Big[\frac{\partial f}{\partial x}\Big],$$

we mean that we first differentiate with respect to x and then differentiate with respect to y. This can be expressed in the alternate notation $f_{xy} = (f_x)_y$.

However, to find the second partial derivative

$$f_{yx} = (f_y)_x,$$

we first differentiate with respect to y and then differentiate with respect to x. This means that

$$\frac{\partial^2 f}{\partial y \partial x} = f_{xy}, \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = f_{yx}.$$

The Leibniz notation and subscript notation for a mixed partial differ in terms of the order in which x and y appear in each notation. However, in both notations, we use the idea of order of operations working from the inside out. With the Leibniz notation, that means our partial derivative operators are written *before* the function we are taking partial derivatives of, whereas with the subscript notation writes order of derivatives *after* the function.

Activity 11.5.2

(a) Find all second-order partial derivatives of the following functions. For each partial derivative you calculate, state explicitly which variable is being held constant.

(i) $f(x, y) = x^2 y^3$

(ii) $y(q, p) = p\cos(q)$

(iii) $g(s, t) = st^3 + s^4$

(b) If $h(x, y, z, t) = 9x^9 z - xyz^9 + 9t$, how many second-order partial derivatives does the function h have? Write a sentence to justify your reasoning on the number of second-order partial derivatives of h. Finally, find h_{xz} and h_{zx} (you do not need to find the other second-order partial derivatives).

In Preview Activity 11.5.1 and Activity 11.5.2, you may have noticed that the mixed second-order partial derivatives are equal. This observation holds generally and is known as Clairaut's Theorem.

Theorem 11.5.1 Clairaut's Theorem. *Let f be a function of two variables for which the partial derivatives f_{xy} and f_{yx} are continuous near the point (a, b). Then*

$$f_{xy}(a, b) = f_{yx}(a, b).$$

11.5.3 Interpreting the Second-Order Partial Derivatives

Recall from single-variable calculus, that the first derivative describes whether a function is increasing or decreasing, but the second derivative measures *how* the function is increasing or decreasing.

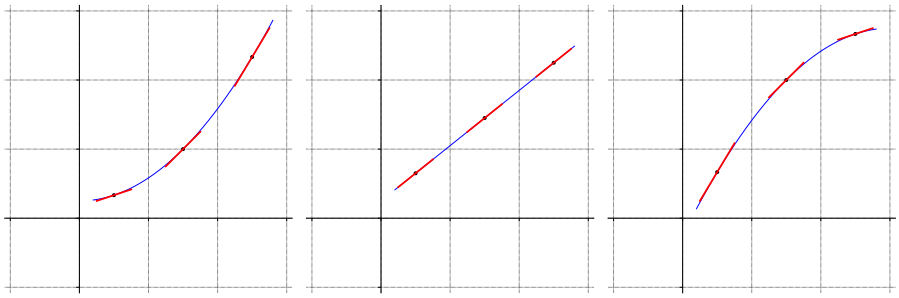


Figure 11.5.2 Plots of three functions that are increasing, but doing so at an increasing rate, a constant rate, and a decreasing rate, respectively

In Figure 11.5.2, the left curve has the slopes of the tangent lines increasing as you move along the graph from left to right, which means that the function is increasing at an increasing rate. In other words, the derivative of the slope of the tangent lines is positive, making $f''(x) > 0$. We say that this curve is concave up. The middle curve in Figure 11.5.2 has tangent lines with the same slope, which means that the function is increasing at a constant rate. Thus the derivative of the slope of the tangent lines will be zero. The right curve in Figure 11.5.2 has the slope of the tangent lines decreasing as you move along the graph from left to right, which means the function is increasing at a decreasing rate. In other words, the derivative of the slope of the tangent line is negative, so $f''(x) < 0$. We say that this curve is concave down.

A similar argument can be applied to the graphs of decreasing functions, as shown in Figure 11.5.3. In particular, we can say that the left plot will have negative slopes of tangent lines but those slopes are increasing from left to right, so $f''(x) > 0$ and the curve is concave up. The middle plot in Figure 11.5.3 has a constant slope for tangent lines along the curve, thus the second derivative is zero. The right plot in Figure 11.5.3 has negative slopes of tangent lines, corresponding to a decreasing graph, and the slopes of those tangent lines are decreasing. Thus $f''(x) < 0$ and we say that the graph is concave down.

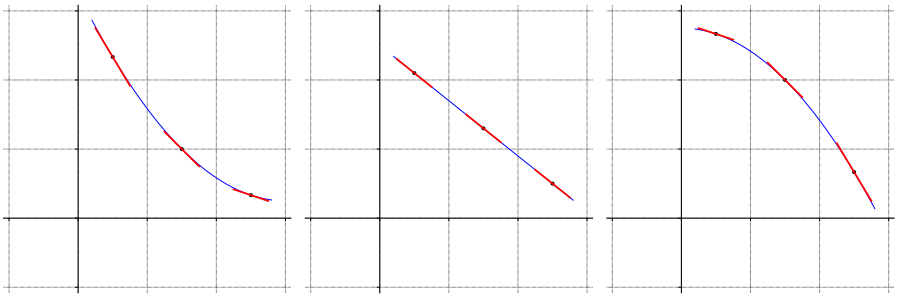


Figure 11.5.3 Plots of three functions that are decreasing, but doing so at an increasing rate, a constant rate, and a decreasing rate, respectively

Remember that $-10 < -5$ and $-5 < 2$, so an increase in negative numbers means the numbers are getting closer to zero, while a decrease in negative numbers means that values are getting "more negative" or farther from zero.

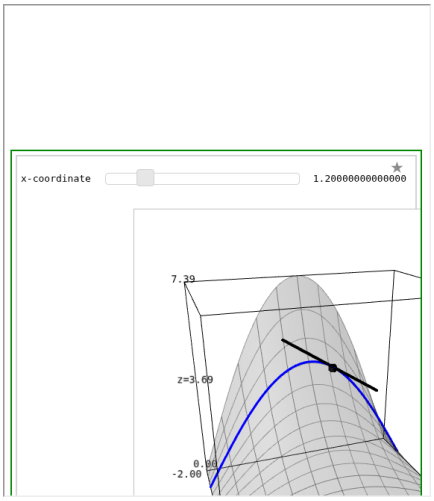
Since the unmixed second-order partial derivative f_{xx} requires us to hold y constant and differentiate twice with respect to x, we can view f_{xx} as the second derivative of a trace of f where y is fixed. As such, f_{xx} will measure the concavity of this trace.

Example 11.5.4 In this example, we will look at the second-order unmixed partial derivative of $f(x, y) = \sin(x)e^{-y}$ with respect to x and how f_{xx} measures the concavity of the surface in the directions parallel to the x-axis.

Figure 11.5.5 shows a plot of the surface defined by $z = f(x, y) = \sin(x)e^{-y}$ with the trace at $y = -1.5$ highlighted (including a tangent line). Move the slider to change the location of the tangent line shown on the trace given by $y = -1.5$. As with *all* plots in three dimensions, you should look at the axes or the coordinate values shown in the frame of the plot to make sure you understand which direction corresponds to an increase in each variable. This is especially important with interactive plots because you may have changed the perspective from the original settings in order to see the plot from a different view.

As you increase the x-coordinate where the tangent line is shown, you should notice that the slope of the tangent line decreases, and just as in single

variable calculus, this is why the trace is concave down.



Standalone
In Context
Embed

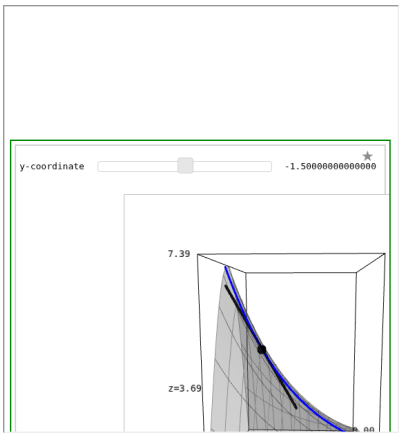
Figure 11.5.5 The surface given by $z = f(x, y) = \sin(x)e^{-y}$ with trace at $y = -1.5$ and tangent line drawn

Algebraically, we can calculate $f_x(x, y) = \cos(x)e^{-y}$ and $f_{xx}(x, y) = -\sin(x)e^{-y}$. If we consider x coordinates from 0 to π, then $f_{xx}(x, y) = -\sin(x)e^{-y} < 0$ since $-\sin(x) \leq 0$ if $0 \leq x \leq \pi$ and $e^{-y} > 0$ for all values of y, including $y = -1.5$. The other second partials of f are explored in Activity 11.5.3. ▲

In the following activity, we further explore what second-order partial derivatives tell us about the geometric behavior of a surface.

Activity 11.5.3 We will continue to work with the function f defined by $f(x, y) = \sin(x)e^{-y}$.

(a) In the interactive graphic below we see the trace of $f(x, y) = \sin(x)e^{-y}$ with x held constant with $x = 1.75$ plotted in blue. Use the slider to investigate how the slope of the tangent line changes as you vary the y-coordinate along this trace. Write a couple of sentences that describe whether the slope of the tangent lines to this curve increase or decrease as y increases along the trace $x = 1.75$. Be sure to pay attention to which direction corresponds to each coordinate increasing.



Standalone
In Context
Embed

(b) Compute $f_{yy}(x, y)$ algebraically and explain how your observations in the previous part are related to the value of $f_{yy}(1.75, y)$. Your response should address the notion of concavity. Be careful to note the directions

in which y is increasing.

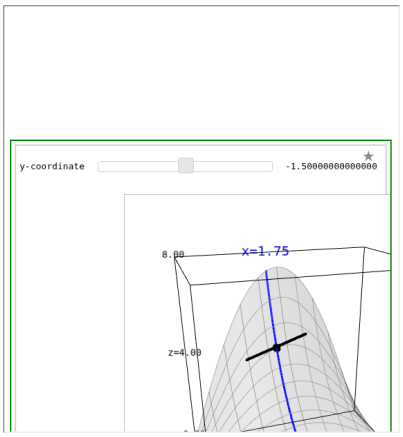
(c) We want to explore what the mixed partial derivative f_{xy} describes. We will do this by focusing on the point $(x, y) = (1.75, -1.5)$. In this part, we will work through the definition for $f_{xy}(1.75, -1.5)$ carefully.

The partial derivative $f_x(1.75, -1.5)$ measures the slope of the line tangent to the trace given by $y = -1.5$. These tangent lines change in the x-direction (parallel to the x-axis). When we consider f_{xy}, we are taking the partial derivative of f_x with respect to y, $\frac{\partial}{\partial y}\left[f_x\right]$. Thus, we are considering how the slope of the tangent line in the x-direction changes when we vary the y-coordinate a small amount. We can approximate $f_{xy}(1.75, -1.5)$ with the difference quotient

$$\frac{f_x(1.75, -1.5 + h) - f_x(1.75, -1.5)}{h}.$$

One way of visualizing this is by thinking of sliding a pencil along the trace with $x = 1.5$ so that the pencil is in the x-direction and tangent to the surface. In the interactive image below, the pencil would be the black tangent line drawn. Use the slider at the top of the interactive to change the y-coordinate of the point where the tangent line is drawn. Examine what happens to the slope of the tangent line as you increase the y-coordinate of the point of tangency by adjusting the slider. The tangent line in the x-direction at the point $(1.75, -1.5)$ is drawn in gray for reference purposes.

Based on your exploration, write a few sentences about whether $f_{xy}(1.75, -1.5)$ is positive or negative and justify your reasoning.

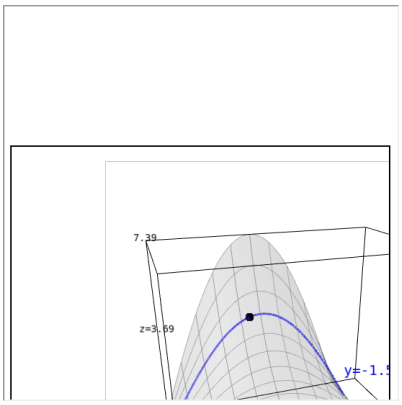


Standalone
In Context
Embed

(d) Compute $f_{xy}(x, y)$ algebraically and evaluate $f_{xy}(1.75, -1.5)$. Write a couple of sentences about how this value compares with your observations in the previous part.

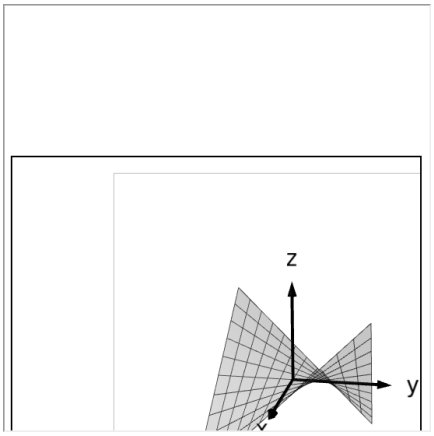
(e) We know that $f_{xx}(1.75, -1.5)$ measures the concavity of the $y = -1.5$ trace, and that $f_{yy}(1.75, -1.5)$ measures the concavity of the $x = 1.75$ trace. What do you think the quantity $f_{xy}(1.75, -1.5)$ measures?

(f) In the interactive image below, the trace with $y = -1.5$ is highlighted with the point $(1.75, -1.5, f(1.75, -1.5))$ drawn in black. Sketch three tangent lines to this trace whose slopes correspond to the value of $f_{yx}(x, -1.5)$ for three different values of x near $x = 1.75$. Use your tangent lines to state whether $f_{yx}(1.75, -1.5)$ is positive or negative. Justify your reasoning and describe what you think $f_{yx}(1.75, -1.5)$ measures.



Standalone
In Context
Embed

As we have seen in Activity 11.5.3, we may think of f_{xy} as measuring the "twist" of the graph as we increase y along a particular trace where x is held constant. In the same way, f_{yx} measures how the graph twists as we increase x. If we remember that Clairaut's Theorem tells us that since f is continuous, $f_{xy} = f_{yx}$, we see that the amount of twisting is the same in both directions. This twisting is perhaps more easily seen in Figure 11.5.6, which shows the graph of $f(x, y) = -xy$, for which $f_{xy} = -1$.



Standalone
In Context
Embed

Figure 11.5.6 The surface given by $z = f(x, y) = xy$

Just as with the first-order partial derivatives, we can approximate second-order partial derivatives in the situation where we have information about the function only at certain inputs, like with tables or contour plots.

Activity 11.5.4 As we saw in Activity 11.4.5, the wind chill $w(v, T)$, in degrees Fahrenheit, is a function of the wind speed, in miles per hour, and the air temperature, in degrees Fahrenheit. Some values of the wind chill are recorded in the table below.

$v \backslash T$	-20	-15	-10	-5	0	5	10
10	-41	-35	-28	-22	-16	-10	-4
15	-45	-39	-32	-26	-19	-13	-7
20	-48	-42	-35	-29	-22	-15	-9
25	-51	-44	-37	-31	-24	-17	-11
30	-53	-46	-39	-33	-26	-19	-12
35	-55	-48	-41	-34	-27	-21	-14

(a) Estimate the partial derivatives $w_T(20, -15)$, $w_T(20, -10)$, and $w_T(20, -5)$. Use these results to estimate the second-order partial $w_{TT}(20, -10)$.

(b) In a similar way, estimate the second-order partial $w_{vv}(20, -10)$.

(c) Estimate the partial derivatives $w_T(20, -10)$, $w_T(25, -10)$, and $w_T(15, -10)$, and use your results to estimate the partial $w_{Tv}(20, -10)$.

(d) In a similar way, estimate the partial derivative $w_{vT}(20, -10)$.

(e) Write a few sentences to explain what the values $w_{TT}(20, -10)$, $w_{vv}(20, -10)$, and $w_{Tv}(20, -10)$ indicate regarding the behavior of $w(v, T)$.

Summary

- There are four second-order partial derivatives of a function f of two variables x and y:

$$f_{xx} = (f_x)_x, \quad f_{xy} = (f_x)_y, \quad f_{yx} = (f_y)_x, \quad \text{and} \quad f_{yy} = (f_y)_y.$$

- The unmixed second-order partial derivatives f_{xx} and f_{yy} tell us about the concavity of the traces. The mixed second-order partial derivatives f_{xy} and f_{yx} tell us how the graph of f twists.

11.5.4 Exercises

1. Calculate all four second-order partial derivatives of $f(x, y) = 2x^2y + 4xy^3$.

$f_{xx}(x, y) =$ _____
$f_{xy}(x, y) =$ _____
$f_{yx}(x, y) =$ _____
$f_{yy}(x, y) =$ _____

2. Find all the first and second order partial derivatives of $f(x, y) = -4\sin(2x + y) + 7\cos(x - y)$.

A. $\frac{\partial f}{\partial x} = f_x =$ _____
B. $\frac{\partial f}{\partial y} = f_y =$ _____
C. $\frac{\partial^2 f}{\partial x^2} = f_{xx} =$ _____
D. $\frac{\partial^2 f}{\partial y^2} = f_{yy} =$ _____
E. $\frac{\partial^2 f}{\partial x \partial y} = f_{yx} =$ _____
F. $\frac{\partial^2 f}{\partial y \partial x} = f_{xy} =$ _____

3. Find the partial derivatives of the function

$$f(x, y) = xye^{-1y}$$

$f_x(x, y) =$ _____
$f_y(x, y) =$ _____
$f_{xy}(x, y) =$ _____
$f_{yx}(x, y) =$ _____

4. Calculate all four second-order partial derivatives of $f(x, y) = \sin\left(\frac{4x}{y}\right)$.

$f_{xx}(x, y) =$ _____
$f_{xy}(x, y) =$ _____
$f_{yx}(x, y) =$ _____
$f_{yy}(x, y) =$ _____

5. Calculate all four second-order partial derivatives and check that $f_{xy} = f_{yx}$. Assume the variables are restricted to a domain on which the function is defined.

$$f(x,y) = e^{4xy}$$

$f_{xx} =$ _____
$f_{yy} =$ _____
$f_{xy} =$ _____
$f_{yx} =$ _____

6. Calculate all four second-order partial derivatives of $f(x,y) = (5x+3y)\,e^{y}$.

$f_{xx}\,(x,y) =$ _____
$f_{xy}\,(x,y) =$ _____
$f_{yx}\,(x,y) =$ _____
$f_{yy}\,(x,y) =$ _____

7. Let $f(x,y) = (4x-y)^5$. Then

$\frac{\partial^2 f}{\partial x \partial y} =$ _____
$\frac{\partial^3 f}{\partial x \partial y \partial x} =$ _____
$\frac{\partial^3 f}{\partial x^2 \partial y} =$ _____

8. If $z_{xy} = 9y$ and all of the second order partial derivatives of z are continuous, then

(a) $z_{yx} =$ _____
(b) $z_{xyx} =$ _____
(c) $z_{xyy} =$ _____

9. If $z = f(x) + yg(x)$, what can we say about z_{yy}?

◯ $z_{yy} = g(x)$

◯ $z_{yy} = y$

◯ $z_{yy} = z_{xx}$

◯ $z_{yy} = 0$

◯ We cannot say anything

10. Assume z is a smooth function of x and y. If $z_{xy} = -4x$, what can you say about each of the following?

(a) $z_{yx} =$ _____
(b) $z_{xyx} =$ _____
(c) $z_{xyy} =$ _____

11. Shown in Figure 11.5.7 is a contour plot of a function f with the values of f labeled on the contours. The point $(2,1)$ is highlighted in red.

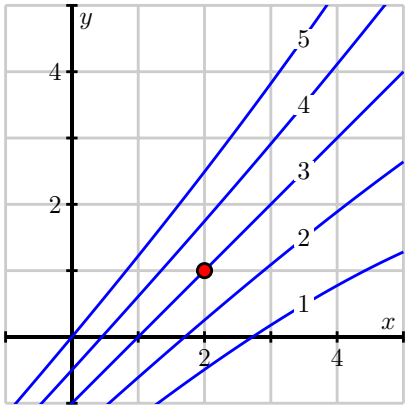


Figure 11.5.7 A contour plot of $f(x, y)$.

(a) Estimate the partial derivatives $f_x(2, 1)$ and $f_y(2, 1)$.

(b) Determine whether the second-order partial derivative $f_{xx}(2, 1)$ is positive or negative, and explain your thinking.

(c) Determine whether the second-order partial derivative $f_{yy}(2, 1)$ is positive or negative, and explain your thinking.

(d) Determine whether the second-order partial derivative $f_{xy}(2, 1)$ is positive or negative, and explain your thinking.

(e) Determine whether the second-order partial derivative $f_{yx}(2, 1)$ is positive or negative, and explain your thinking.

(f) Consider a function g of the variables x and y for which $g_x(2, 2) > 0$ and $g_{xx}(2, 2) < 0$. Sketch possible behavior of some contours around $(2, 2)$ on the left axes in Figure 11.5.8.

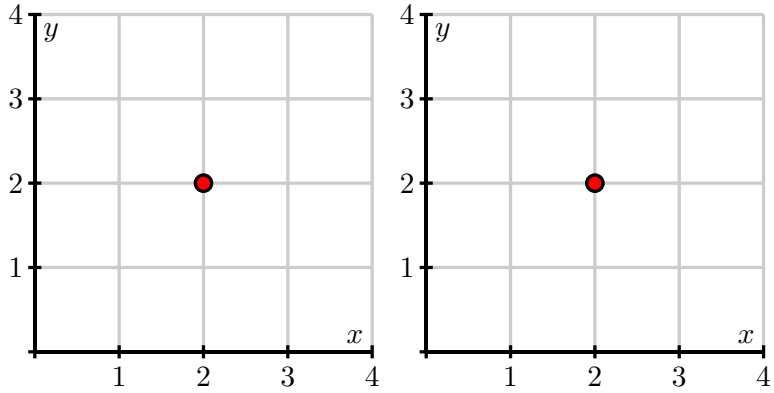


Figure 11.5.8 Plots for contours of g and h.

(g) Consider a function h of the variables x and y for which $h_x(2, 2) > 0$ and $h_{xy}(2, 2) < 0$. Sketch possible behavior of some contour lines around $(2, 2)$ on the right axes in Figure 11.5.8.

12. The Heat Index, I, (measured in *apparent degrees F*) is a function of the actual temperature T outside (in degrees F) and the relative humidity H (measured as a percentage). A portion of the table which gives values for this function, $I(T, H)$, is reproduced in Table 11.5.9.

Table 11.5.9 Heat index.

$T \downarrow \setminus H \rightarrow$	70	75	80	85
90	106	109	112	115
92	112	115	119	123
94	118	122	127	132
96	125	130	135	141

(a) State the limit definition of the value $I_{TT}(94, 75)$. Then, estimate $I_{TT}(94, 75)$, and write one complete sentence that carefully explains the meaning of this value, including units.

(b) State the limit definition of the value $I_{HH}(94, 75)$. Then, estimate $I_{HH}(94, 75)$, and write one complete sentence that carefully explains the meaning of this value, including units.

(c) Finally, do likewise to estimate $I_{HT}(94, 75)$, and write a sentence to explain the meaning of the value you found.

13. The temperature on a heated metal plate positioned in the first quadrant of the xy-plane is given by

$$C(x, y) = 25e^{-(x-1)^2-(y-1)^3}.$$

Assume that temperature is measured in degrees Celsius and that x and y are each measured in inches.

(a) Determine $C_{xx}(x, y)$ and $C_{yy}(x, y)$. Do not do any additional work to algebraically simplify your results.

(b) Calculate $C_{xx}(1.1, 1.2)$. Suppose that an ant is walking past the point $(1.1, 1.2)$ along the line $y = 1.2$. Write a sentence to explain the meaning of the value of $C_{xx}(1.1, 1.2)$, including units.

(c) Calculate $C_{yy}(1.1, 1.2)$. Suppose instead that an ant is walking past the point $(1.1, 1.2)$ along the line $x = 1.1$. Write a sentence to explain the meaning of the value of $C_{yy}(1.1, 1.2)$, including units.

(d) Determine $C_{xy}(x, y)$ and hence compute $C_{xy}(1.1, 1.2)$. What is the meaning of this value? Explain, in terms of an ant walking on the heated metal plate.

14. Let $f(x, y) = 8 - x^2 - y^2$ and $g(x, y) = 8 - x^2 + 4xy - y^2$.

(a) Determine f_x, f_y, f_{xx}, f_{yy}, f_{xy}, and f_{yx}.

(b) Evaluate each of the partial derivatives in (a) at the point $(0, 0)$.

(c) What do the values in (b) suggest about the behavior of f near $(0, 0)$? Plot a graph of f and compare what you see visually to what the values suggest.

(d) Determine g_x, g_y, g_{xx}, g_{yy}, g_{xy}, and g_{yx}.

(e) Evaluate each of the partial derivatives in (d) at the point $(0, 0)$.

(f) What do the values in (e) suggest about the behavior of g near $(0, 0)$? Plot a graph of g and compare what you see visually to what the values suggest.

(g) What do the functions f and g have in common at $(0,0)$? What is different? What do your observations tell you regarding the importance of a certain second-order partial derivative?

15. Let $f(x,y) = \frac{1}{2}xy^2$ represent the kinetic energy in Joules of an object of mass x in kilograms with velocity y in meters per second. Let (a,b) be the point $(4,5)$ in the domain of f.

(a) Calculate $\frac{\partial^2 f}{\partial x^2}$ at the point (a,b). Then explain as best you can what this second-order partial derivative tells us about kinetic energy.

(b) Calculate $\frac{\partial^2 f}{\partial y^2}$ at the point (a,b). Then explain as best you can what this second-order partial derivative tells us about kinetic energy.

(c) Calculate $\frac{\partial^2 f}{\partial y \partial x}$ at the point (a,b). Then explain as best you can what this second-order partial derivative tells us about kinetic energy.

(d) Calculate $\frac{\partial^2 f}{\partial x \partial y}$ at the point (a,b). Then explain as best you can what this second-order partial derivative tells us about kinetic energy.

11.6 Linearization: Tangent Planes and Differentials

Motivating Questions

- What does it mean for a function of two variables to be locally linear at a point?
- How do we find the equation of the plane tangent to a locally linear function at a point?
- How can a linearization be used to approximate the output of a multivariable function?

11.6.1 Introduction

Throughout single-variable calculus, the tangent line serves as a geometric and conceptual tool to understand the derivative at a point. While the value of a derivative at a point measures the instantaneous rate of change, we visualize the value of the derivative as the slope of the line tangent to the function's graph. This allowed us to "see" the derivative's value in terms of an important behavior of the function's graph.

In Subsection 10.3.4, we saw that a curve in space, given as the graph of a vector-valued function of one variable, is locally linear and approximated well by a tangent line provided that the function is differentiable. So the vector-valued function of one variable that describes a curve in space will be differentiable if the curve looks like a line on a small scale.

In this section, we will explore the various ways in which we can extend to functions of two variables the related ideas of differentiability and local linearity. Geometrically, this means we will use a tangent plane as a "flat" analog for surfaces of the form $z = f(x,y)$.[1] We will also introduce a couple of new tools

[1] For graphs in the cartesian plane, there was only one type of flat graph: a line. When we generalized to higher dimensions, we saw how both lines and planes can be considered "flat" graphs in three dimensions. However, we noted that lines are 1-dimensional flat graphs because there is only direction to move along the graph. On the other hand, planes are

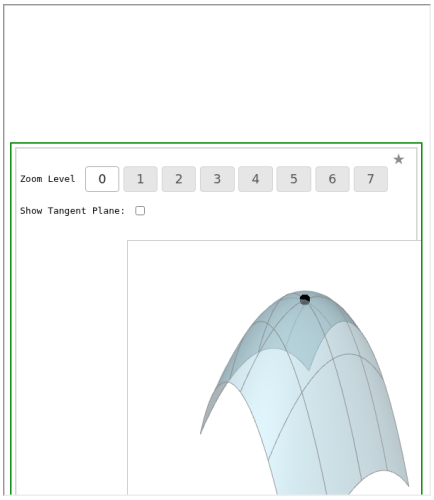
to help us algebraically and numerically describe the change in the output of $f(x, y)$.

To start our investigation, let's consider the graph of a "nice" two-variable function on a small scale. We will look at

$$f(x, y) = 6 - \frac{x^2}{2} - y^2,$$

whose graph is shown in Figure 11.6.1.

You should examine the behavior of f near the input $(x_0, y_0) = (1, 1)$, which has output of $f(1, 1) = 4.5$. You can change how closely Figure 11.6.1 shows the graph of $z = f(x, y)$ on smaller and smaller scales around the point $(1, 1, 4.5)$. With zoom set to 0, the plot clearly reveals that the surface is curved. However, once you increase the zoom level to 3 or higher, the surface looks more like a tilted plane than a curved surface.



Standalone
In Context
Embed

Figure 11.6.1 A plot of $f(x, y) = 6 - \frac{x^2}{2} - y^2$ with the point $(1, 1, 4.5)$ highlighted and the option to zoom in around the highlighted point.

Just as the graph of a differentiable single-variable function or a differentiable vector-valued function of one variable looks like a line when viewed on a small scale, the graph of this two-variable function looks like a plane when viewed on a small scale. We will call functions that have this property **locally linear**. The Preview Activity guides you through finding an equation for this plane, but we first state a key idea that will be a helpful algebraic form to use when thinking about tangent planes to surfaces of the form $z = f(x, y)$.

Key Idea 11.6.2 *A non-vertical plane that goes through the point (x_0, y_0, z_0) can be expressed in the form $z = z_0 + a(x - x_0) + b(y - y_0)$, where a and b are constants.*

Proof. As we saw in Section 9.7, the equation of a plane passing through the point (x_0, y_0, z_0) may be written in the form $A(x-x_0)+B(y-y_0)+C(z-z_0) = 0$. If the plane is not vertical, then $C \neq 0$, and we can rearrange this equation and write $C(z - z_0) = -A(x - x_0) - B(y - y_0)$ and thus

$$\begin{aligned} z &= z_0 - \frac{A}{C}(x - x_0) - \frac{B}{C}(y - y_0) \\ &= z_0 + a(x - x_0) + b(y - y_0) \end{aligned}$$

where $a = -A/C$ and $b = -B/C$, respectively.

2-dimensional flat graphs.

After our work in the Preview Activity, we will look at how to obtain this equation in terms of the normal vector of this plane that best describes the surface at a particular point. ■

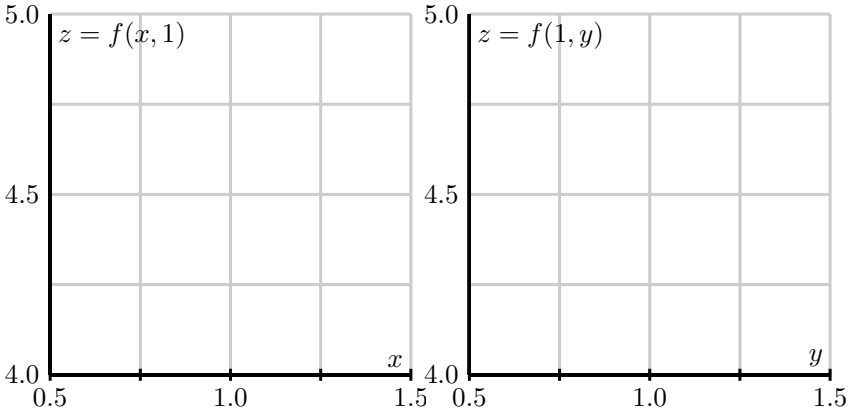
Preview Activity 11.6.1 We want to find the equation of the plane, using the form given in Key Idea 11.6.2, that best describes the surface given by $z = f(x, y) = 6 - \frac{x^2}{2} - y^2$ for input values around $(x_0, y_0) = (1, 1)$. In particular, we will need to find how the values of z_0, a, and b are related to $f(x, y)$.

(a) Find $f(1, 1)$, $f_x(1, 1)$, and $f_y(1, 1)$.

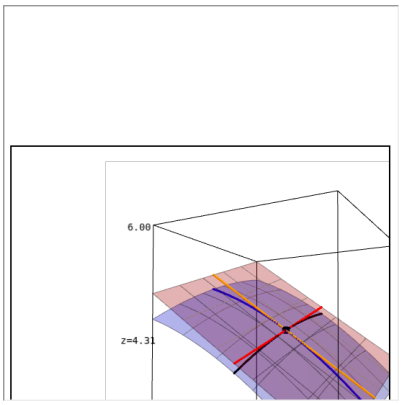
(b) We want our plane to match the height of the surface at $(1, 1)$. For this to happen, the value z_0 from Key Idea 11.6.2 must be the z-coordinate value where the plane intersects the surface. What is the z-coordinate of the point where the tangent plane and the surface intersect?

(c) For the plane to have the same behavior as the surface $z = f(x, y)$ near $(1, 1)$, the plane must match the behavior of the traces given by $x = 1$ and $y = 1$ near this point.

Sketch the traces of $f(x, y) = 6 - \frac{x^2}{2} - y^2$ for $y = y_0 = 1$ and $x = x_0 = 1$ below. Draw the tangent lines to the each of the traces when the appropriate input is 1.



(d) What are the slopes of the tangent lines to the traces that you drew in the previous part? Write a couple of sentences to explain why the tilt of the tangent plane in the x-direction is given by the partial derivative f_x and the tilt of the tangent plane in the y-direction is given by the partial derivative f_y. Your answer should discuss how each of these slopes/partial derivatives relates to the traces of the surface and the plane illustrated below. You should identify whether the red/orange traces along the plane and the black/blue traces along the surface correspond to constant values of x or y.



Standalone
In Context
Embed

(e) Fill in the blanks below with the proper values to give the tangent plane to the graph of $f(x,y) = 6 - x^2/2 - y^2$ at the point $(x_0, y_0) = (1,1)$.

$$z = z_0 + a(x - x_0) + b(y - y_0) = \quad + \quad (x - \quad) + \quad (y - \quad).$$

11.6.2 The Tangent Plane

As the preview activity suggests, the graph of most functions we have encountered will be approximated well by a plane tangent to the graph at a point of interest. In Subsection 11.6.4, we will talk more about the technical conditions for the tangent plane to be a good approximation to the surface, but for now we will use the following general formula for the tangent plane from our final result of Preview Activity 11.6.1.

The Tangent Plane.

If f is a function of x and y for which both f_x and f_y exist and are continuous in an open disk containing the point (x_0, y_0), then the equation of the plane tangent to the graph of f at the point $(x_0, y_0, f(x_0, y_0))$ is

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \qquad (11.6.1)$$

Having the equation for a tangent plane in the form given by $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$ allows us to quickly identify important information about the function f at the point (x_0, y_0). For example, if a function f has a tangent plane given by $z = 7 - 2(x - 3) + 4(y + 1)$, then we can immediately read the following information from the given form of the tangent plane equation:

- $f(3, -1) = 7$,
- $f_x(3, -1) = -2$ is the slope of the trace to both f and the tangent plane in the x-direction at $(3, -1)$, and
- $f_y(3, -1) = 4$ is the slope of the trace of both f and the tangent plane in the y-direction at $(3, -1)$.

Activity 11.6.2

(a) Find the equation of the tangent plane to $f(x,y) = x^2 y$ at the point $(1, 2)$.

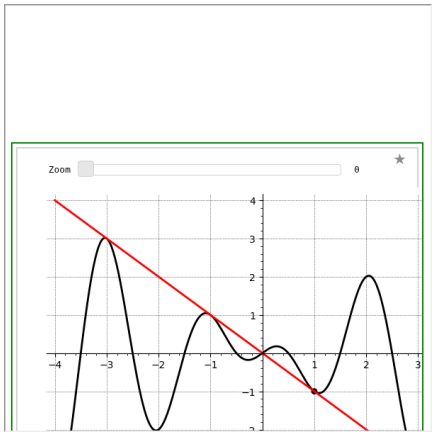
(b) Suppose that the tangent plane to the graph of a continuously differen-

tiable function $z = g(x, y)$ is given in the form

$$z = 5 - 3(x + 2) + (y - 3).$$

Use the equation of the tangent plane to identify a point on the graph as well as a value of g_x and a value of g_y. Be sure to identify at what point(s) you have found the values of the partial derivatives.

In single-variable calculus, an important use of the tangent line is to approximate a differentiable function. Near the point x_0, the tangent line to the graph of f at x_0 is close to the graph of f for input values close to x_0, as shown in Figure 11.6.3. Adjust the Zoom slider to examine a small region around the highlighted point to see how the tangent line is a good approximation for a small region around that point. In fact, the smaller the scale around the point, the better the approximation is.



Standalone
In Context
Embed

Figure 11.6.3 A plot of a locally linear function f

In a similar way, we say that a two-variable function f **is locally linear near** (x_0, y_0) provided that the graph of f looks like a plane (its **tangent plane**) when viewed on a small scale near (x_0, y_0). Determining when a function of two variables is locally linear at a point involves more nuance than in the single-variable setting. In Subsection 11.6.4, we discuss the technical details and present examples of functions that are not locally linear.

The tangent plane is a geometric way of visualizing and approximating a surface of the form $z = f(x, y)$ near the point of tangency. In the next subsection, we will discuss the linearization, a functional form of the tangent plane that facilitates approximations near the point of tangency.

11.6.3 Linearization

In single-variable calculus, we let L denote the function whose graph is the tangent line, and thus

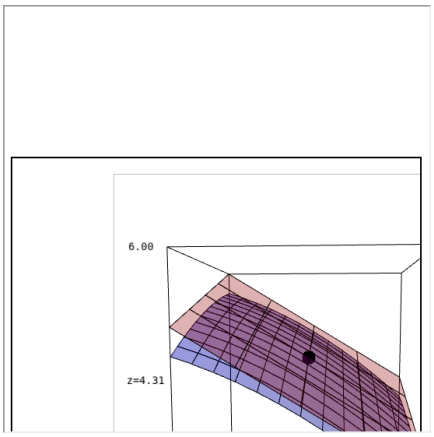
$$L(x) = f(x_0) + f'(x_0)(x - x_0).$$

Furthermore, $f(x) \approx L(x)$ near x_0. We call L the **linearization** of f.

Similarly, the tangent plane to the graph of $f(x, y)$, a locally linear function of two variables, at a point (x_0, y_0) provides a good approximation near (x_0, y_0). We define the **linearization**, L, to be the two-variable function whose graph is the tangent plane. Thus,

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \qquad (11.6.2)$$

is the linearization of f at (x_0, y_0). Note that $f(x, y) \approx L(x, y)$ for points near (x_0, y_0), as illustrated in Figure 11.6.4.



Standalone
In Context
Embed

Figure 11.6.4 A plot of a locally linear function $f(x, y)$

Using a function's linearization when you already have an algebraic expression for the function provides little added value. In many applications the function of interest does not have a known algebraic formula. For instance, if you were working for a mining company and were trying to map a pocket of some resource underground, it would be *very* expensive to drill a bunch of samples to make a chart with amounts of the resource at a comprehensive grid of locations. Instead, you could be more selective in obtaining data via other sampling techniques and then use linearization to make estimates for intermediate locations. Linearization is a valuable *first* step in estimating values between different input locations. In later math courses, you may encounter more sophisticated ways to estimate between data points and discuss the advantages and drawbacks of these ideas.

Example 11.6.5 In this example, we will give the equation of the tangent plane to $f(x, y) = 6 - \frac{x^2}{2} - y^2$ at the point $(1, 1)$, state the associated linearization, and use this linearization to estimate the output of f for a nearby input.

The partial derivatives of f are $f_x = -x$ and $f_y = -2y$. This gives the values $f_x(1, 1) = -1$ and $f_y(1, 1) = -2$. Using $f(1, 1) = 4.5$, the tangent plane to f at the point $(1, 1)$ is

$$z = 4.5 + (-1)(x - 1) + (-2)(y - 1).$$

The linearization of f at $(1, 1)$ is

$$L(x, y) = 4.5 + (-1)(x - 1) - 2(y - 1).$$

Notice that the same information is used to find the tangent plane and the linearization. Generally speaking, we consider the tangent plane to be a geometric tool, while the linearization is a function or algebraic tool. However, both describe f near $(1, 1)$. We can use the linearization to estimate $f(0.9, 1.2)$:

$$\begin{aligned} f(0.9, 1.2) \approx L(0.9, 1.2) &= 4.5 + (-1)(0.9 - 1) + (-2)(1.2 - 1) \\ &= 4.5 + (0.1) - 0.4 = 4.2. \end{aligned}$$

▲

Because we had an algebraic expression for f in the previous example, we

could have evaluated $f(0.9, 1.2)$ directly. However, as the next activity shows, many applications of multivariable functions use tables or graphs where it is not possible to have the function's output at every possible input.

Activity 11.6.3 In this activity, we will find the linearization of several different functions that are given in algebraic, tabular, or graphical form.

(a) Find the linearization $L(x, y)$ for the function g defined by

$$g(x, y) = \frac{x}{x^2 + y^2}$$

at the point $(1, 2)$. Use the linearization to estimate the value of $g(0.8, 2.3)$.

(b) The table below provides a collection of values of the wind chill $w(v, T)$, in degrees Fahrenheit, as a function of wind speed, in miles per hour, and temperature, also in degrees Fahrenheit.

$v \backslash T$	-20	-15	-10	-5	0	5	10
10	-41	-35	-28	-22	-16	-10	-4
15	-45	-39	-32	-26	-19	-13	-7
20	-48	-42	-35	-29	-22	-15	-9
25	-51	-44	-37	-31	-24	-17	-11
30	-53	-46	-39	-33	-26	-19	-12
35	-55	-48	-41	-34	-27	-21	-14

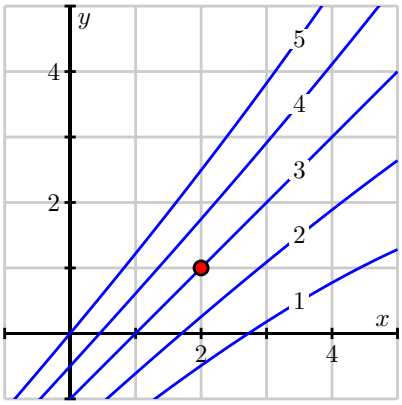
i Use the data to estimate the appropriate partial derivatives of $w(v, T)$ at the point $(20, -10)$.

ii Then find the linearization $L(v, T)$ at the point $(20, -10)$.

iii Use the linearization to estimate $w(10, -10)$, $w(20, -12)$, and $w(18, -12)$.

iv Compare your results to what you obtained in Activity 11.4.5.

(c) The image below gives a contour plot of a continuously differentiable function f. After estimating appropriate partial derivatives, determine the linearization $L(x, y)$ at the point $(2, 1)$, and use it to estimate $f(2.2, 1)$, $f(2, 0.8)$, and $f(2.2, 0.8)$.



Example 11.6.6 Suppose we have a machine that cuts rectangles of width $w = 20$ cm and height $h = 10$ cm. However, the machine isn't perfect, and

therefore the width could be off by $dw = \Delta w = 0.2$ cm and the height could be off by $dh = \Delta h = 0.4$ cm.

The area of the rectangle is

$$A(w, h) = wh,$$

which means that the area of a perfectly-manufactured rectangle is $A(20, 10) = 200$ square centimeters. Since the machine isn't perfect, we would like to know how much the area of a given rectangle could differ from the perfect rectangle. We will estimate the uncertainty in the area using steps in each variable along the linearization to estimate the change in the area. In particular, this looks like

$$\Delta A \approx A_w(20, 10)\ \ \Delta w + A_h(20, 10)\ \ \Delta h$$

By computing partial derivatives, we find that $A_w = h$ and $A_h = w$. This implies that

$$\Delta A \approx 10\ \ \Delta w + 20\ \ \Delta h = 10(0.2) + 20(0.4) = 10.$$

That is, we estimate that the area in our rectangles could be off by as much as 10 square centimeters. ▲

Activity 11.6.4

(a) Suppose that the elevation of a plot of land is given by the function h, where we additionally know that $h(3, 1) = 4.35$, $h_x(3, 1) = 0.27$, and $h_y(3, 1) = -0.19$. Assume that x and y are measured in miles in the east and north directions, respectively, from $(0, 0)$.

Your GPS device says that you are currently at the point $(3, 1)$. However, you know that the coordinates are only accurate to within 0.2 miles; that is, $dx = \Delta x = 0.2$ and $dy = \Delta y = 0.2$. Estimate the uncertainty in your elevation using the linearization of h for the input (3,1).

(b) The pressure, volume, and temperature of an ideal gas are related by the equation

$$P = P(T, V) = 8.31\frac{T}{V},$$

where P is measured in kilopascals, V in liters, and T in kelvin. Find the pressure when the volume is 12 liters and the temperature is 310 K. Use the linearization of P at the point $(310, 12)$ to estimate the change in the pressure when the volume increases to 12.3 liters and the temperature decreases to 305 K.

(c) Use the table of values for the wind chill $w(v, T)$, in degrees Fahrenheit, as a function of temperature, also in degrees Fahrenheit, and wind speed, in miles per hour provided in part 11.6.3.b for this part. Suppose your anemometer[2] says the wind is blowing at 25 miles per hour and your thermometer[3] shows a reading of $-15°$ degrees. However, you know your thermometer is only accurate to within 2° degrees and your anemometer is only accurate to within 3 miles per hour. What is the wind chill based on your measurements? Estimate the uncertainty in your measurement of the wind chill.

In Example 11.6.6, we see how the same form is being used to approximate the change in the output of a multivariable function across very different input

[2] An instrument for measuring wind speed.
[3] An instrument for measuring the temperature.

values. The total differential is an abstraction of this form, namely that the change in the output of a function of two variables is a linear combination of the changes in the inputs with the coefficients being the values of the partial derivatives.

The change in the output of our function can be approximated at a particular input point (x_0, y_0) using

$$\Delta f \approx f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y \tag{11.6.3}$$

We can use the following abstraction of the linearization idea to express the form used to measure changes in output of a multivariable function based on small changes in the input

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{11.6.4}$$

We call the quantities dx, dy, and df **differentials**, and we think of them as measuring small changes in the quantities x, y, and f. In fact, Equation (11.6.4) is called the total differential of f.

The linearization of a function approximates the function's output around a particular point, but a total differential, df, shows the algebraic form corresponding to a change in the output of f in terms of linear combinations of the changes in inputs (not just the values at a particular point).

For example, we can apply our differential idea to the example of cutting boxes (Example 11.6.6) and estimate the area estimate error of a box of a different size as well. If we wanted to make rectangles that are 63 cm wide and 25 cm high and our machines had the same tolerances ($dw = \Delta w = 0.2$ cm and $dh = \Delta h = 0.4$ cm), then our target rectangle would have area 1,575 square centimeters but the uncertainty in the area would be 30.2 cm^2:

$$\begin{aligned}\Delta A \approx dA =& A_w(63, 25)\ dw + A_h(63, 25)\ dh \\ =& 25dw + 63dh \\ =& 25(0.2) + 63(0.4) = 30.2\end{aligned}$$

We close this subsection by highlighting an idea that will be used time and again in this chapter. For a "nice" function (e.g., locally linear, differentiable, or continuously differentiable), the rate of change in the output of the function near a specific point is the same as the rate of change along the linearization. In other words, we can approximate change near a point on a surface by looking at the change on the linearization. We can also evaluate the rates of change along our surface by evaluating the corresponding rate of change on the linearization. Regardless of the direction, these rates of change will be the same at the specified point.

11.6.4 Differentiability and Local Linearity

In our equation for a tangent plane, we gave the general formula for the equation of the tangent plane at a particular point on the graph of $z = f(x, y)$, but we had a condition involving partial derivatives that we needed to satisfy in order for the tangent plane to make sense. We will briefly discuss some ways in which a function can fail to be locally linear; in other words, we will look at several examples of functions of two variables where the tangent plane is either not defined or is ***not*** a good approximation of the surface near a particular point.

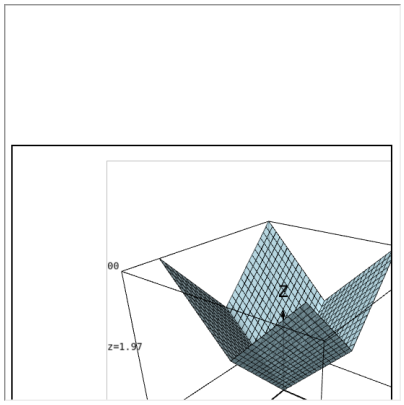
In single-variable calculus, one of the first functions you studied that was continuous but had a point where the graph was not locally linear was the

absolute value function $f(x) = |x|$. When zooming in on the graph of f near the point $(0, 0)$, the graph maintains a sharp corner, and thus f is not locally linear at the point on its graph with $x = 0$.

Remember that a two-variable function f **is locally linear near** (x_0, y_0) provided that the graph of f looks like a plane (its **tangent plane**) when viewed on a small scale near (x_0, y_0). Determining when a function of two variables is locally linear at a point involves more nuance than in the single-variable setting, as the next example illustrates.

Example 11.6.7

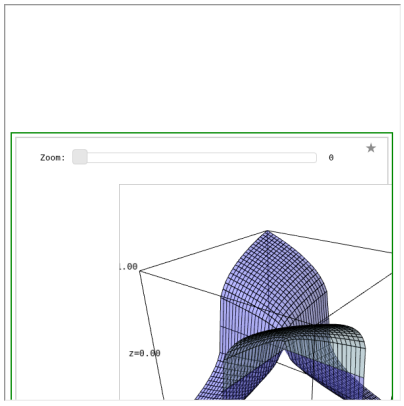
(a) The function $f(x, y) = |x| + |y|$ is graphed in Figure 11.6.8. No matter how much you zoom in around the origin, the surface will not look like a plane. You can also see this if you zoom in near any point on the surface with either $x = 0$ or $y = 0$. While this function is continuous at every point, neither partial derivative will be defined when either x or y is zero.



Standalone
In Context
Embed

Figure 11.6.8 A plot of $f(x, y) = |x| + |y|$

(b) Your experience from single-variable calculus may lead you to reasonably expect that if $f_x(a, b)$ and $f_y(a, b)$ both exist at a point (a, b), then f is locally linear at (a, b). ***This is not sufficient for multivariable functions, however.*** To illustrate this, consider the function defined by $f(x, y) = x^{1/3}y^{1/3}$. You can see from the figure below that as you zoom in around the origin, the graph does not flatten out and look like a plane. Exercise 24 guides you through using the limit definition of the partial derivative to show that $f_x(0, 0)$ and $f_y(0, 0)$ both exist, but that f is not locally linear at $(0, 0)$.



Standalone
In Context
Embed

Figure 11.6.9 A plot of $f(x, y) = x^{1/3}y^{1/3}$

▲

A precise discussion of differentiability of functions of more than one variable is beyond the scope of this text, but if you are interested in exploring these ideas a bit more, see Exercise 28. We will be content to define a stronger, but more easily verified, set of conditions called continuous differentiability that ensure local linearity. For our purposes in this text, continuous differentiability is the only condition we will need to use. All of the results we encounter will apply to differentiable functions, and so also apply to continuously differentiable functions.

Differentiability.

If f is a function of x and y for which both f_x and f_y exist and are continuous in an open disk containing the point (x_0, y_0), then f is **continuously differentiable** at (x_0, y_0).

If a function f of two variables is continuously differentiable at a point (x_0, y_0), then f has a tangent plane at (x_0, y_0). This means that when viewed up close, the tangent plane and the graph of the function are virtually indistinguishable. You can see this illustrated in Figure 11.6.1 if you enable the Show Tangent Plane check box.

Important Note: As discussed before, the function $f(x, y) = x^{1/3}y^{1/3}$, which we saw in Figure 11.6.9, is not locally linear as we zoom in near the point $(0, 0)$. However, as Exercise 24 demonstrates, both $f_x(0, 0)$ and $f_y(0, 0)$ exist for this function. Thus, we can define the plane $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$. However, this plane is *not* "tangent" to the graph in any reasonable sense. This is why we have specified that we require that f be continuously differentiable at a point in order to discuss the tangent plane at that point. The function $f(x, y) = x^{1/3}y^{1/3}$ is not continuously differentiable (or differentiable) at $(0, 0)$! Differentiability for a function of two variables implies the existence of a tangent plane. However, the existence of the two first-order partial derivatives of a function at a point does *not* imply differentiability. This is quite different from single-variable calculus, where the derivative existing at a point is equivalent to the function being locally linear at that point.

Summary

- A function f of two independent variables is locally linear at a point (x_0, y_0) if the graph of $z = f(x, y)$ looks like a plane as we zoom in on the graph (around the point (x_0, y_0)). In this case, the equation of the tangent plane is given by

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

- The function $L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$ is called the linearization of a differentiable function f at (x_0, y_0) and may be used to estimate values of $f(x, y)$; that is, $f(x, y) \approx L(x, y)$ for points (x, y) near (x_0, y_0). The graph of the linearization function is the tangent plane.

- A function f of two independent variables is (continuously) differentiable at (x_0, y_0) provided that both f_x and f_y exist and are continuous in an open disk containing the point (x_0, y_0).

- The differential df of a function $f = f(x, y)$ is related to the differentials

dx and dy by

$$df = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy.$$

We can use this relationship to approximate small changes in the output of f that result from small changes in x and y.

11.6.5 Exercises

1. Find the linearization $L(x, y)$ of the function $f(x, y) = \sqrt{88 - 4x^2 - 9y^2}$ at $(3, -2)$.

$L(x, y) =$ _____

Note: Your answer should be an expression in x and y; e.g. "3x - 5y + 9"

2. Find the equation of the tangent plane to the surface $z = e^{-1x/17} \ln(2y)$ at the point $(-3, 3, 2.138)$.

z = _____

Note: Your answer should be an expression of x and y; e.g. "5x + 2y - 3"

3. A student was asked to find an equation for the plane that is tangent plane to the surface

$z = x^3 - y^4$

at the point $(x, y) = (5, 3)$. The student's answer was

$z = 44 + 3x^2(x - 5) - 4y^3(y - 3)$.

a. At a glance, how do you know this is wrong? What mistakes did the student make? Select all that apply.

- ☐ The answer is not a linear function.
- ☐ The partial derivatives were not evaluated at the point.
- ☐ The 44 should not be in the answer.
- ☐ The (x - 5) and (y - 3) should be x and y.
- ☐ All of the above

b. Find a correct equation for the tangent plane.

$z =$_

4. ***(a)*** Check the local linearity of $f(x, y) = e^x \cos(y)$ near $x = -1,\ y = 2.5$ by filling in the following table of values of f for $x = -1.1,\ -1,\ -0.9$ and $y = 2.4,\ 2.5,\ 2.6$. Express values of f with 4 digits after the decimal point.

$x =$	-1.1	-1	-0.9
$y = 2.4$	_____	_____	_____
$y = 2.5$	_____	_____	_____
$y = 2.6$	_____	_____	_____

(b) Next, fill in the table for the values $x = -1.01,\ -1,\ -0.99$ and $y = 2.49,\ 2.5,\ 2.51$, again showing 4 digits after the decimal point.

$x =$	-1.01	-1	-0.99
$y = 2.49$	_____	_____	_____
$y = 2.5$	_____	_____	_____
$y = 2.51$	_____	_____	_____

Notice if the two tables look nearly linear, and whether the second looks more linear than the first (in particular, think about how you would decide if they were linear, or if the one were more closely linear than the other).

(c) Give the local linearization of $f(x, y) = e^x \cos(y)$ at $(-1, 2.5)$:

Using the second of your tables:

$f(x, y) \approx$ _____

Using the fact that $f_x(x, y) = e^x \cos(y)$ and $f_y(x, y) = -e^x \sin(y)$:

$f(x, y) \approx$ _____

5. Suppose that z is a linear function of x and y with slope 3 in the x direction and slope -2 in the y direction.

(a) A change of -0.1 in x and 0.2 in y produces what change in z?

change in $z =$ _____

(b) If $z = 4$ when $x = 3$ and $y = 4$, what is the value of z when $x = 2.8$ and $y = 4.2$?

$z =$ _____

6. The table below gives the number of calories burned per minute for someone roller-blading, as a function of the person's weight in pounds and speed in miles per hour [from the August 28,1994, issue of *Parade Magazine*].

calories burned per minute

weight\speed	8	9	10	11
120	4.2	5.8	7.4	8.9
140	5.1	6.7	8.3	9.9
160	6.1	7.7	9.2	10.8
180	7	8.6	10.2	11.7
200	7.9	9.5	11.1	12.6

(a) Suppose that a 180 lb person and a 200 person both go 12 miles, the first at 10 mph and the second at 9 mph.

How many calories does the 180 lb person burn?

_____

How many calories does the 200 lb person burn?

_____

(b) We might also be interested in the number of calories each person burns per pound of their weight.

How many calories per pound does the 180 lb person burn?

_____

How many calories per pound does the 200 lb person burn?

_____

7. The dimensions of a closed rectangular box are measured as 100 centimeters, 80 centimeters, and 60 centimeters, respectively, with the error in each measurement at most .2 centimeters. Use differentials to estimate the maximum error in calculating the surface area of the box.

_____ square centimeters

8. One mole of ammonia gas is contained in a vessel which is capable of changing its volume (a compartment sealed by a piston, for example). The total energy U (in Joules) of the ammonia is a function of the volume V (in cubic meters) of the container, and the temperature T (in degrees Kelvin) of the gas. The differential dU is given by $dU = 840dV + 27.32dT$.

(a) How does the energy change if the volume is held constant and the temperature is decreased slightly?

◯ it decreases slightly

◯ it increases slightly

◯ it does not change

(b) How does the energy change if the temperature is held constant and the volume is increased slightly?

◯ it increases slightly

◯ it decreases slightly

◯ it does not change

(c) Find the approximate change in energy if the gas is compressed by 450 cubic centimeters and heated by 3 degrees Kelvin.

Change in energy = _____. Please include in your answer.

9. An unevenly heated metal plate has temperature $T(x, y)$ in degrees Celsius at a point (x, y). If $T(2, 1) = 104$, $T_x(2, 1) = 15$, and $T_y(2, 1) = -14$, estimate the temperature at the point $(2.02, 0.95)$.

$T(2.02, 0.95) \approx$ _____. Please include in your answer.

10. Find the equation of the tangent plane to the surface $z = 4y^2 - 1x^2$ at the point $(0, -2, 16)$.

z = _____

Note: Your answer should be an expression of x and y; e.g. "3x - 4y + 6"

11. Find the equation of the tangent plane to the surface $z = e^{-1x/17} \ln(1y)$ at the point $(4, 1, 0)$.

z = _____

Note: Your answer should be an expression of x and y; e.g. "5x + 2y - 3"

12. Find the linearization $L(x, y)$ of the function $f(x, y) = \sqrt{184 - 9x^2 - 16y^2}$ at $(2, -3)$.

$L(x, y) =$ _____

Note: Your answer should be an expression in x and y; e.g. "3x - 5y + 9"

13. The dimensions of a closed rectangular box are measured as 80 centimeters, 80 centimeters, and 90 centimeters, respectively, with the error in each measurement at most .2 centimeters. Use differentials to estimate the maximum error in calculating the surface area of the box.

_____ square centimeters

14. Find the differential of the function $z = e^{-y} \cos(x)$.

$dz =$ _____

15. Find the differential of the function $f(x, y) = xe^y$ at $(2, 0)$.

$df =$ _____

16. Find the differential of $f(x, y) = \sqrt{x^3 + y^2}$ at the point $(2, 1)$.

$df =$ _____

Then use the differential to estimate $f(2.04, 0.9)$.

$f(2.04, 0.9) \approx$ _____

17. The gas equation for one mole of oxygen relates its pressure, P (in atmospheres), its temperature, T (in K), and its volume, V (in cubic decimeters,

dm^3):

$$T = 16.574 \cdot \frac{1}{V} - 0.52754 \cdot \frac{1}{V^2} - 0.3879\,P + 12.187\,V\,P.$$

(a) Find the temperature T and differential dT if the volume is 20 dm^3 and the pressure is 0.5 atmosphere.

$T =$ _____

$dT =$ _____

(b) Use your answer to part (a) to estimate how much the volume would have to change if the pressure increased by 0.1 atmosphere and the temperature remained constant.

change in volume = _____

18. Find the equation for a linear function $f(x, y)$ with the values given in the table below:

$x\backslash y$	-1	0	1	2
0	2.5	4	5.5	7
1	1	2.5	4	5.5
2	-0.5	1	2.5	4
3	-2	-0.5	1	2.5

$f(x, y) =$ _____

19. The table below gives a partial table of values for a linear function. Fill in the blanks.

$x\backslash y$	-1.0	0.0	1.0
2.0	8	_____	_____
3.0	_____	9	8

20. Suppose that $f(x, y)$ is a smooth function and that its partial derivatives have the values, $f_x(-2, -8) = 2$ and $f_y(-2, -8) = -5$. Given that $f(-2, -8) = 6$, use this information to estimate the value of $f(-1, -7)$. Note this is analogous to finding the tangent line approximation to a function of one variable. In fancy terms, it is the first Taylor approximation.

Estimate of (integer value) $f(-2, -7)$ _____

Estimate of (integer value) $f(-1, -8)$ _____

Estimate of (integer value) $f(-1, -7)$ _____

21. Estimate $f(1.98, -3.03)$ given that $f(2, -3) = -2, f_x(2, -3) = 3$ and $f_y(2, -3) = 4$.

$f(1.98, -3.03) \approx$ _____

22. ***(a)*** For $g(x, y) = \sqrt{x^2 + 3y + 14}$,

$\nabla g(4, 2) =$ _____

(b) Find the best linear approximation of $g(x, y)$ for (x, y) near $(4, 2)$.

Linear approximation = _____

(c) Use the approximation in part (b) to estimate $g(4.03, 1.97)$.

$g(4.03, 1.97) \approx$ _____

23. Let f be the function defined by $f(x, y) = 2x^2 + 3y^3$.

(a) Find the equation of the tangent plane to f at the point $(1, 2)$.

(b) Use the linearization to approximate the values of f at the points $(1.1, 2.05)$ and $(1.3, 2.2)$.

(c) Compare the approximations form part (b) to the exact values of

$f(1.1, 2.05)$ and $f(1.3, 2.2)$. Which approximation is more accurate. Explain why this should be expected.

24. Let f be the function defined by $f(x, y) = x^{1/3}y^{1/3}$, whose graph is shown in Figure 11.6.10.

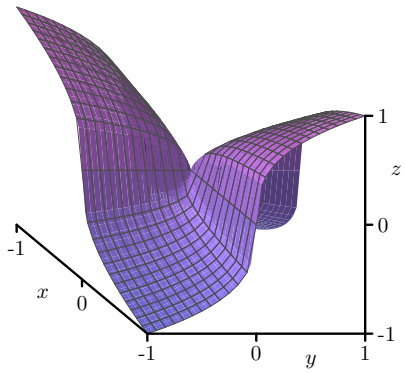


Figure 11.6.10 The surface for $f(x, y) = x^{1/3}y^{1/3}$.

(a) Determine

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

What does this limit tell us about $f_x(0, 0)$?

(b) Note that $f(x, y) = f(y, x)$, and this symmetry implies that $f_x(0, 0) = f_y(0, 0)$. So both partial derivatives of f exist at $(0, 0)$. A picture of the surface defined by f near $(0, 0)$ is shown in Figure 11.6.10. Based on this picture, do you think f is locally linear at $(0, 0)$? Why?

(c) Show that the curve where $x = y$ on the surface defined by f is not differentiable at 0. What does this tell us about the local linearity of f at $(0, 0)$?

(d) Is the function f defined by $f(x, y) = \frac{x^2}{y^2+1}$ locally linear at $(0, 0)$? Why or why not?

25. Let g be a function that is differentiable at $(-2, 5)$ and suppose that its tangent plane at this point is given by $z = -7 + 4(x + 2) - 3(y - 5)$.

(a) Determine the values of $g(-2, 5)$, $g_x(-2, 5)$, and $g_y(-2, 5)$. Write one sentence to explain your thinking.

(b) Estimate the value of $g(-1.8, 4.7)$. Clearly show your work and thinking.

(c) Given changes of $dx = -0.34$ and $dy = 0.21$, estimate the corresponding change in g that is given by its differential, dg.

(d) Suppose that another function h is also differentiable at $(-2, 5)$, but that its tangent plane at $(-2, 5)$ is given by $3x + 2y - 4z = 9$. Determine the values of $h(-2, 5)$, $h_x(-2, 5)$, and $h_y(-2, 5)$, and then estimate the value of $h(-1.8, 4.7)$. Clearly show your work and thinking.

26. In the following questions, we determine and apply the linearization for several different functions.

(a) Find the linearization $L(x, y)$ for the function f defined by $f(x, y) = \cos(x)(2e^{2y} + e^{-2y})$ at the point $(x_0, y_0) = (0, 0)$. Use the linearization to estimate the value of $f(0.1, 0.2)$. Compare your estimate to the actual value of $f(0.1, 0.2)$.

(b) The Heat Index, I, (measured in apparent degrees F) is a function of the actual temperature T outside (in degrees F) and the relative humidity H (measured as a percentage). A portion of the table which gives values for this function, $I = I(T, H)$, is provided in Table 11.6.11.

Table 11.6.11 Heat index.

$T \downarrow \setminus H \rightarrow$	70	75	80	85
90	106	109	112	115
92	112	115	119	123
94	118	122	127	132
96	125	130	135	141

Suppose you are given that $I_T(94, 75) = 3.75$ and $I_H(94, 75) = 0.9$. Use this given information and one other value from the table to estimate the value of $I(93.1, 77)$ using the linearization at $(94, 75)$. Using proper terminology and notation, explain your work and thinking.

(c) Just as we can find a local linearization for a differentiable function of two variables, we can do so for functions of three or more variables. By extending the concept of the local linearization from two to three variables, find the linearization of the function $h(x, y, z) = e^{2x}(y+z^2)$ at the point $(x_0, y_0, z_0) = (0, 1, -2)$. Then, use the linearization to estimate the value of $h(-0.1, 0.9, -1.8)$.

27. In the following questions, we investigate two different applied settings using the differential.

(a) Let f represent the vertical displacement in centimeters from the rest position of a string (like a guitar string) as a function of the distance x in centimeters from the fixed left end of the string and y the time in seconds after the string has been plucked. (An interesting video of this can be seen at https://www.youtube.com/watch?v=TKF6nFzpHBUA.) A simple model for f could be

$$f(x, y) = \cos(x)\sin(2y).$$

Use the differential to approximate how much more this vibrating string is vertically displaced from its position at $(a, b) = \left(\frac{\pi}{4}, \frac{\pi}{3}\right)$ if we decrease a by 0.01 cm and increase the time by 0.1 seconds. Compare to the value of f at the point $\left(\frac{\pi}{4} - 0.01, \frac{\pi}{3} + 0.1\right)$.

(b) Resistors used in electrical circuits have colored bands painted on them to indicate the amount of resistance and the possible error in the resistance. When three resistors, whose resistances are R_1, R_2, and R_3, are connected in parallel, the total resistance R is given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}.$$

Suppose that the resistances are $R_1 = 25\Omega$, $R_2 = 40\Omega$, and $R_3 = 50\Omega$. Find the total resistance R. If you know each of R_1, R_2, and R_3 with a possible error of 0.5%, estimate the maximum error in your calculation of R.

28. In this section we argued that if $f = f(x, y)$ is a function of two variables and if f_x and f_y both exist and are continuous in an open disk containing the point (x_0, y_0), then f is differentiable at (x_0, y_0). This condition ensures that f is differentiable at (x_0, y_0), but it does not define what it means for f to be differentiable at (x_0, y_0). In this exercise we explore the definition of differentiability of a function of two variables in more detail. Throughout, let g be the function defined by $g(x, y) = \sqrt{|xy|}$.

(a) Use appropriate technology to plot the graph of g on the domain $[-1, 1] \times [-1, 1]$. Explain why g is not locally linear at $(0, 0)$.

(b) Show that both $g_x(0, 0)$ and $g_y(0, 0)$ exist. If g is locally linear at $(0, 0)$, what must be the equation of the tangent plane L to g at $(0, 0)$?

(c) Recall that if a function $f = f(x)$ of a single variable is differentiable at $x = x_0$, then

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. We saw in single variable calculus that the existence of $f'(x_0)$ means that the graph of f is locally linear at $x = x_0$. In other words, the graph of f looks like its linearization $L(x) = f(x_0) + f'(x_0)(x - x_0)$ for x close to x_0. That is, the values of $f(x)$ can be closely approximated by $L(x)$ as long as x is close to x_0. We can measure how good the approximation of $L(x)$ is to $f(x)$ with the error function

$$E(x) = L(x) - f(x) = f(x_0) + f'(x_0)(x - x_0) - f(x).$$

As x approaches x_0, $E(x)$ approaches $f(x_0) + f'(x_0)(0) - f(x_0) = 0$, and so $L(x)$ provides increasingly better approximations to $f(x)$ as x gets closer to x_0.

Show that, even though $g(x, y) = \sqrt{|xy|}$ is not locally linear at $(0, 0)$, its error term

$$E(x, y) = L(x, y) - g(x, y)$$

at $(0, 0)$ has a limit of 0 as (x, y) approaches $(0, 0)$. (Use the linearization you found in part (b).) This shows that just because an error term goes to 0 as (x, y) approaches (x_0, y_0), we cannot conclude that a function is locally linear at (x_0, y_0).

(d) As the previous part illustrates, having the error term go to 0 does not ensure that a function of two variables is locally linear. Instead, we need a notation of a relative error. To see how this works, let us return to the single variable case for a moment and consider $f = f(x)$ as a function of one variable. If we let $x = x_0 + h$, where $|h|$ is the distance from x to x_0, then the relative error in approximating $f(x_0 + h)$ with $L(x_0 + h)$ is

$$\frac{E(x_0 + h)}{h}.$$

Show that, for a function $f = f(x)$ of a single variable, the limit of the relative error is 0 as h approaches 0.

(e) Even though the error term for a function of two variables might have a limit of 0 at a point, our example shows that the function may not be locally linear at that point. So we use the concept of relative error to define differentiability of a function of two variables. When we consider differentiability of a function $f = f(x, y)$ at a point (x_0, y_0), then if $x = x_0 + h$ and $y = y_0 + k$, the distance from (x, y) to (x_0, y_0) is $\sqrt{h^2 + k^2}$.

Definition 11.6.12 A function $f = f(x, y)$ is **differentiable at a point** (x_0, y_0) if there is a linear function $L = L(x, y) = f(x_0, y_0) + m(x - x_0) + n(y - y_0)$ such that the relative error

$$\frac{E(x_0 + h, y_0 + k)}{\sqrt{h^2 + k^2}},$$

has at limit of 0 at $(h, k) = (0, 0)$, where $E(x, y) = f(x, y) - L(x, y)$, $h = x - x_0$, and $k = y - y_0$. ◆

A function f is differentiable if it is differentiable at every point in its domain. The function L in the definition is the linearization of f at (x_0, y_0). Verify that $g(x, y) = \sqrt{|xy|}$ is not differentiable at $(0, 0)$ by showing that the relative error at $(0, 0)$ does not have a limit at $(0, 0)$. Conclude that the existence of partial derivatives at a point is not enough to ensure differentiability at that point. (Hint: Consider the limit along different paths.)

29. Suppose that a function $f = f(x, y)$ is differentiable at a point (x_0, y_0). Let $L = L(x, y) = f(x_0, y_0) + m(x - x_0) + n(y - y_0)$ as in the conditions of Definition 11.6.12. Show that $m = f_x(x_0, y_0)$ and $n = f_y(x_0, y_0)$. (Hint: Calculate the limits of the relative errors when $h = 0$ and $k = 0$.)

30. We know that if a function of a single variable is differentiable at a point, then that function is also continuous at that point. In this exercise we determine that the same property holds for functions of two variables. A function f of the two variables x and y is continuous at a point (x_0, y_0) in its domain if

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$$

or (letting $x = x_0 + h$ and $y = y_0 + k$,

$$\lim_{(h,k)\to(0,0)} f(x_0 + h, y + k) = f(x_0, y_0).$$

Show that if f is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0). (Hint: Multiply both sides of the equality that comes from differentiability by $\lim_{(h,k)\to(0,0)} \sqrt{h^2 + k^2}$.)

11.7 The Multivariable Chain Rule

Motivating Questions

- How can we take derivatives involving compositions of multivariable functions?
- For locally linear functions, how can a partial derivative of a composition of multivariable functions separate into dependence on the input variables?

11.7.1 Introduction

In single-variable calculus, we encountered situations in which some quantity was related through a composition of related quantities. For instance, suppose P is the price of a product and P depends on the wages x of people making the product. The wages of people making the product will also change over time because of seasonal availability and other factors. So a change in the value of time t produces a change in the value of x, which produces a change in the value of P. This means that we can write P as a function of x and x as a function of t; thus we can express P's dependence on t by a composition of functions: $(P \circ x)(t) = P(x(t))$.

Applying the chain rule from single-variable calculus to find the rate of change of the price P with respect to time t likely leads you to write

$$\frac{dP}{dt} = P'(x(t))x'(t).$$

However, this form of the chain rule is only useful when we have algebraic expressions for $P(x)$ and $x(t)$. It may be that we know about the rate of change of P with respect to x at a specific value x_0 and the rate of change of x with respect to t at a specific value t_0 for which we have $x(t_0) = x_0$. Using the language of differentials and local linearity we can express these changes as follows. For a small change in x, the change in the price of the product is approximately proportional to the change in x:

$$dP = \left.\frac{dP}{dx}\right|_{x=x_0} dx,$$

where $\left.\frac{dP}{dx}\right|_{x=x_0}$ denotes the instantaneous rate of change in the product price per unit of labor cost at $x = x_0$. Similarly, a small change in the value of t will also cause a proportional change in the value of x:

$$dx = \left.\frac{dx}{dt}\right|_{t=t_0} dt.$$

We can combine these relationships to express how a change in t produces a change in P:

$$dP = \left.\frac{dP}{dx}\right|_{x=x_0} dx = \left(\left.\frac{dP}{dx}\right|_{x=x_0}\right)\left(\left.\frac{dx}{dt}\right|_{t=t_0}\right) dt$$

We should interpret this as telling us that the change in the price of the product is approximately the current instantaneous rate of change of the price per step in x times the current rate of change in the price of labor times the step in time.

As we look at functions of two (or more) variables in this text, we will encounter a situation where $z = f(x, y)$ and both x and y depend on another variable t. A change in t then produces changes in *both* x and y, which then causes z to change. In this section, we will see how to describe the change in z that is caused by a change in t, leading us to multivariable versions of the chain rule involving both regular and partial derivatives.

Preview Activity 11.7.1 Your self-driving car company, *Steer Clear*, is doing well and almost ready to launch its first car. Some of your engineers have reported that the car has problems when the air intake encounters a large quantity of large particulate matter (sand, dust, large pollen, smog, etc.) in the air. To fix this issue, you have created a new type of filter that uses a

sophisticated mesh and gravity to filter out large particulate matter. However, this new type of filter gets clogged if exposed to too much large particulate matter too quickly. To determine the viability of this filter, you must measure the rate at which your car's intake will be exposed to large particulate matter per unit time.

You consult a friend Alex who does atmospheric modeling of large particulate matter in your area. Alex created a function that describes the amount of large particulate matter in the air in terms of location relative to her lab. This function is expressed algebraically as $P(x, y) = 10 - \frac{1}{2}x^2 - \frac{1}{5}y^2$. Here x is the distance east/west from Alex's lab and y is the distance north/south from her lab. Both distances are measured in kilometers and P is measured in parts per million.

Using the self-driving feature of your car, you specify that your car will move along a test course that can be described as

$$\vec{r}(t) = \langle x(t), y(t) \rangle = \langle 2 - t^2, t^3 + 1 \rangle,$$

where the x and y coordinates are the same as measured by Alex's function and t is measured in minutes.

(a) Substitute the path component functions $x(t)$ and $y(t)$ into the expression for $P(x, y)$ but ***do not simplify*** the resulting expression for $P(t)$. This gives a one-variable function that describes the amount of large particulate matter encountered at each location on your test course as a function of *time*.

(b) Use derivative rules from single-variable calculus to find $P'(t) = \frac{dP}{dt}$. ***Do not simplify this expression.*** Note that this is *not* a partial derivative because $P(t)$ only has *one* dependent variable.

(c) Alex has told you that the amount of large particulate matter in the air changes in the y-direction (north/south) because of the proximity to industrial pollution from a factory complex. Also, the amount of large particulate matter in the air changes in the x-direction (east/west) because of tree pollen from a species of tree in a nearby forest. You would like to understand how the rate of large particulate matter encountered on the car's drive is split into these two factors.

Alex suggests that you measure $\frac{dP}{dt}$ at a location $(x, y) = (a, b)$ by taking the derivative of the linearization for P at (a, b). The linearization of P at (a, b) is given by

$$L(x, y) = P(a, b) + P_x(a, b)(x - a) + P_y(a, b)(y - b)$$

Remember that $x = x(t)$ and $y = y(t)$ are being considered as functions of time in this scenario.

Write a couple of sentences to explain why

$$\frac{dP}{dt} \approx \frac{dL}{dt} = 0 + P_x(a, b)\frac{dx}{dt} + P_y(a, b)\frac{dy}{dt} \tag{11.7.1}$$

and how $P_x(a, b)\frac{dx}{dt}$ and $P_y(a, b)\frac{dy}{dt}$ measure the rates of change in pollution coming from tree pollen and industrial pollution, respectively.

(d) Use the formulas provided at the beginning of the Preview Activity for $P(x, y)$, $x(t)$, and $y(t)$ to compute each of the following. Then substitute your expressions into Alex's approximation given in equation (11.7.1). ***Do not simplify this expression.***

$$P_x(a, b) \qquad P_y(a, b) \qquad \frac{dx}{dt} \qquad \frac{dy}{dt}$$

(e) Compare your results for part (b) and part (d) and write a couple of sentences that describe how each part of your answer for part (d) corresponds to different parts of part (b). Remember your work for part (b) is completely in terms of t so you will need to translate some terms from t into its meaning in other quantities.

11.7.2 The Chain Rule

As Preview Activity 11.7.1 suggests, the following version of the chain rule holds in general.

The Chain Rule.

If $z = f(x, y)$ is a differentiable function of x and y and x and y are each differentiable functions of t, then

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{11.7.2}$$

It is important to note the differences among the derivatives in (11.7.2). Since z is a function of the two variables x and y, the derivatives in the chain rule for z with respect to x and y are *partial* derivatives. However, since $x = x(t)$ and $y = y(t)$ are functions of the single variable t, their derivatives are the standard derivatives of functions of one variable. When we compose z with $x(t)$ and $y(t)$, z is represented as a function of the single variable t. Therefore, the derivative of z with respect to t is a derivative from single-variable calculus as well.

To understand why this chain rule works in general, suppose that z depends on x and y so that we can express the change in z in terms of the differential as

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy. \tag{11.7.3}$$

Further suppose that x and y each depend on t, so that

$$dx = \frac{dx}{dt}\,dt \text{ and } dy = \frac{dy}{dt}\,dt. \tag{11.7.4}$$

Combining equations (11.7.3) and (11.7.4), we find that

$$dz = \frac{\partial z}{\partial x}\frac{dx}{dt}\,dt + \frac{\partial z}{\partial y}\frac{dy}{dt}\,dt = \frac{dz}{dt}\,dt,$$

which is the chain rule in this particular context, as expressed in equation (11.7.2).

This approach to understanding the change in z using the differential or linearization also separates how much of the change in z is coming through each of the intermediate variables x and y. In the context of the Preview Activity 11.7.1 where P is the amount of the large particulate matter in the air as function of location (x, y) and location is changing according to time along the path $\vec{r}(t) = \langle x(t), y(t)\rangle$, the chain rule expresses the rate of change in P as a function of time as

$$\frac{dP}{dt} = \frac{\partial P}{\partial x}\frac{dx}{dt} + \frac{\partial P}{\partial y}\frac{dy}{dt}.$$

Because the amount of large particulate matter in terms of the x-coordinate is governed by tree pollen, the term $\frac{\partial P}{\partial x}\frac{dx}{dt}$ describes the rate of change in the large particulate matter per unit time along our drive attributable to tree

pollen. Similarly, the amount of large particulate matter in terms of the y-coordinate is governed by industrial pollution. Hence, the second term $\frac{\partial P}{\partial y}\frac{dy}{dt}$ describes the rate of change in the large particulate matter per unit time along our drive attributable to industrial pollution. Because this function is locally linear, the total rate of change in large particulate matter is the sum of these individual rates of change.

Activity 11.7.2 In the following questions, we apply the chain rule in several different contexts.

(a) Suppose $z(x, y) = x^2 + xy^3$. In addition, suppose that x and y are restricted to points that move around the plane by following a circle of radius 2 centered at the origin that is parameterized by

$$x(t) = 2\cos(t), \quad \text{and} \quad y(t) = 2\sin(t).$$

i. Use the chain rule to find the instantaneous rate of change $\frac{dz}{dt}$.

ii. Substitute $x(t)$ for x and $y(t)$ for y in the rule for z to write z in terms of t and calculate $\frac{dz}{dt}$ using only techniques from single-variable calculus.

iii. Write a couple of sentences comparing your answers above.

(b) Suppose that the temperature on a metal plate is given by the function T with

$$T(x, y) = 100 - (x^2 + 4y^2)$$

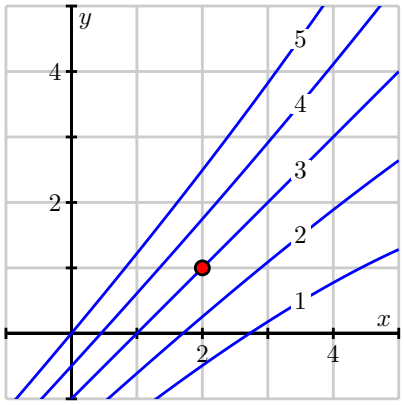
where the temperature is measured in degrees Fahrenheit and x and y are each measured in feet.

i. Find T_x and T_y. What are the units on these partial derivatives?

ii. Suppose an ant is walking along the x-axis at the rate of 2 feet per minute toward the origin. When the ant is at the point $(2, 0)$, what is the instantaneous rate of change in the temperature dT/dt that the ant experiences. Include units on your response.

iii. Suppose instead that the ant walks along the ellipse $\vec{r}(t) = \langle 6\cos(t), 3\sin(t)\rangle$, where t is measured in minutes. Use the multivariable chain rule to find $\frac{dT}{dt}$. What does this tell you about the path along which the ant is walking?

(c) Suppose that you are walking along a surface whose elevation is given by a function f. Some contours of f are as shown in the image below. Furthermore, suppose that if you consider how your location corresponds to points in the xy-plane, you know that when you pass the point $(2, 1)$, your velocity vector is $\vec{v} = \langle -1, 2\rangle$. Estimate the rate of change df/dt when you pass through $(2, 1)$.



The previous activity and discussion have all dealt with a multivariable function where each input of the function depends on a single variable t. The next example generalizes the concept of the chain rule for multivariable functions. We show how the multivariable chain rule can be applied to functions with more than two input variables as well as the situation where the input variables depend on more than one other variable.

Example 11.7.1 Let C be the cost to manufacture flibertygibbits. The cost to produce flibertygibbits depends on the costs of two parts, which we will call p and q, and the cost of labor to construct the flibertygibbits, which we will call L. This means that C is a multivariable function of p, q, and L, which we will denote $C(p, q, L)$. The part costs and the labor costs each depend on the distance s to nearby suppliers and cities and time t. This means that p, q, and L are themselves multivariable functions of s and t. In other words, we need to consider $p(s, t)$, $q(s, t)$, and $L(s, t)$.

To find how the cost to produce flibertygibbits depends on changes in s or t, we can use the chain rule. We can look at the differential of C to determine how these dependencies affect each other. This differential can be expressed as

$$dC = \frac{\partial C}{\partial p}dp + \frac{\partial C}{\partial q}dq + \frac{\partial C}{\partial L}dL,$$

and it describes how a change in C depends on changes in each variable p, q, and L.

We can look at how this change in C depends on either s or t considering the differentials dC, dp, dq, and dL via the appropriate partial derivatives with respect to s or t. In particular, we have

$$\frac{\partial C}{\partial s} = \frac{\partial C}{\partial p}\frac{\partial p}{\partial s} + \frac{\partial C}{\partial q}\frac{\partial q}{\partial s} + \frac{\partial C}{\partial L}\frac{\partial L}{\partial s}$$

and

$$\frac{\partial C}{\partial t} = \frac{\partial C}{\partial p}\frac{\partial p}{\partial t} + \frac{\partial C}{\partial q}\frac{\partial q}{\partial t} + \frac{\partial C}{\partial L}\frac{\partial L}{\partial t}.$$

These chain rule expressions have the same separation of intermediate variables as in our earlier discussion. For instance, the term $\frac{\partial C}{\partial q}\frac{\partial q}{\partial s}$ describes the rate of change in the cost of making flibertygibbits per unit of distance from other cities that depends on the cost of part q. The term $\frac{\partial C}{\partial L}\frac{\partial L}{\partial t}$ describes the rate of change in the cost of making flibertygibbits per unit time that depends on the cost of labor L. In both cases, the partial derivatives of C with respect to s and t are a linear combination of the changes through each of the intermediate variables p, q, and L. ▲

The chain rule and local linearity are related, conceptually-significant ideas. Every version of the chain rule we have seen, including the chain rule from single-variable calculus, ultimately states that the change in an output variable is a linear combination of the instantaneous rates of change of each input variable and the step size of that variable. The chain rule is an algebraic statement of this linear combination. Recall from the previous section that when a function is locally linear, the instantaneous rate of change of the function is the same as the corresponding measurement of change along the linearization/ tangent plane.

Summary

- The Chain Rule is a tool for differentiating composite functions. In its simplest form, it says that if $f(x, y)$ is a function of two variables and $x(t)$ and $y(t)$ depend on t, then

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

- The Chain Rule generalizes to functions of more than two variables and functions with more than one independent variable. If $g(x, y, z)$ is a function of three variables and each of x, y, and z depend on s and t, then

$$\frac{\partial g}{\partial s} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial g}{\partial z}\frac{\partial a}{\partial s}$$

and

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial t}$$

11.7.3 Exercises

1. Use the chain rule to find $\frac{dz}{dt}$, where

$$z = x^2y + xy^2, \quad x = 3 - t^7, \quad y = 5 + t^3$$

First the pieces:
$\frac{\partial z}{\partial x} =$ _____
$\frac{\partial z}{\partial y} =$ _____
$\frac{dx}{dt} =$ _____
$\frac{dy}{dt} =$ _____
End result (in terms of just t):
$\frac{dz}{dt} =$ _____

2. Use the chain rule to find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$, where

$$z = e^{xy}\tan y, x = 5s + 4t, y = \frac{4s}{4t}$$

First the pieces:
$\frac{\partial z}{\partial x} =$_____ $\frac{\partial z}{\partial y} =$_____
$\frac{\partial x}{\partial s} =$_____ $\frac{\partial x}{\partial t} =$_____
$\frac{\partial y}{\partial s} =$_____ $\frac{\partial y}{\partial t} =$_____
And putting it all together:
$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$ and $\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$

3. Suppose $w = \dfrac{x}{y} + \dfrac{y}{z}$, where

$x = e^{5t},\ y = 2 + \sin(2t)$, and $z = 2 + \cos(6t)$.

A) Use the chain rule to find $\frac{dw}{dt}$ as a function of x, y, z, and t. Do not rewrite x, y, and z in terms of t, and do not rewrite e^{5t} as x.

$\frac{dw}{dt} =$ _____

Note: You may want to use exp() for the exponential function. Your answer should be an expression in x, y, z, and t; e.g. "3x - 4y"

B) Use part A to evaluate $\frac{dw}{dt}$ when $t = 0$.

4. If $z = (x + y)\, e^y$ and $x = u^2 + v^2$ and $y = u^2 - v^2$, find the following partial derivatives using the chain rule. Enter your answers as functions of u and v.

$\dfrac{\partial z}{\partial u} =$ _____

$\dfrac{\partial z}{\partial v} =$ _____

5. If

$$z = \sin\left(x^2 + y^2\right), \qquad x = u\cos(v), \qquad y = u\sin(v),$$

find $\partial z/\partial u$ and $\partial z/\partial v$. The variables are restricted to domains on which the functions are defined.

$\partial z/\partial u =$ _____

$\partial z/\partial v =$ _____

6. Let $z = g(u, v, w)$ and $u(r, s), v(r, s), w(r, s)$. How many terms are there in the expression for $\partial z/\partial r$?

_____ terms

7. Let $W(s, t) = F(u(s, t), v(s, t))$ where

$$u(1, 0) = -2, u_s(1, 0) = 9, u_t(1, 0) = -1$$

$$v(1, 0) = -9, v_s(1, 0) = 8, v_t(1, 0) = -4$$

$$F_u(-2, -9) = -1, F_v(-2, -9) = -8$$

$W_s(1, 0) =$ _____ $W_t(1, 0) =$ _____

8. The radius of a right circular cone is increasing at a rate of 4 inches per second and its height is decreasing at a rate of 4 inches per second. At what rate is the volume of the cone changing when the radius is 50 inches and the height is 40 inches?

_____ cubic inches per second

9. In a simple electric circuit, Ohm's law states that $V = IR$, where V is the voltage in volts, I is the current in amperes, and R is the resistance in ohms. Assume that, as the battery wears out, the voltage decreases at 0.03 volts per second and, as the resistor heats up, the resistance is increasing at 0.02 ohms per second. When the resistance is 200 ohms and the current is 0.02 amperes, at what rate is the current changing?

_____ amperes per second

10. Suppose $z = x^2 \sin y,\ x = 3s^2 + 3t^2,\ y = 2st$.

A. Use the chain rule to find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ as functions of x, y, s and t.

$\frac{\partial z}{\partial s} =$ _____

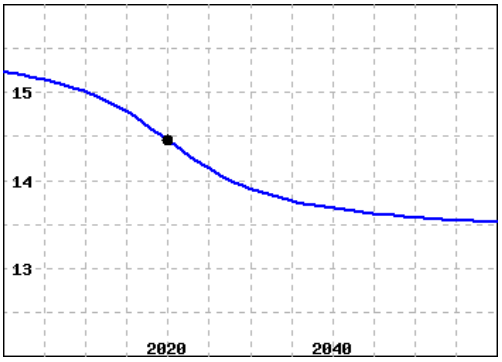
$\frac{\partial z}{\partial t} =$ _____

B. Find the numerical values of $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ when $(s, t) = (-2, 3)$.

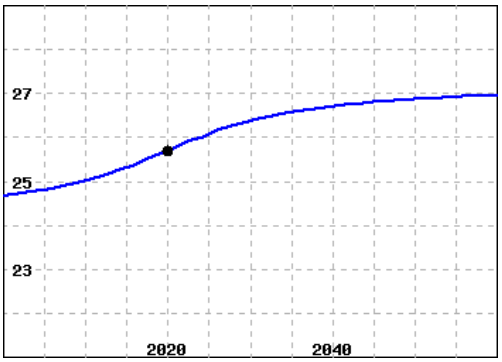
$\frac{\partial z}{\partial s}(-2, 3) =$ _____

$\frac{\partial z}{\partial t}(-2,3) =$ _____

11. Corn production, C, is a function of rainfall, R (in inches), and temperature, T (in degrees C). The first figure below shows how rainfall is predicted to vary with time because of global warming; the second, how temperature is predicted to vary with time because of global warming.



Graph of rainfall, R vs. time, t



Graph of temperature, T vs. time, t
Suppose that we know that (currently)

$$\Delta C \approx 3.8\Delta R - 5.1\Delta T$$

Use this to estimate the change in corn production between the year 2020 and the year 2021.
Change = _____
Hence, estimate dC/dt when $t = 2020$.
$dC/dt =$ _____

12. Find the indicated derivative. In each case, state the version of the Chain Rule that you are using.

(a) $\frac{df}{dt}$, if $f(x,y) = 2x^2y$, $x = \cos(t)$, and $y = \ln(t)$.

(b) $\frac{\partial f}{\partial w}$, if $f(x,y) = 2x^2y$, $x = w + z^2$, and $y = \frac{2z+1}{w}$

(c) $\frac{\partial f}{\partial v}$, if $f(x,y,z) = 2x^2y + z^3$, $x = u - v + 2w$, $y = w2^v - u^3$, and $z = u^2 - v$

13. Let $z = u^2 - v^2$ and suppose that

$$u = e^x \cos(y)$$
$$v = e^x \sin(y)$$

(a) Find the values of u and v that correspond to $x = 0$ and $y = 2\pi/3$.

(b) Use the Chain Rule to find the general partial derivatives

$$\frac{\partial z}{\partial x} \text{ and } \frac{\partial z}{\partial y}$$

and then determine both $\left.\frac{\partial z}{\partial x}\right|_{(0,\frac{2\pi}{3})}$ and $\left.\frac{\partial z}{\partial y}\right|_{(0,\frac{2\pi}{3})}$

In this section, we have applied the Chain Rule when x and y each depend on more than one quantity, or when z is a function of more than two variables. It can be challenging to keep track of all the dependencies among the variables, and thus a tree diagram can be a useful tool to organize our work. For example, suppose that z depends on x and y, and x and y both depend on t. We may represent these relationships using the tree diagram shown at left Figure 11.7.2. We place the dependent variable at the top of the tree and connect it to the variables on which it depends one level below. We then connect each of those variables to the variable on which each depends.

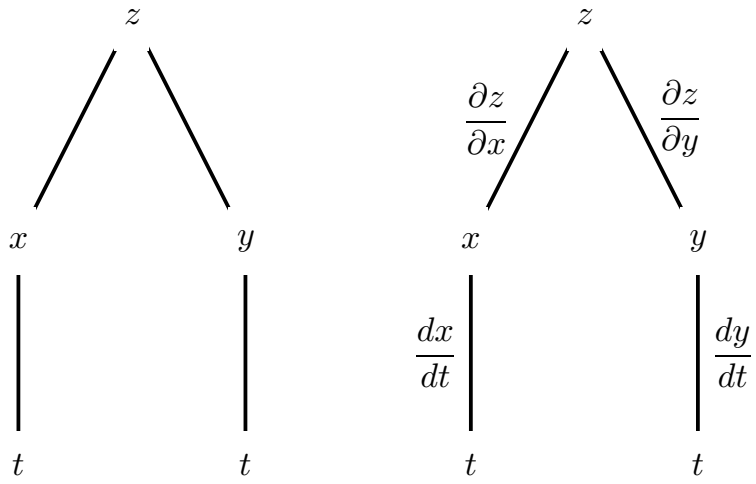


Figure 11.7.2 A tree diagram illustrating dependencies

To represent the Chain Rule, we label every edge of the diagram with the appropriate derivative or partial derivative, as seen at right in Figure 11.7.2. To calculate an overall derivative according to the Chain Rule, we construct the product of the derivatives along all paths connecting the variables and then add all of these products. For example, the diagram at right in Figure 11.7.2 illustrates the Chain Rule

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

14. Suppose that $T = x^2 + y^2 - 2z$ where

$$\begin{aligned} x &= \rho\sin(\phi)\cos(\theta) \\ y &= \rho\sin(\phi)\sin(\theta) \\ z &= \rho\cos(\phi) \end{aligned}$$

Note that we are essentially converting our expression for T to spherical coordinates.

(a) Construct a tree diagram representing the dependencies among the variables.

(b) Apply the chain rule to find the partial derivatives

$$\frac{\partial T}{\partial \rho}, \frac{\partial T}{\partial \phi}, \text{ and } \frac{\partial T}{\partial \theta}.$$

15. Suppose that the temperature on a metal plate is given by the function T with

$$T(x, y) = 100 - (x^2 + 4y^2),$$

where the temperature is measured in degrees Fahrenheit and x and y are each measured in feet. Now suppose that an ant is walking on the metal plate in such a way that it walks in a straight line from the point $(1, 4)$ to the point $(5, 6)$.

(a) Find parametric equations $(x(t), y(t))$ for the ant's coordinates as it walks the line from $(1, 4)$ to $(5, 6)$.

(b) What can you say about $\frac{dx}{dt}$ and $\frac{dy}{dt}$ for every value of t?

(c) Determine the instantaneous rate of change in temperature with respect to t that the ant is experiencing at the moment it is halfway from $(1, 4)$ to $(5, 6)$, using your parametric equations for x and y. Include units on your answer.

16. There are several proposed formulas to approximate the surface area of the human body. One model[1] uses the formula

$$A(h, w) = 0.0072h^{0.725}w^{0.425},$$

where A is the surface area in square meters, h is the height in centimeters, and w is the weight in kilograms.

Since a person's height h and weight w change over time h and w are functions of time t. Let us think about what is happening to a child whose height is 60 centimeters and weight is 9 kilograms. Suppose, furthermore, that h is increasing at an instantaneous rate of 20 centimeters per year and w is increasing at an instantaneous rate of 5 kg per year.

Determine the instantaneous rate at which the child's surface area is changing at this point in time.

17. Let $z = f(x, y) = 50 - (x + 1)^2 - (y + 3)^2$ and $z = h(x, y) = 24 - 2x - 6y$.

Suppose a person is walking on the surface $z = f(x, y)$ in such a way that she walks the curve which is the intersection of f and h.

(a) Show that $x(t) = 4\cos(t)$ and $y(t) = 4\sin(t)$ is a parameterization of the "shadow" in the xy-plane of the curve that is the intersection of the graphs of f and h.

(b) Use the parameterization from part (a) to find the instantaneous rate at which her height is changing with respect to time at the instant $t = 2\pi/3$.

18. The voltage V (in volts) across a circuit is given by Ohm's Law: $V = IR$, where I is the current (in amps) in the circuit and R is the resistance (in ohms). Suppose we connect two resistors with resistances R_1 and R_2 in parallel as shown in Figure 11.7.3. The total resistance R in the circuit is

[1] DuBois D, DuBois DF. A formula to estimate the approximate surface area if height and weight be known. *Arch Int Med* 1916;17:863-71.

then given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$

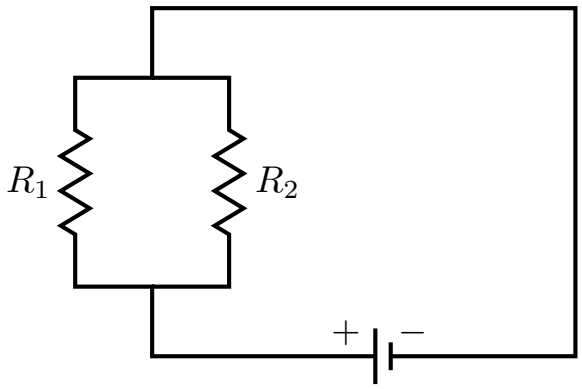


Figure 11.7.3 Resistors in parallel.

(a) Assume that the current, I, and the resistances, R_1 and R_2, are changing over time, t. Use the Chain Rule to write a formula for $\frac{dV}{dt}$.

(b) Suppose that, at some particular point in time, we measure the current to be 3 amps and that the current is increasing at $\frac{1}{10}$ amps per second, while resistance R_1 is 2 ohms and decreasing at the rate of 0.2 ohms per second and R_2 is 1 ohm and increasing at the rate of 0.5 ohms per second. At what rate is the voltage changing at this point in time?

11.8 Directional Derivatives and the Gradient

Motivating Questions

- The partial derivatives of a function f tell us the rate of change of f in the direction of the coordinate axes. How can we measure the rate of change of f in other directions?
- What is the gradient of a multivariable function and what does it tell us about the rate of change in different directions?

11.8.1 Introduction

The partial derivatives of a multivariable function tell us the instantaneous rate at which the function's output changes as we hold all but one input variable constant (allowing the remaining input variable to change). It is natural to wonder how we can measure the rate at which a function changes in a direction other than parallel to a coordinate axis. In this section, we investigate this question and will connect the rates of change in other directions to the rates of change given by the partial derivatives. The Preview Activity investigates these concepts in terms of a contour map representing elevation in terms of location.

We can use cardinal directions to specify the direction of displacement vectors. These directions can be described by a compass rose. The compass rose

given in Figure 11.8.1 is an example of a sixteen point compass rose. Directions of the form ESE are read as "east-southeast" and point in the direction halfway between east and southeast.

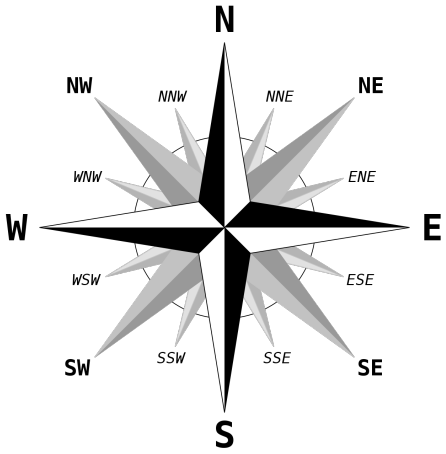
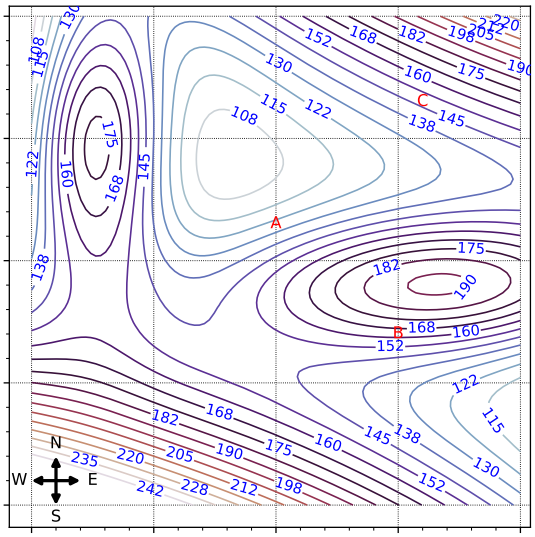


Figure 11.8.1 An example of a sixteen-point compass rose (from By Brosen~commonswiki - Own work, CC BY 2.5)

Preview Activity 11.8.1 Below is a contour plot showing the elevations for a region of a nearby park. We will be referring to $h(x, y)$ as the function of two variables that gives the elevation as a function of x-coordinate (location in the east-west or horizontal direction) and y-coordinates (location in the north-south or vertical direction).



(a) Using the contour plot and treating the elevation as a multivariable function $h(x, y)$, state whether each of the following is positive, negative, or zero. Write a sentence to justify your reasoning.

i. $h_x(A)$
ii. $h_y(A)$
iii. $h_x(B)$
iv. $h_y(B)$
v. $h_x(C)$
vi. $h_y(C)$

(b) Suppose you are standing at point B. Do you expect the elevation to increase, decrease, or remain constant if you take a step in the northeast direction? Write a sentence to explain your reasoning.

(c) Suppose you are standing at point B. Do you expect the elevation to increase, decrease, or remain constant if you take a step in the southeast direction? Write a sentence to explain your reasoning.

(d) Suppose you are standing at point A. Do you expect the elevation to increase, decrease, or remain constant if you take a step in the southwest direction? Write a sentence to explain your reasoning.

(e) Suppose you are standing at point C. In what direction would you take a step to move in the steepest downhill direction? Write a sentence to explain your reasoning.

(f) Suppose you are standing at point A. In what direction would you take a step to move in the steepest uphill direction? Write a sentence to explain your reasoning.

(g) Suppose you are standing at point B. Rank the following directions in order of steepness (from most steep and uphill to level to most steep downhill): northeast, north, east, west, south

The Preview Activity demonstrates how visual information such as a contour plot or a surface plot can be helpful in determining information about how quickly the output of a function changes as its inputs change in different directions. In this section, we will focus on how to calculate this kind of directional derivative using algebraic or numerical representations of functions and relate this calculation to other ways we have described change for a function.

11.8.2 Directional Derivatives

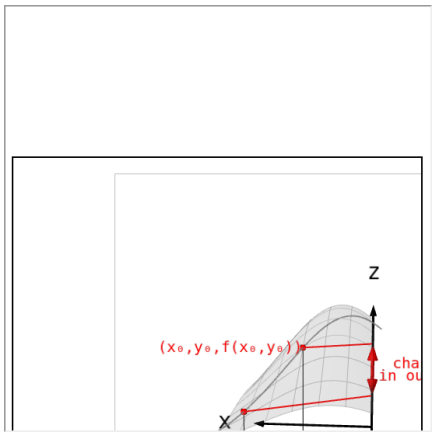
We begin by determining how to measure the rate of change for a function of two variables in a direction that is not parallel to one of the coordinate directions. We use the classic calculus approach to set up this measurement: approximate the measurement, quantify how the approximation changes on a smaller scale, and finally use a limit to find the value. To measure the rate of change in the output of a function, we use a difference quotient:

$$\text{rate of change} = \frac{\text{change in output}}{\text{change in input}}.$$

Let f be a function of the variables x and y. We wish to measure the rate of change of f at the base point (x_0, y_0) in the direction of the vector $\langle a, b\rangle$. To do this, we write the difference quotient using the points $(x_0 + a, y_0 + b)$ and (x_0, y_0):

$$\frac{f(x_0 + a, y_0 + b) - f(x_0, y_0)}{\sqrt{a^2 + b^2}}.$$

The denominator comes from using the distance formula to measure the change in the input as distance in the xy-plane. Figure 11.8.2 shows the plot of a surface given by $z = f(x, y)$ with the points used in the difference quotient labeled.



Standalone
In Context
Embed

Figure 11.8.2 A plot of $f(x, y)$ with the measurements of change in input and output labeled for a change of inputs given by $\langle a, b\rangle$

The difference quotient completes is the first step in the classic calculus approach and provides a good start to approximating the rate of change for the output of f in the direction $\langle a, b\rangle$. However, it is not clear how this approximation changes when we look at smaller scales. Specifically, we need to look at what happens to this approximation as we shrink the step size $\sqrt{a^2 + b^2}$ to 0 while maintaining the same direction of change. In order to maintain our direction *and* look at smaller step sizes, we must separate the length of the vector $\langle a, b\rangle$ from its direction. To do this, we use the unit vector in the direction of $\langle a, b\rangle$.

Let $\vec{u} = \langle u_1, u_2\rangle$ be the unit vector in the same direction as $\langle a, b\rangle$. Vectors in the same direction as $\langle a, b\rangle$ can be written in the form $t\langle u_1, u_2\rangle$, where $t > 0$ is the step size and $\vec{u} = \langle u_1, u_2\rangle$ is the unit vector. Taking a step of length t in the direction $\langle u_1, u_2\rangle$, the difference quotient can be written as

$$\frac{f(x_0 + t(u_1), y_0 + t(u_2)) - f(x_0, y_0)}{t}.$$

We will look at how this difference quotient changes at smaller scales, i.e., when the step size t gets smaller. This accomplishes the second step of the classic calculus approach because it quantifies how the approximation works on smaller scales. We can now take the limit as $t \to 0$ to define the directional derivative.

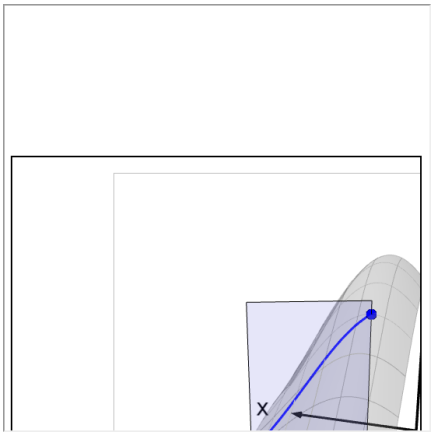
Definition 11.8.3 Let $f(x, y)$ be a function of two variables. The **derivative of f at the point (x, y) in the direction of the unit vector $\vec{u} = \langle u_1, u_2\rangle$** is denoted $D_{\vec{u}} f(x, y)$ and is given by

$$D_{\vec{u}} f(x, y) = \lim_{t \to 0} \frac{f(x + u_1 t, y + u_2 t) - f(x, y)}{t} \tag{11.8.1}$$

for those values of x and y for which the limit exists. ◆

The notation $D_{\vec{u}} f(x, y)$ can also be read as "the directional derivative of f in the direction $\vec{u}$ at the point (x, y)." While this may seem like a mouthful, each piece of this phrase is necessary to specify a directional derivative: we must specify the function, the location at which to measure the rate of change, *and* the direction in which the inputs change. When we evaluate the directional derivative $D_{\vec{u}} f(x, y)$ at a point (x_0, y_0), the result $D_{\vec{u}} f(x_0, y_0)$ tells us the instantaneous rate at which f changes at (x_0, y_0) per unit increase in the direction of the vector $\vec{u}$.

We can connect equation (11.8.1) to our work on traces earlier in this chapter. When initially exploring surface plots, we simplified our approach by looking at slices that were parallel to one of the coordinate axes, obtained by holding the other input variable constant. This allowed us to use ideas from single-variable calculus to understand more about the surface and to motivate the definitions of partial derivatives. Similarly, we can cut the surface in the $\langle u_1, u_2 \rangle$ direction through the base point and then think about the directional derivative as a single-variable calculus problem. In particular, we can parameterize the trace along the surface by noticing our inputs correspond to the line in the xy-plane given by $\langle x_0 + u_1 t, y_0 + u_2 t \rangle$, which is shown in red on Figure 11.8.4.



Standalone
In Context
Embed

Figure 11.8.4 A plot of $f(x, y)$ with a trace shown from the point (x_0, y_0) in the direction $\langle u_1, u_2 \rangle$

When $t = 0$, the location is (x_0, y_0). The trace along the surface is given by $\vec{r}(t) = \langle x_0 + u_1 t, y_0 + u_2 t, f(x_0 + u_1 t, y_0 + u_2 t) \rangle$. This trace is shown in blue in Figure 11.8.4. We can also plot the trace of this surface in the direction of $\langle u_1, u_2 \rangle$ as a function of t using a two-dimensional plot. This plot is shown in Figure 11.8.5. We can visualize the directional derivative as the slope of a tangent line at the base point. This tangent line is drawn in black on Figure 11.8.5. The scalar $D_{\vec{u}} f(x_0, y_0)$ tells us the slope of the line tangent to the surface in the direction of $\vec{u}$ at the point $(x_0, y_0, f(x_0, y_0))$.

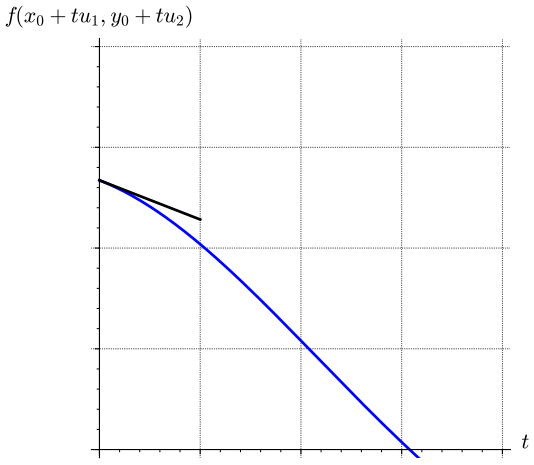


Figure 11.8.5 A plot of the trace of f along the direction $\langle u_1, u_2 \rangle$ with tangent line drawn at point corresponding to (x_0, y_0)

11.8.3 Efficiently Computing the Directional Derivative

We want to find a way to evaluate directional derivatives without resorting to evaluating the limit definition in Equation (11.8.1) every time. We will look at this algebraic approach from two perspectives. First, we will use the chain rule from Section 11.7 to define a simple algebraic rule which will allow us to efficiently calculate directional derivatives. Second, we will use the linearization of a locally linear/differentiable function to compute directional derivatives.

We are interested in the instantaneous rate of change of f at a point (x_0, y_0) in the direction of a unit vector $\vec{u} = \langle u_1, u_2 \rangle$. In particular, the input variables x and y are changing according to

$$x = x_0 + u_1 t \quad \text{and} \quad y = y_0 + u_2 t.$$

A trace along the surface in the direction of $\langle u_1, u_2 \rangle$ is given by $\vec{r}(t) = \langle x_0 + u_1 t, y_0 + u_2 t, f(x_0 + u_1 t, y_0 + u_2 t) \rangle$. Observe that $\frac{dx}{dt} = u_1$ and $\frac{dy}{dt} = u_2$ for all values of t. Since $\vec{u}$ is a unit vector in the xy-plane, a unit change in the parameter t corresponds to moving one unit in the $\vec{u}$ direction. This allows us to use the multivariable chain rule to calculate the directional derivative as a measure of the instantaneous rate of change of f in the direction $\vec{u}$.

The output of f along the given direction can be written as a composition of functions, $f(t) = f(\vec{r}(t)) = f(x(t), y(t))$, which means we can apply the multivariable chain rule:

$$\begin{aligned} D_{\vec{u}} f(x_0, y_0) &= \frac{d}{dt} \left[f(\vec{r}(t)) \right] \\ &= f_x(x_0, y_0) \frac{dx}{dt} + f_y(x_0, y_0) \frac{dy}{dt} \\ &= f_x(x_0, y_0) u_1 + f_y(x_0, y_0) u_2 \end{aligned}$$

This allows us to compute the directional derivative at an arbitrary point according to the following formula.

Theorem 11.8.6 Calculating Directional Derivatives. *Given a differentiable function $f = f(x, y)$ and a unit vector $\vec{u} = \langle u_1, u_2 \rangle$, the directional derivative $D_{\vec{u}} f(x, y)$ can be computed by*

$$D_{\vec{u}} f(x, y) = f_x(x, y) u_1 + f_y(x, y) u_2. \tag{11.8.2}$$

To use equation (11.8.2), we must have a *unit* vector $\vec{u} = \langle u_1, u_2 \rangle$ in the direction of motion. In the event that we have a direction prescribed by a non-unit vector, we must first rescale the vector to have length 1.

Example 11.8.7 We will look at an example that is algebraically and geometrically simple but allows us to use Theorem 11.8.6. We will consider the rate of change for the function $f(x, y) = 2.5 - \frac{(x-1)^2}{2} - \frac{(y+1)^2}{9}$ at the input point $(2, 2)$ in several directions. Because $f(2, 2) = 2.5 - \frac{(2-1)^2}{2} - \frac{(2+1)^2}{9} = 2.5 - 0.5 - 1 = 1$, this point lies on the level curve with value 1 as shown in Figure 11.8.8.

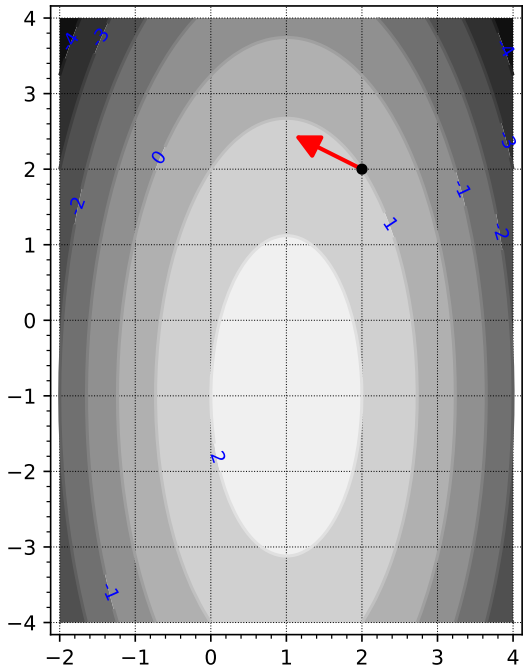


Figure 11.8.8 A contour plot of $f(x,y) = 2.5 - \frac{(x-1)^2}{2} - \frac{(y+1)^2}{9}$ with the point $(2,2)$ highlighted

We will compute the directional derivative of f at the point $(2,2)$ in the direction of $\langle -2,-1\rangle$. Since $\langle -2,1\rangle$ is not a unit vector, we use $\vec{u} = \frac{1}{\sqrt{5}}\langle -2,1\rangle$ as the unit vector in this direction. To use equation (11.8.2), we will need the partial derivatives of f:

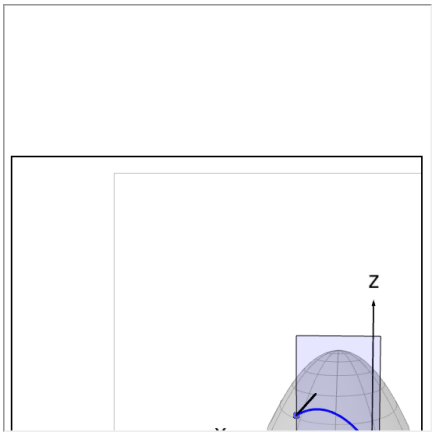
$$f_x = -\frac{2(x-1)}{2} \quad \text{and} \quad f_y = -\frac{2(y+1)}{9}$$

At $(2,2)$, we have $f_x(2,2) = -1$ and $f_y(2,2) = -\frac{2}{3}$. By equation (11.8.2), the directional derivative of f at $(2,2)$ in the direction $\vec{u}$ is

$$Df_{\vec{u}}(2,2) = f_x(2,2)u_1 + f_y(2,2)u_2 = (-1)\left(-\frac{2}{\sqrt{5}}\right) + \left(-\frac{2}{3}\right)\left(\frac{1}{\sqrt{5}}\right) = \frac{4}{3\sqrt{5}}.$$

We can interpret this as saying that for a small step in the direction of $\vec{u}$ at $(2,2)$, we would expect the output of f to increase by $\frac{4}{3\sqrt{5}}$ times the step size.

This may be a bit difficult to see based on the contour plot in Figure 11.8.8. However, the surface plot combined with a plot of the path of inputs in the direction $\vec{u}$ gives a better visualization. Figure 11.8.9 shows a surface plot of $z = f(x,y)$ with the trace in the $\vec{u}$-direction shown in blue and the direction vector $\vec{u}$ shown in red on the xy-plane. The black line segment shows how $Df_{\vec{u}}(2,2)$ is positive because the z-coordinate of this tangent line increases as we move away from $(2,2)$ in the $\vec{u}$-direction.



Standalone
In Context
Embed

Figure 11.8.9 A plot of $z = f(x, y)$ with trace in the direction $\vec{u}$ through $(2, 2, 1)$ shown including the tangent line

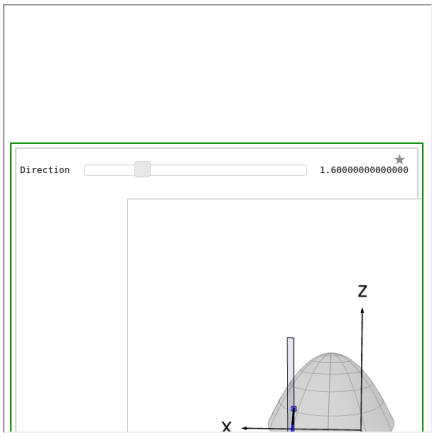
Remember that the directional derivative is a local measurement and only applies in a small neighborhood of a point and only in the direction $\vec{u}$. In other words, $Df_{\vec{u}}(2, 2)$ does not describe the rate of change at other points along the blue curve, only the rate of change along the blue curve for a small step away from our base point in the red direction. With plots like Figure 11.8.8, you may be tempted to use information far away from the point of interest to try to figure out the rate of change. However, like most calculus measurements, we must only use plots to provide information on what is happening near a specific point.

If we compute $Df_{\hat{\jmath}}(2, 2)$, the rate of change for the output of f in the direction $\hat{\jmath} = \langle 0, 1 \rangle$ at $(2, 2)$, using equation (11.8.2), then we have

$$Df_{\hat{\jmath}}(2, 2) = f_x(2, 2)u_1 + f_y(2, 2)u_2 = (-1)(0) + (-\frac{2}{3})(1) = -\frac{2}{3}.$$

This corresponds to the partial derivative of f with respect to y.

Figure 11.8.10 allows you to adjust the direction of the trace shown. By default, it shows the direction vector $\hat{\jmath}$, the cooresponding trace, and tangent line to the surface in the $\hat{\jmath}$ direction. Use the slider at the top of Figure 11.8.10 to change the direction in which you want to evaluate the directional derivative. As you do so, examine how the steepness of the black tangent segment varies depending on the direction.



Standalone
In Context
Embed

Figure 11.8.10 An interactive plot of $z = f(x, y)$ with trace in the direction $\vec{u}$ through (2,2,1) shown including the tangent line

▲

In the following activity, we use algebraic techniques to calculate and interpret the of directional derivative of a function.

Activity 11.8.2 In this activity, we will use equation (11.8.2) to calculate and interpret directional derivatives for the function $f(x, y) = 3xy - x^2y^3$.

(a) Calculate $f_x(x, y)$ and $f_y(x, y)$.

(b) Use equation (11.8.2) to determine $D_{\hat{\imath}}f(x, y)$ and $D_{\hat{\jmath}}f(x, y)$. Write a couple of sentences to describe what familiar functions $D_{\hat{\imath}}f$ and $D_{\hat{\jmath}}f$ are.

Hint. Remember that $\hat{\imath}$ is the unit vector in the positive x-direction and $\hat{\jmath}$ is the unit vector in the positive y-direction.

(c) Use equation (11.8.2) to find the derivative of f in the direction of the vector $\vec{v} = \langle 2, 3\rangle$ at the point $(1, -1)$.

Hint. Remember that a *unit* direction vector is needed.

(d) Find the derivative of f in the direction of the vector $\vec{v} = \langle 4, 6\rangle$ at the point $(1, -1)$.

(e) Use equation (11.8.2) to find the derivative of f in the direction of the vector $\vec{v} = \langle -2, -3\rangle$ at the point $(1, -1)$. Write a couple of sentences to explain why this result is different from your answer to the previous two tasks, even though the direction vectors are parallel.

Equation (11.8.2) arose through consideration of the definition for the directional derivative as a composition of functions and application of the chain rule. When considering multivariable functions that are locally linear, recall that we can evaluate the limit of such a function near a point by analyzing the linearization or tangent plane. On a small scale, there is virtually no difference between the locally linear function and the linearization as demonstrated by Figure 11.6.1.

The equation of the plane tangent to the graph of f at the point $(x_0, y_0, f(x_0, y_0))$ is

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

If f is a locally linear function, then the directional derivative $D_{\vec{u}}f(x_0, y_0)$ is the same as the rate of change along the tangent plane in the direction of $\vec{u}$. This is convenient because on the tangent plane, the rate of change is the same regardless of step size, a fact that is *not* true on the original surface $z = f(x, y)$. We now consider the difference quotient along the tangent plane. We use the notation of the linearization below because the linearization, $L(x, y)$ is the functional form of the tangent plane at (x_0, y_0).

$$\begin{aligned}
&\frac{z(x_0 + u_1, y_0 + u_2) - z(x_0, y_0)}{\sqrt{u_1^2 + u_2^2}} \\
&= \frac{(f(x_0, y_0) + f_x(x_0, y_0)(x_0 + u_1 - x_0) + f_y(x_0, y_0)(y_0 + u_2 - y_0)) - f(x_0, y_0)}{1} \\
&= f_x(x_0, y_0)(u_1) + f_y(x_0, y_0)(u_2),
\end{aligned}$$

which is exactly the result in Equation (11.8.2).

11.8.4 The Gradient

We have used the chain rule and linearization to see see that the instantaneous rate of change a function $f = f(x, y)$ in the direction of a unit vector $\vec{u} =$

$\langle u_1, u_2 \rangle$ is given by

$$D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2. \tag{11.8.3}$$

You may recognize the form of (11.8.3) as being similar to a dot product. We can also view this equation as stating that the rate of change along the linearization in a given direction can be expressed as a linear combination of the rates of change in the coordinate directions by using partial derivatives with the weight of each partial derivative specified by the components of the unit vector $\vec{u}$.

We can think about Equation (11.8.3) in a way that will have geometric meaning related to the dot product. The directional derivative $D_{\vec{u}} f(x_0, y_0)$ is the dot product of $\langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$ and $\vec{u} = \langle u_1, u_2 \rangle$. Notice that the vector $\langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$ comes from simple calculations that tell us how the function f is changing near the input (x_0, y_0).

Definition 11.8.11 The **gradient** of f is vector formed by partial derivatives of f. The gradient of f is denoted

$$\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle .$$

◆

We read ∇f as "the gradient of f," "grad f" or "del f".[1] Notice that ∇f varies from point to point, and also provides an alternate formulation of the directional derivative.

Key Idea 11.8.12 The Directional Derivative and the Gradient. *Given a differentiable function $f = f(x, y)$ and a unit vector $\vec{u} = \langle u_1, u_2 \rangle$, we may compute $D_{\vec{u}} f(x, y)$ by*

$$D_{\vec{u}} f(x, y) = \nabla f(x, y) \cdot \vec{u}. \tag{11.8.4}$$

In the following activity, we investigate some of what the gradient tells us about the behavior of a function f.

Activity 11.8.3 Let $f(x, y) = \frac{1}{3}(x^2 - y^2)$. Some contours for this function are shown in Figure 11.8.13.

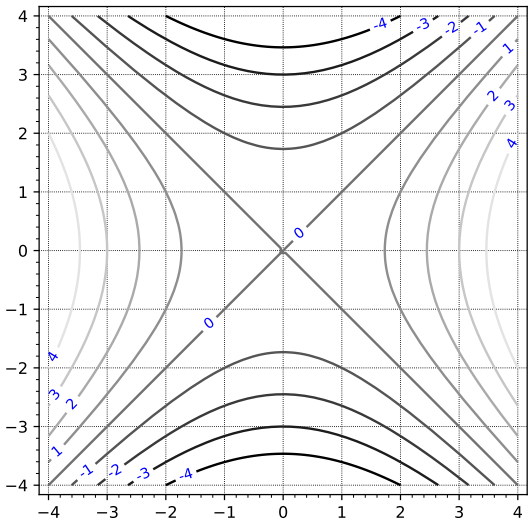


Figure 11.8.13 A contour plot of $f(x, y) = \frac{1}{3}(x^2 - y^2)$

[1] The symbol ∇ is called **nabla**, which comes from a Greek word for a certain type of harp that has a similar shape.

(a) Find the gradient $\nabla f(x, y)$.

(b) For each of the following points (x_0, y_0), evaluate the gradient $\nabla f(x_0, y_0)$ and sketch the gradient vector with its tail at (x_0, y_0). Some of the vectors are too long to fit onto the plot. To draw them to scale, you should scale each vector by a factor of $1/2$.

- $(x_0, y_0) = (2, 0)$
- $(x_0, y_0) = (0, 2)$
- $(x_0, y_0) = (2, 2)$
- $(x_0, y_0) = (2, 1)$
- $(x_0, y_0) = (-3, 2)$
- $(x_0, y_0) = (-2, -4)$
- $(x_0, y_0) = (0, 0)$

(c) Write a few sentences about how the direction of the gradient at each of these points is related to the the contour passing through that point.

(d) Does the output of f increase or decrease in the direction of $\nabla f(x_0, y_0)$? Use examples from the points above to write a couple of sentences that justify your answer.

Hint. You may wish to think about the contour plot as a topographical map and whether your elevation would increase or decrease if you took a step in the direction of the gradient.

Because $\nabla f(x_0, y_0)$ is a vector, we can consider its direction and length separately. The next subsection discusses the information conveyed by these geometric aspects about the behavior of f near (x_0, y_0).

11.8.5 Vector Properties of the Gradient

Key Idea 11.8.12 shows how we can separate the directional derivative of a function of two variables into two separate parts: the gradient vector evaluated at the point of interest and the unit vector in the direction we want to change the inputs of the function. Recall from equation (9.4.1) that the dot product of two vectors depends on the lengths of the vectors and the angle between the vectors. If θ is the angle between $\nabla f(x_0, y_0)$ and $\vec{u}$ (where $\vec{u}$ is a unit vector), then combining (11.8.4) and (9.4.1) gives

$$\begin{aligned} D_{\vec{u}} f(x_0, y_0) &= \nabla f(x_0, y_0) \cdot \vec{u} \\ &= \|\nabla f(x_0, y_0)\| \, \|\vec{u}\| \cos(\theta). \end{aligned}$$

Remember that $\|\vec{u}\| = 1$ because $\vec{u}$ is a unit vector. Hence, the directional derivative is the length of the gradient vector times the cosine of the angle between $\nabla f(x_0, y_0)$ and $\vec{u}$:

$$D_{\vec{u}} f(x_0, y_0) = \|\nabla f(x_0, y_0)\| \cos(\theta) \tag{11.8.5}$$

Equation (11.8.5) will be *extremely* useful in interpreting the gradient geometrically.

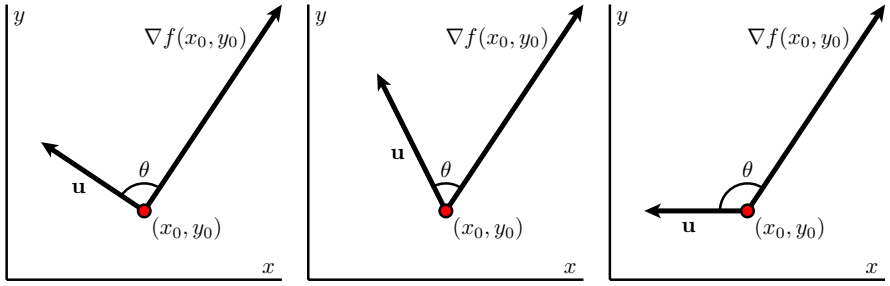


Figure 11.8.14 The sign of $D_{\vec{u}} f(x_0, y_0)$ is determined by θ

Because the magnitude of a vector is always non-negative, the sign of a directional derivative depends on $\cos(\theta)$. Figure 11.8.14 graphically shows examples of the following statements (from left to right):

- If the angle between the gradient and $\vec{u}$ is a right angle, then the directional derivative is zero.
- If the angle between the gradient and $\vec{u}$ is acute, then the directional derivative is positive.
- If the angle between the gradient and $\vec{u}$ is obtuse, then the directional derivative is negative.

The first statement explains why the gradient is perpendicular to the level curve through the point of interest. Because a level curve is the set of points for which the function has a particular output value, the output will not change along the level curve. Thus, the directional derivative in a direction tangent to the level curve must be zero. We can expand this explanation to the other statements as well. The output of f increases in any direction that makes an acute angle with the gradient vector and the output of f decreases in any direction that makes an obtuse angle with the gradient.

Activity 11.8.4 For this activity, we will look at the value of the directional derivative for temperature as a function of location. When looking at a plot of thermoclines (locations with the same temperature), you notice that your town is exactly on the thermocline labeled as 70 degrees, shown as the point P on Figure 11.8.15

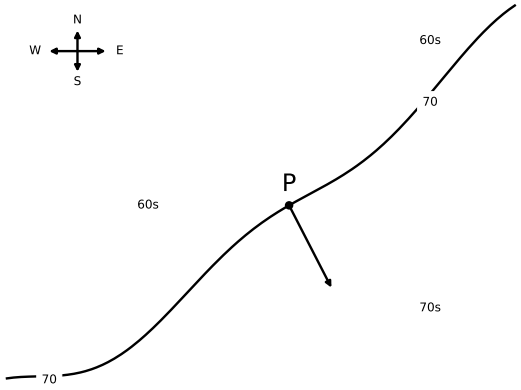


Figure 11.8.15 A plot of the thermocline corresponding to 70 degrees going through the point P

(a) Write a few sentences about why the gradient at P goes in the southeast direction and not the northwest (which is also perpendicular to the

thermocline/ level curve at P).

(b) Draw a vector in each of the following directions starting at P and state whether the directional derivative will be positive/negative/0 in each of these directions. You should write a couple of sentences for each direction vector about how the sign of the directional derivative is related to the angle between the direction and the gradient vectors.

i. West

ii. South

iii. Northeast

(c) Write a few sentences about why any direction moving up and left will have negative rate of change for the temperature in the direction given.

(d) Explain what direction you would need to move in to have the greatest positive rate of change for the temperature. Write a couple of sentences to relate your answer to the gradient using Equation (11.8.5).

Here are some questions about how the directional derivative changes at a point that we will address:

- What direction corresponds to the *largest* value for the directional derivative?
- What is the largest value possible for the directional derivative at (x_0, y_0)?
- What direction corresponds to the *smallest* (or most negative) value for the directional derivative?
- What is the smallest value possible for the directional derivative at (x_0, y_0)?

All of these questions can be answered by considering equation (11.8.5). Because the length of the gradient vector does not change when changing direction, the $\cos(\theta)$ factor tells us when we have the maximum and minimum values. Because of how we define the angle between vectors, we only need to consider values of θ between 0 and π. Therefore, the largest value of the directional derivative occurs when $\theta = 0$ and $\cos(\theta)$ attains its maximum value of 1. This means that when the direction vector $\vec{u}$ is in the same direction as the gradient, the directional derivative is maximized. Furthermore, the largest value of the directional derivative is the length of the gradient. In equation form, we can express that if $\vec{u}$ is in the same direction as $\nabla f(x_0, y_0)$, then

$$D_{\vec{u}} f(x_0, y_0) = \|\nabla f(x_0, y_0)\| \cos(0) = \|\nabla f(x_0, y_0)\| .$$

By a parallel argument, the smallest (or most negative) value the directional derivative can take is when $\theta = \pi$. In this case, the direction vector is in the *opposite* direction of the gradient vector. Thus if $\theta = \pi$, then

$$D_{\vec{u}} f(x_0, y_0) = \|\nabla f(x_0, y_0)\| \cos(\pi) = -\|\nabla f(x_0, y_0)\|$$

because $\cos(\pi) = -1$.

We summarize our most recent work by stating important facts about the gradient.

The Meaning of the Gradient as a Vector.

Let f be a differentiable function and (x_0, y_0) a point for which $\nabla f(x_0, y_0) \neq \vec{0}$.

- The gradient points in a direction perpendicular to the level curve $z = k$ where $k = f(x_0, y_0)$.
- The gradient $\nabla f(x_0, y_0)$ points in the direction of greatest rate of increase for f at (x_0, y_0), and the instantaneous rate of change of f in that direction is the length of the gradient vector.
- If $\vec{u} = \frac{1}{\|\nabla f(x_0,y_0)\|}\nabla f(x_0, y_0)$, then $\vec{u}$ is a unit vector in the direction of greatest increase of f at (x_0, y_0), and $D_{\vec{u}} f(x_0, y_0) = \|\nabla f(x_0, y_0)\|$.
- The gradient $\nabla f(x_0, y_0)$ points in the opposite direction of greatest rate of decrease for f at (x_0, y_0), and the instantaneous rate of change of f in that direction is the length of the gradient vector times -1.
- If $\vec{u} = -\frac{1}{\|\nabla f(x_0,y_0)\|}\nabla f(x_0, y_0)$, then $\vec{u}$ is a unit vector in the direction of greatest decrease of f at (x_0, y_0), and $D_{\vec{u}} f(x_0, y_0) = -\|\nabla f(x_0, y_0)\|$.

Note that the third and fifth bullets above are algebraic statements of the second and fourth bullets, respectively.

Activity 11.8.5 We will look algebraically calculating and explaining the vector properties of the gradient using the function $f(x, y) = xy - 2y^2 + \frac{x^2}{2} + 2y - x$ at the input $(1, 1)$.

(a) Calculate the gradient of f

(b) Calculate $\nabla f(1, 1)$.

(c) Calculate $D_{\vec{u}} f(1, 1)$ for each of the following vectors:

i. $\vec{u} = \langle -1, 0\rangle$

ii. $\vec{u} = \langle \frac{3}{5}, -\frac{4}{5}\rangle$

iii. $\vec{u} = \langle -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\rangle$

(d) What direction will correspond to the greatest rate of increase for f at the input $(1, 1)$? Find the value of the directional derivative in this direction.

Exercises 21–23 look at how the directional derivative changes as a function of the direction used or of the point where the directional derivative is being evaluated.

11.8.6 Applications

The gradient has many natural applications. For example, situations often arise where we are interested in knowing the direction in which a function is increasing or decreasing most rapidly. This type of question is natural when constructing a road through the mountains or planning the flow of water across a landscape. In the next activity, we examine how the gradient can help with

navigation to the top of a mountain in foggy conditions.

Activity 11.8.6 This activity considers directional derivatives and gradients for a function that measures elevation. Suppose you are hiking in a foggy park and can only see a few feet in front of you. There is nothing blocking you from walking in any particular direction, but because of the fog you cannot see where the highest point on the mountain is. You want to try to find the top of the mountain, but you don't have a map, trail, or line of sight to other landmarks. You do have a compass, which works in the fog. This allows you to identify the cardinal directions.

In order to use calculational tools from multivariable calculus, you overlay an xy-coordinate system on the park. The x-coordinate measures east/west location, with eastward movement associated with increasing x-values. The y-coordinate measures north/south location, with northward movement associated with increasing y-values. The elevation in meters above sea level of the park's terrain at the location (x, y) is given by a function $h(x, y)$.

(a) Let P_1 be your current location in the foggy park. You use your compass to find the east and north directions. At P_1, you find that the ground rises 1 meter per 50 meters traveled to the east and the ground rises 2.5 meters per 50 meters traveled to the north.

Use this information to find $\nabla h(P_1)$.

(b) Use your answer to the previous part to say what direction is "uphill" at P_1 and state the rate of elevation increase in this direction.

(c) You decide to walk uphill from your location P_1 in order to try to find the top of the mountain. After walking in the same direction for a while, you are at a point P_2 and notice that you are no longer walking in the steepest direction. You again locate east and north and measure the steepness of the mountain in these directions. You find that the ground rises 1.5 meters per 75 meters traveled to the east and the ground goes down 0.5 meters per 100 meters traveled to the north.

Use this new information to calculate $\nabla h(P_2)$, find the uphill direction, and find how steep the mountain is in the uphill direction at P_2.

(d) Suppose you continue this method of walking in the uphill direction for a while, finding the new uphill direction, and walking in the new uphill direction. Do you think you must eventually find the top of the mountain? How you will know that you have reached the top of the mountain? Remember that you can't see very far in front of you. Write a few sentences to explain your reasoning.

The technique described in the previous activity has many applications related to maximizing or minimizing functions. For example, consider a two-dimensional version of how a heat-seeking missile might work. (This application is borrowed from United States Air Force Academy Department of Mathematical Sciences.) Suppose that the temperature surrounding a fighter jet can be modeled by the function T defined by

$$T(x, y) = \frac{100}{1 + (x - 5)^2 + 4(y - 2.5)^2},$$

where (x, y) is a point in the plane of the fighter jet and $T(x, y)$ is measured in degrees Celsius. Some contours and gradients ∇T are shown on the left in Figure 11.8.16.

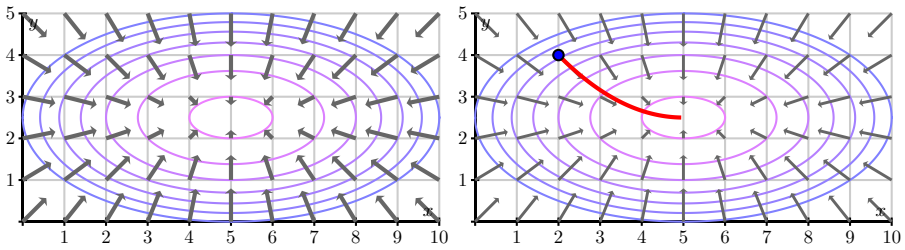


Figure 11.8.16 Contours and gradient for $T(x, y)$ and the missile's path.

A heat-seeking missile will always travel in the direction in which the temperature increases most rapidly; that is, it will always travel in the direction of the gradient ∇T. If a missile is fired from the point $(2, 4)$, then its path will be that shown on the right in Figure 11.8.16.

This type of strategy is sometimes called **gradient ascent**. The gradient decent method is used to find relative minimums for functions by taking steps in the direction opposite the gradient. Gradient decent methods are used extensively in economics and machine learning to find where the difference between predictions from a model and data are as small as possible (minimize error).

Summary

- The directional derivative of f at the point (x, y) in the direction of the unit vector $\vec{u} = \langle u_1, u_2 \rangle$ is

$$D_{\vec{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + u_1 h, y + u_2 h) - f(x, y)}{h}$$

  for those values of x and y for which the limit exists. In addition, $D_{\vec{u}} f(x, y)$ measures the slope of the graph of f when we move in the direction $\vec{u}$. Alternatively, $D_{\vec{u}} f(x_0, y_0)$ measures the instantaneous rate of change of f in the direction $\vec{u}$ at (x_0, y_0).

- The gradient of a function $f = f(x, y)$ at a point (x_0, y_0) is the vector

$$\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$$

- The directional derivative in the direction $\vec{u}$ may be computed by

$$D_{\vec{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \vec{u}$$

- At any point where the gradient is nonzero, the gradient is orthogonal to the contour through that point and points in the direction in which f increases most rapidly; moreover, the slope of f in this direction equals the length of the gradient $\|\nabla f(x_0, y_0)\|$. Similarly, the direction opposite of the gradient is the direction of greatest decrease, and that rate of decrease is the negative length of the gradient.

11.8.7 Exercises

1. If $f(x, y) = 4x^2 - 3y^2$, find the value of the directional derivative at the point $(1, 3)$ in the direction given by the angle $\theta = \frac{2\pi}{5}$.

2. ***(a)*** What is the rate of change of $f(x, y) = 2xy + y^2$ at the point $(4, 4)$ in the direction $\vec{v} = -\boldsymbol{i} + 2\boldsymbol{j}$?

$f_{\vec{v}} =$ _____

(b) What is the direction of maximum rate of change of f at $(4, 4)$?
direction = _____
(Give your answer as a vector.)

(c) What is the maximum rate of change?
maximum rate of change = _____

3. Find the gradient of the given function. Assume the variables are restricted to a domain on which the function is defined.

$$z = (x + 3y)\, e^{2y}$$

$\nabla z =$ _____

4. Find the gradient of the given function. Assume the variables are restricted to a domain on which the function is defined.

$$f(r, t) = r^3 \cos(t)\,.$$

$\nabla f =$ _____

5. Find the gradient of the given function. Assume the variables are restricted to a domain on which the function is defined.

$$f(a, b) = \frac{3a + 5b}{3a - 5b}.$$

$\nabla f =$ _____

6. Find the gradient of the function $f(r, h) = 2\pi rh + \pi r^2$ at the point $(3, 3)$.
$\nabla f(3, 3) =$ _____

7. Find the differential df of the function f given the gradient.

$$\nabla f = \left(4x + 2e^{-y}\right) \boldsymbol{i} + 2xe^{-y}\boldsymbol{j}.$$

$df =$ _____

8. Find the gradient ∇f of the function f given the differential.

$$df = \left(4x^3 + 1\right) ye^x dx + x^4 e^x dy$$

$\nabla f =$ _____

9. Suppose $f\,(x, y) = 2x^2 + 2xy + 2y^2$, $P = (2, 1)$, and $\mathbf{u} = \left(\frac{3}{5}, \frac{4}{5}\right)$.
A. Compute the gradient of f.
$\nabla f =$ _____ **i**+ _____ **j**
Note: Your answers should be expressions of x and y; e.g. "3x - 4y"
B. Evaluate the gradient at the point P.
$(\nabla f)\,(2, 1) =$ _____ **i**+ _____ **j**
Note: Your answers should be numbers
C. Compute the directional derivative of f at P in the direction $\boldsymbol{u}$.
$(D_{\mathbf{u}} f)\,(P) =$ _____
Note: Your answer should be a number

10. Suppose that $f(x, y) = \sin(4x + 5y)$ is a function of two variables.
1. Find the gradient of $f(x, y)$.
Answer: $\nabla f(x, y) =$ _____
2. Evaluate the gradient at the point $P(-10, 8)$.
Answer: $\nabla f(-10, 8) =$ _____
3. Find the rate of change of $f(x, y)$ at P in the direction of the vector

$\mathbf{u} = \frac{1}{2}\left(\sqrt{3}\mathbf{i} - \mathbf{j}\right)$.

Answer: $D_{\mathbf{u}} f(-10, 8) =$ _____

11. Suppose that $f(x, y, z) = x^2yz - xyz^3$ is a function of three variables.

1. Find the gradient of $f(x, y, z)$.

Answer: $\nabla f(x, y, z) =$ _____

2. Evaluate the gradient at the point $P(-2, 1, 1)$.

Answer: $\nabla f(-2, 1, 1) =$ _____

3. Find the rate of change of $f(x, y, z)$ at P in the direction of the vector $\mathbf{u} = \left\langle 0, \frac{4}{5}, -\frac{3}{5} \right\rangle$.

Answer: $D_{\mathbf{u}} f(-2, 1, 1) =$ _____

12. Find the directional derivative of the function $f(x, y) = \dfrac{x}{x^2 + y^2}$ at the point $(-2, -1)$ in the direction of the vector $\mathbf{v} = \langle 3, 4 \rangle$.

Answer: $D_{\mathbf{v}} f(-2, -1) =$ _____

13. The temperature at any point in the plane is given by $T(x, y) = \dfrac{130}{x^2 + y^2 + 1}$.

(a) What shape are the level curves of T?

◯ parabolas

◯ circles

◯ hyperbolas

◯ lines

◯ ellipses

◯ none of the above

(b) At what point on the plane is it hottest? _____

What is the maximum temperature? _____

(c) Find the direction of the greatest increase in temperature at the point $(-1, 1)$. _____

What is the value of this maximum rate of change, that is, the maximum value of the directional derivative at $(-1, 1)$? _____

(d) Find the direction of the greatest decrease in temperature at the point $(-1, 1)$. _____

What is the value of this most negative rate of change, that is, the minimum value of the directional derivative at $(-1, 1)$? _____

14. At a certain point on a heated metal plate, the greatest rate of temperature increase, 4 degrees Celsius per meter, is toward the northeast. If an object at this point moves directly north, at what rate is the temperature increasing?

_____ degrees Celsius per meter

15. Suppose that you are climbing a hill whose shape is given by $z = 774 - 0.06x^2 - 0.04y^2$, and that you are at the point (50, 90, 300).

In which direction (unit vector) should you proceed initially in order to reach the top of the hill fastest?

$\langle$_____, _____$\rangle$

If you climb in that direction, at what angle above the horizontal will you be climbing initially (radian measure)?

16. Are the following statements true or false?

(a) $f_{\vec{u}}(a,b) = ||\nabla f(a,b)||$.

(b) The gradient vector $\nabla f(a,b)$ is tangent to the contour of f at (a,b).

(c) $f_{\vec{u}}(a,b)$ is parallel to $\vec{u}$.

(d) $\nabla f(a,b)$ is a vector in 3-dimensional space.

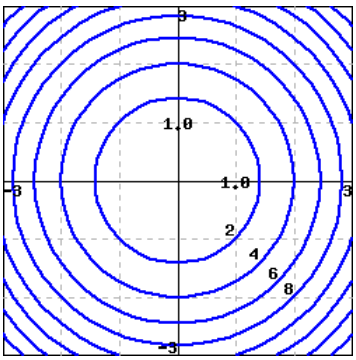
(e) If $\vec{u}$ is perpendicular to $\nabla f(a,b)$, then $f_{\vec{u}}(a,b) = \langle 0,0\rangle$.

(f) If $f(x,y)$ has $f_x(a,b) = 0$ and $f_y(a,b) = 0$ at the point (a,b), then f is constant everywhere.

(g) If $\vec{u}$ is a unit vector, then $f_{\vec{u}}(a,b)$ is a vector.

(h) Suppose $f_x(a,b)$ and $f_y(a,b)$ both exist. Then there is always a direction in which the rate of change of f at (a,b) is zero.

17. Use the contour diagram of f in the figure below to decide if the specified directional derivatives below are positive, negative, or approximately zero.



(a) At point $(-2,2)$, in direction $\vec{i}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

(b) At point $(0,-2)$, in direction $-\vec{j}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

(c) At point $(-1,1)$, in direction $-\vec{i}+\vec{j}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

(d) At point $(-1,1)$, in direction $\vec{i}+\vec{j}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

(e) At point $(0,-2)$, in direction $\vec{i}+2\vec{j}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

(f) At point $(0,-2)$, in direction $-\vec{i}-2\vec{j}$: $f_{\vec{u}}$ is (☐ positive ☐ negative ☐ approximately zero)

18. Find the maximum rate of change of $f(s,t) = te^{st}$ at the point $(0,4)$.

Answer: _____

19. Let $f(x,y) = x^3y^3$ and $c(t) = \left(t^2, 2t^3\right)$

(a) Calculate:

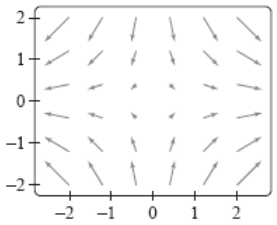
$\nabla f \cdot c'(t) =$ _____

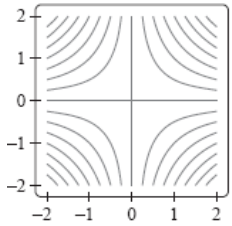
(b) Use the Chain Rule for Paths to evaluate $\frac{d}{dt}f(c(t))$ at $t = 1$.

$\frac{d}{dt} f(c(1)) =$_____

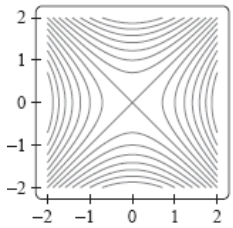
20. Which Figure is the contour plot of a potential function for the vector field **F**?

Recall that the gradient vectors are perpendicular to the level curves.





⇐ ***Figure 1***



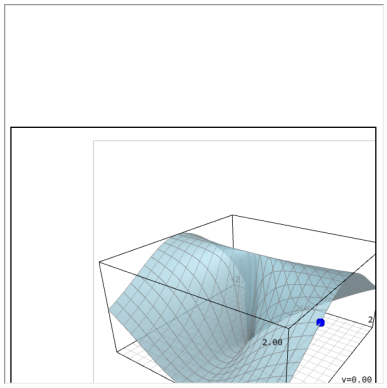
⇐Figure 2
Answer : Figure

◯ 2

◯ 1

Directional Derivative Sense-Making. The next three exercises consider how the directional derivative changes as a function of the direction used or of the point where the directional derivative is computed. The goal of these exercises is to understand the directional derivative as a geometric measurement using visually-based sense-making prompts in terms of the location and directions being used. In particular, Exercise 21 asks students to visually assess whether the directional derivative is positive/negative/zero at a few locations and directions. Exercise 22 fixes the location on a surface and steps students about how to construct a function that gives value of the directional derivative as a function of the direction. Exercise 23 examines the relationships between the value of the directional derivative and ideas like the slope of a secant line or dependence on the length of the direction vector. These exercises are adapted from materials developed by Rafael Martinez-Planell.

21. Figure 11.8.17 shows a graph of $z = f(x, y)$ that you should use for the following tasks.



Standalone
In Context
Embed

Figure 11.8.17 A plot of $f(x, y)$ with the measurements of change in input and output labeled for a change of inputs given by $\langle a, b\rangle$

(a) On the xy-plane of the figure below, draw the direction vector $\langle 1, 1\rangle$ starting at the point $(2, -1)$.

(b) Draw a piece of the tangent line to the graph of f at the point $(2, -1, f(2, -1))$ which is in the direction $\langle 1, 1\rangle$. The slope of the tangent line in the $\langle 1, 1\rangle$ direction is called the directional derivative of f at $(2, -1)$ in the $\langle 1, 1\rangle$ direction and is denoted $D_{\langle 1,1\rangle} f(2, -1)$.

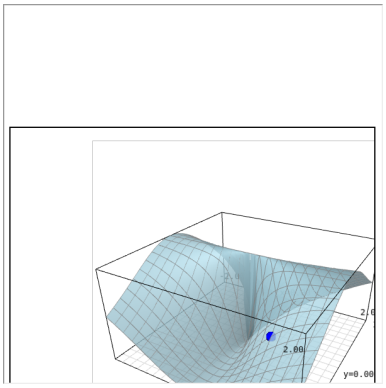
(c) Is $D_{\langle 1,1\rangle} f(2, -1)$ positive, negative, or zero? Justify your answer.

(d) On the xy-plane of (a new plot of) the figure below, draw the direction vector $\langle -\frac{1}{2}, 3\rangle$ starting at the point $(2, -1)$.

(e) Draw a piece of the tangent line to the graph of f at the point $(2, -1, f(2, -1))$ which is in the direction $\langle -\frac{1}{2}, 3\rangle$. The slope of the tangent line in the $\langle -\frac{1}{2}, 3\rangle$ direction is called the directional derivative of f at $(2, -1)$ in the $\langle -\frac{1}{2}, 3\rangle$ direction and is denoted $D_{\langle -\frac{1}{2},3\rangle} f(2, -1)$.

(f) Is $D_{\langle -\frac{1}{2},3\rangle} f(2, -1)$ positive, negative, or zero? Justify your answer.

22. Let f be the function whose graph appears in Figure 11.8.18. For each value of t the vector $\langle \cos(t), \sin(t)\rangle$ is a direction vector. The value of $D(t) = D_{\langle \cos(t),\sin(t)\rangle} f(1, -1)$ a scalar.



Standalone
In Context
Embed

Figure 11.8.18 A plot of $f(x, y)$ with the measurements of change in input and output labeled for a change of inputs given by $\langle a, b\rangle$

(a) What direction is the vector $\langle \cos(t), \sin(t)\rangle$ in for

i. $t = 0$
ii. $t = \dfrac{\pi}{2}$
iii. $t = \dfrac{3\pi}{2}$
iv. $t = \pi$

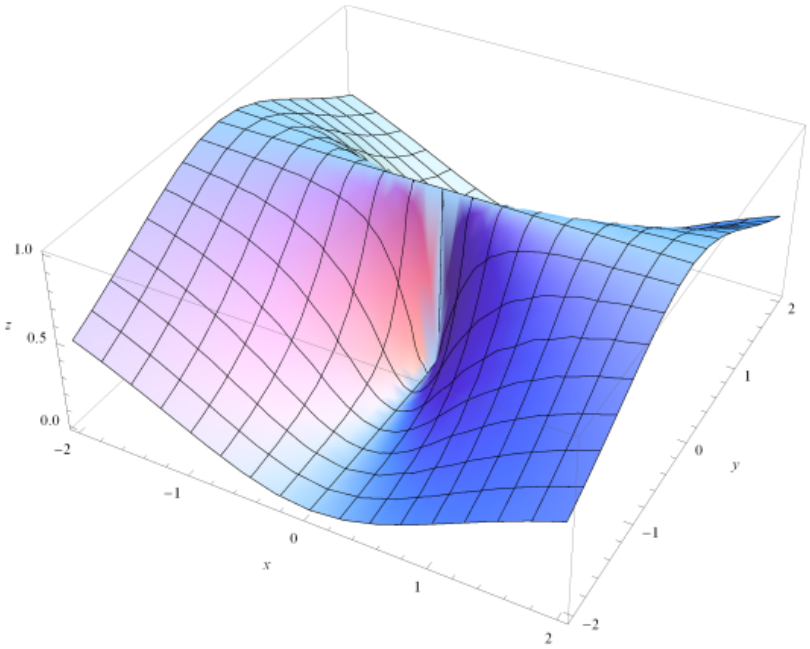
(b) Is the value of $D(t) = D_{\langle \cos(t), \sin(t)\rangle} f(1, -1)$ positive, negative, or zero for:

i. $t = 0$
ii. $t = \dfrac{\pi}{2}$
iii. $t = \dfrac{3\pi}{2}$
iv. $t = \pi$

(c) Draw the graph of $y = D(t)$ for values of t in the interval $[0, 2\pi]$. You should think about how the value of $D(t)$ changes between the values you considered above to help you draw the graph of $D(t)$. Your graph doesn't have to have exact values but should correctly identify where $D(t)$ is positive, negative, and zero.

(d) Draw the graph of $y = G(t) = D_{\langle 1,1\rangle} f(t, -t)$ for values of t in the interval $(0, 2]$. You should probably go through the same process you did above of looking at what $G(t)$ will be for a few values of t between 0 and 2, then think about how $G(t)$ will vary for values in between. The function $G(t)$ is different than $D(t)$.

23. The graph of $z = f(x, y)$ is as given below. In this problem, use geometric arguments to justify your answers.



(a) Draw the line segment that goes from $(1, -1, f(1, -1))$ to $(2, 0, f(2, 0))$. How does the slope of this line segment in that direction compare (smaller, equal, larger) with the value of $D_{\langle 1,1\rangle} f(1, -1)$?

(b) How does $D_{\langle 1,1\rangle} f(1, -1)$ compare with $D_{\langle 0.01,0.01\rangle} f(1, -1)$? Justify your answer.

(c) Which is closer to $D_{\langle 1,1\rangle} f(1, -1)$, the slope of the line segment that goes from $(1, -1, f(1, -1))$ to $(2, 0, f(0, 2))$ or the slope of the line segment that goes from $(1, -1, f(1, -1))$ to $(1.01, -0.99, f(1.01, -0.99))$? Explain your answer.

24. Let $E(x, y) = \frac{100}{1+(x-5)^2+4(y-2.5)^2}$ represent the elevation on a land mass at location (x, y). Suppose that E, x, and y are all measured in meters.

(a) Find $E_x(x, y)$ and $E_y(x, y)$.

(b) Let $\vec{u}$ be a unit vector in the direction of $\langle -4, 3\rangle$. Determine $D_{\vec{u}} E(3, 4)$. What is the practical meaning of $D_{\vec{u}} E(3, 4)$ and what are its units?

(c) Find the direction of greatest increase in E at the point $(3, 4)$.

(d) Find the instantaneous rate of change of E in the direction of greatest decrease at the point $(3, 4)$. Include units on your answer.

(e) At the point $(3, 4)$, find a direction $\vec{w}$ in which the instantaneous rate of change of E is 0.

25. Find all directions in which the directional derivative of $f(x, y) = ye^{-xy}$ is 1 at the point $(0, 2)$.

26. Find, if possible, a function f such that

$$\nabla f = \left\langle \sin(yz), xz\cos(yz) + 2y, xy\cos(yz) + \frac{5}{z} \right\rangle.$$

If not possible, explain why.

27. Let $f(x, y) = x^2 + 3y^2$.

(a) Find $\nabla f(x, y)$ and $\nabla f(1, 2)$.

(b) Find the direction of greatest increase in f at the point $(1, 2)$. Explain. A graph of the surface defined by f is shown at left in Figure 11.8.19. Illustrate this direction on the surface.

(c) A contour diagram of f is shown at right in Figure 11.8.19. Illustrate your calculation from (b) on this contour diagram.

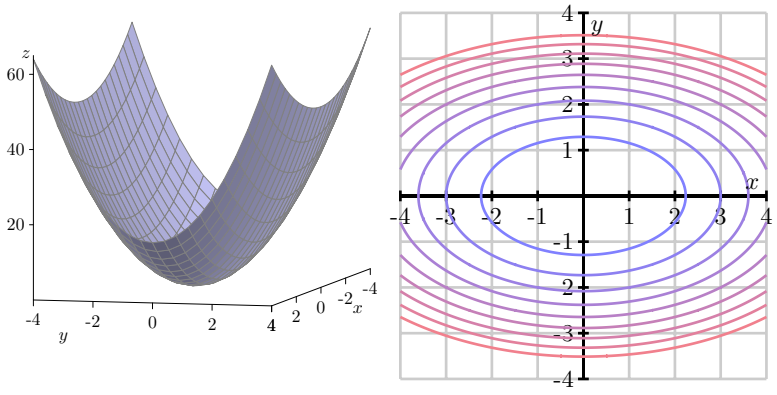


Figure 11.8.19 Left: Graph of $f(x, y) = x^2 + 3y^2$. Right: Contours.

(d) Find a direction $\vec{w}$ for which the derivative of f in the direction of $\vec{w}$ is zero.

28. Figure 11.8.20 shows a plot of the gradient ∇f at several points for some function $f = f(x, y)$.

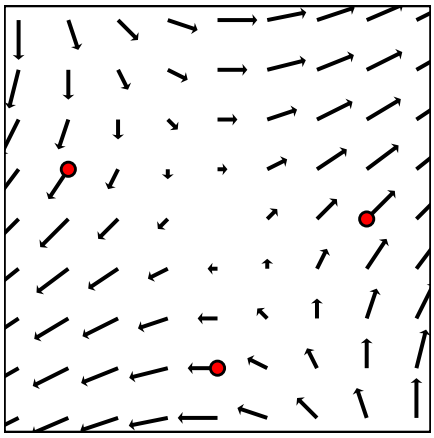


Figure 11.8.20 The gradient ∇f.

(a) Consider each of the three indicated points, and draw, as best as you can, the contour through that point.

(b) Beginning at each point, draw a curve on which f is continually decreasing.

11.9 Higher Dimensions

Motivating Questions

- How does the idea of a level curve for a function of two variables generalize to functions of three or more variables?
- How do measures of change such as partial derivatives, directional derivatives, and gradients generalize to functions of more than two variables?

11.9.1 Introduction

In this chapter, we have primarily worked with functions of two variables including understanding graphs of the form $z = f(x, y)$ and measuring change with functions of two variables. In this section, we explore how many of the tools we have developed for understanding two-variable functions can be generalized to functions of three or more variables.

A significant impediment to expanding the ideas in this chapter to functions of three or more variables will be that is it difficult (or impossible) to draw plots of these functions, which means we must rely on simplifications or analogies to understand these measurements geometrically. For this reason, most efforts to understand functions of three or more variables is algebraically-focused.

We start in the Preview Activity by generalizing the idea of level curves.

Preview Activity 11.9.1

(a) State the equation and shape of the level curves for the function $f(x, y) = x^2 + y^2$ for the values $-2, -1, 0, 1, 2$.

Hint. Some of these level sets will be empty, meaning that there are no (x, y) points that have the particular k-value as output.

(b) State the equation and shape of the level curves for the function $g(x, y) = x^2 - y^2$ for the values $-2, -1, 0, 1, 2$.

(c) For a constant k and a function f, the **level set** is the set of points P for which $f(P) = k$. The remainder of this Preview Activity asks you to consider level sets of some functions of three variables. We use the term "level set" here because the set of inputs for a function of three or more variables that gives a particular output value will likely not be a curve.

State the equation and shape of the level sets for the function $f(x, y, z) = x^2 + y^2 + z^2$ for $-2, -1, 0, 1, 2$.

Hint. Some of these level sets will be empty.

(d) State the equation and shape of the level sets for the function $g(x, y, z) = x^2 + y^2 - z^2$ for $-2, -1, 0, 1, 2$. These level set values should give three different types of surfaces.

(e) State the equation and shape of the level sets for the function $h(x, y, z) = x + y - z$ for $-2, -1, 0, 1, 2$.

In this section, we will often use the idea of level sets to explore functions of three or more variables. The topics we explore are not likely to surprise you: domain and range, graphs, limits, and rates of change.

11.9.2 Inputs and Outputs with Functions of Three or More Variables

A function is a rule that assigns an single output for each allowed input. For instance, your school likely uses a function that takes your student ID number as an input and outputs your name. While you may not be the only student at your school with your name, this function should not associate more than one name with each student ID number. A function of n variables can be thought of as taking points in $\mathbb{R}^n$ as input and outputs a real number for each valid input.

Recall from Definition 11.2.2 that the domain of a function is the set of input values for which a function is defined. The range of a function is the set of values actually output by the function.

Example 11.9.1 Consider the function $g(x, y, z) = \sqrt{1-x} + e^{y^2+z^2}$. We see here that we need $\sqrt{1-x}$ to be defined, which requires that $1-x \geq 0$ or $x \leq 1$. However, there is is no restriction on the values of y or z. Thus, we can say that the domain of g is all points (x, y, z) for which $x \leq 1$. You may be tempted to say the range of g is all real numbers since g outputs scalars. However, a closer look shows that g cannot output negative numbers. Since the outputs of the exponential portion of g is always positive and the square root portion is always nonnegative, the range of g is the interval $(0, \infty)$. ▲

11.9.3 Visualizing Functions of Three or More Variables

Given a function f of several variables the **level set** of f at the value k is the collection of input points P for which $f(P) = k$. A level set is a subset of the domain of the function f because the level set is a collection of input points.

If $f(x, y)$ is a function of two variables, then a level set is a contour or a level curve of the form $f(x, y) = k$. Notice that the level set in this case is a set of points in the xy-plane. As we saw in Section 11.2, we can make a contour plot by graphing a collection of these level curves on the same two-dimensional plot, which will give us an idea of what the surface corresponding to $z = f(x, y)$ looks like. In other words, we were able to give enough information in a two-dimensional plot to describe a three-dimensional surface.

If $g(x, y, z)$ is a function of three variables, then a graph of g function would require four dimensions: three dimensions for the inputs and one dimension for the output. Drawing plots in three dimensions is challenging, but we will use three-dimensional graphs to understand functions of three variables rather than attempting to visualize four-dimensional graphs directly. We will do this by generalizing our approach of using contour plots to express a three-dimensional plot in a two-dimensional setting. Specifically, we will use level sets of $g(x, y, z)$ to understand a graph of $w = g(x, y, z)$ and help measure change in the output of g.

A level set of g is a collection of points in xyz-space that corresponds to a surface of the form $g(x, y, z) = k$. As you saw in Preview Activity 11.9.1, the level sets of a function of three variables are often surfaces, which we call **level surfaces**. For example, if $g(x, y, z) = x^2 + y^2 + z^2$, then the level surface corresponding to the value 1 is all points that satisfy the equation $x^2 + y^2 + z^2 = 1$. This set of points forms the sphere of radius 1 centered at the origin.

These level surfaces will often be a different kind of surface than we have been working with throughout Chapter 11. Surfaces of the form $z = f(x, y)$ are called **explicit surfaces** because one of the coordinates can be solved

explicitly as a function of the other coordinates. Surfaces such as spheres or hyperboloids are *not* explicit surfaces because there is no way to solve for one variable as a function of the other coordinates. Two perspectives from which you can recognize that are first that these surfaces fail the vertical line test in $\mathbb{R}^3$ and algebraically, the $\pm$ that comes from needing to take square roots prevents this for many quadric surfaces.

Surfaces like spheres and hyperboloids are called **implicit surfaces** because they can be described as the set of points that satisfy an implicit equation. For instance, a hyperboloid of one sheet can be described by an equation of the form

$$\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1+\frac{(z-z_0)^2}{c^2}.$$

Conveniently, level sets for functions of several variables have an associated implicit equation, $g(x,y,z)=k$.

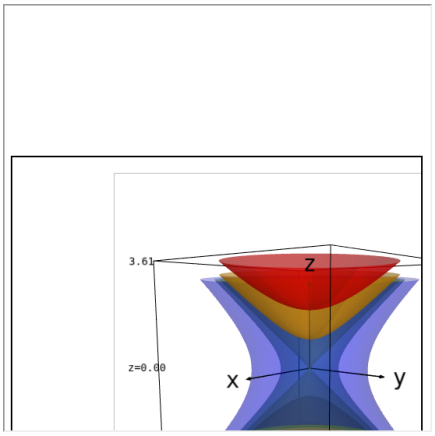
This insight goes the other way as well. Any implicit surface can be thought of as a level set for some function of three variables. If S_1 is the surface described by $x^2-xyz=y^2z^3-2$, then we can also think of S_1 as the level set of the three-variable function $F(x,y,z)=x^2-xyz-y^2z^3$ for the output value -2 because any point in three dimensions that satisfies $x^2-xyz=y^2z^3-2$ must also satisfy $x^2-xyz-y^2z^3=-2$.

Example 11.9.2 Analogous to how a two-dimensional contour plot provides useful information about the corresponding three-dimensional surface plot, we would like to see if we can graph multiple level sets of a three variable function to understand the four-dimensional graph of a function of the form $w=g(x,y,z)$. In this example, we will work with the three-variable function $g(x,y,z)=x^2+y^2-z^2$.

If we consider the level sets corresponding to the values $k=\{-4,-1,0,1,4\}$, we get the following level surfaces:

- If $k=-4$, then we are looking at the implicit equation $x^2+y^2-z^2=-4$ which corresponds to a hyperboloid of two-sheets given by $\frac{z^2}{4}-\frac{x^2}{4}-\frac{y^2}{4}=1$.
- If $k=-1$, then we are looking at the implicit equation $x^2+y^2-z^2=-1$ which corresponds to a hyperboloid of two-sheets given by $z^2-x^2-y^2=1$.
- If $k=0$, then we are looking at the implicit equation $x^2+y^2-z^2=0$ which corresponds to a cone given by $x^2+y^2=z^2$.
- If $k=1$, then we are looking at the implicit equation $x^2+y^2-z^2=1$ which corresponds to a hyperboloid of one-sheet given by $x^2+y^2-z^2=1$.
- If $k=4$, then we are looking at the implicit equation $x^2+y^2-z^2=-4$ which corresponds to a hyperboloid of one-sheet given by $\frac{x^2}{4}+\frac{y^2}{4}-\frac{z^2}{4}=1$.

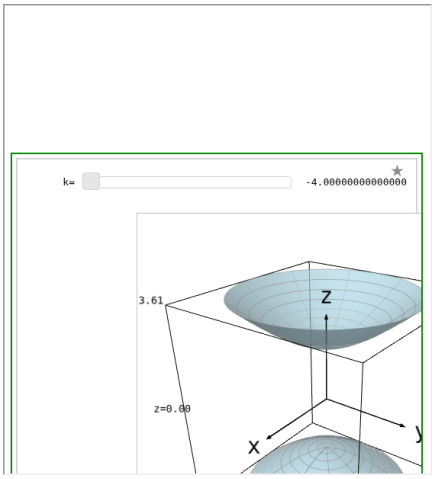
We can plot these level sets together to create the analogous plot to a contour plot. Figure 11.9.3 shows a graph of these surfaces with colors that go from red (at $k=-4$) to blue (at $k=4$). You can see how a three-dimensional plot of multiple level sets gets visually cluttered very quickly, but you can see how the output of g increases as you move away from the z-axis or if you move closer to the xy-plane. Later in this section, we will explore how these level surfaces are related to questions such as "In what direction will the output of g have the greatest rate of increase?"



Standalone
In Context
Embed

Figure 11.9.3 A plot of the level surfaces for $g(x, y, z) = x^2 + y^2 - z^2$ for values $k = \{-4, 1, 0, 1, 4\}$ in colors from red to blue, respectively.

While the plot of multiple level surfaces may be a bit overwhelming, we can show each of these level sets in a separate plot. Figure 11.9.4 plots a level surface for the value given by the slider at the top. Use the slider to change the value of k over the entire range from -4 to 4. As you do so, look at how the shape of the corresponding level surface morphs from a hyperboloid of two sheets to a cone to a hyperbola of one sheet.



Standalone
In Context
Embed

Figure 11.9.4 A plot of the level surface for $g(x, y, z) = x^2 + y^2 - z^2$ with a slider for changing the level curve value from -4 to 4

Figure 11.9.4 may look like a three-dimensional plot, but we really have a four-dimensional plot with slider providing the axis for the fourth dimension. We saw a similar case in Figure 10.2.3 where the plot of $\vec{r}(t)$ was a four dimensional idea. In that plot, one dimension was the input changing, with the t-axis for the slider. The other three dimensions were represented by the vector $\vec{r}$ in standard position, which we can also think of as being a plot of the vector's terminal point. With vector-valued functions of one variables, we simplified the four-dimensional plot to include *all* of the three-dimensional outputs, frequently ignoring plotting how those outputs depended on the value of the parameter t. ▲

Activity 11.9.2 In this activity, we will look at level surfaces created by two three-variable functions and consider the direction in which the output is increasing or decreasing as fast as possible.

(a) Let $f(x, y, z) = 2x - y + z$.

(i) Write a couple of sentences to describe shape and characteristics will the level surfaces of f. Be sure to include all possibilities for different types of level surfaces. Compare what is the same or different about these level surfaces and relate the differences to the level value, k.

(ii) If you are at the point $(1, 1, -1)$, in what direction should you change the input of f to see the greatest increase in the output of f? You should think about how a change in each variable will change the output of f and talk about this in your reasoning.

(b) Let $g(x, y, z) = x^2 + y^2 + z^2$.

(i) Write a couple of sentences to describe shape and characteristics will the level surfaces of g. Be sure to include all possibilities for different types of level surfaces. Compare what is the same or different about these level surfaces and relate the differences to the level value, k.

(ii) If you at the point $(2, 0, -2)$, in what direction should you change the input of g to see the greatest increase in the output of g? You should think about how a change in each variable will change the output of g and talk about this in your reasoning.

11.9.4 Measuring Change with Functions of Three or More Variables

Conceptually, limits for functions of three or more variables work the same as for functions of two variables. Given a function f, we say that f **has limit** L **as the inputs approach** P_0 provided that we can make $f(P)$ as close to L as we like by taking P sufficiently close (but not equal) to P_0. We write

$$\lim_{P \to P_0} f(P) = L.$$

If the limit of a function f along every path through an input point P_0 exists and all of those limits are the same value L then we say the limit of f at P is L. There is not much more insight into limits of multivariable functions to be had at this point, so we will move on to measuring the change in output of our multivariable functions of three or more variables.

The partial derivative of a multivariable function measures the rate of change in the output of the function when one variable is changed and all others are held constant. Our definition and notation of partial derivatives given for functions of two variables only needs to be updated to account for three or more input variables.

Example 11.9.5 If we consider $T(x, y, z, t)$ to be a function that measures the air temperature at a location with spatial coordinates (x, y, z) at time t, then we have four first partial derivatives:

$$T_x = \frac{\partial T}{\partial x} \qquad T_y = \frac{\partial T}{\partial y} \qquad T_z = \frac{\partial T}{\partial z} \qquad T_t = \frac{\partial T}{\partial t}$$

The limit definition of the partial derivative with respect to t at a point (x_0, y_0, z_0, t_0) is

$$\frac{\partial T}{\partial z} = \lim_{h \to 0} \frac{T(x_0, y_0, z_0, t_0 + h) - T(x_0, y_0, z_0, t_0)}{h}.$$

This partial derivative measures the instantaneous rate of change in temperature with respect to time at the location (x_0, y_0, z_0) at time $t = t_0$.

Second partial derivatives work the same as for functions with two variable functions. For example,

$$T_{yt} = \frac{\partial T_y}{\partial t} = \frac{\partial}{\partial t}\left[\frac{\partial T}{\partial y}\right]$$

measure the rate of change with respect to time of T_y. This function has sixteen second partial derivatives: four unmixed partials and 12 mixed partials. ▲

Clairaut's Theorem also generalizes to higher dimensions: if all mixed partials are continuous near an input point of a function of three or more variables, then the mixed partials at that point are equal.

In Section 11.6, we saw how the ideas of locally linear functions and differentiability meant that for small scales, the change in the output of a function could be expressed as a linear combination of the changes in each variable separately. This connected to the idea that the linearization of a function approximates the original function well for small neighborhoods around the point at which the linearization was created. We visualized this linearization with the tangent plane and saw that geometrically speaking, the tangent plane was a good approximation for the surface at small scales.

Unfortunately, we will not be able to use this kind of visualization for functions of three or more variables because our graphing techniques are limited to looking at plots of level surfaces. However, we can use linearizations to approximate the change in functions of three or more variables. For example, if T is the four-variable function of Example 11.9.5, then the linearization of this function at location (a, b, c) and time $t = d$ is

$$\begin{aligned} L(x, y, z, t) = T(a, b, c, d) &+ T_x(a, b, c, d)(x - a) \\ &+ T_y(a, b, c, d)(y - b) + T_z(a, b, c, d)(z - c) \\ &+ T_t(a, b, c, d)(t - d). \end{aligned}$$

Notice this is the same form as we have used for all of our linear ideas: the change in the output is a linear combination of the changes with respect to a change in a single input variable, while holding all other inputs constant.

We can easily generalize related ideas like the differential to functions of three or more variables. The differential for the temperature function is given by

$$dT = \frac{\partial T}{\partial x}dx + \frac{\partial T}{\partial y}dy + \frac{\partial T}{\partial z}dz + \frac{\partial T}{\partial t}dt.$$

Other important ideas, like the classic calculus approach or our interpretation of the gradient, did not depend on using a function of two variables, so we can adapt these arguments and results to higher dimensional cases as well. Again, the downside to this abstraction is that we do not have the geometric tools to help us interpret and understand these measurements in higher dimensions.

11.9.5 Directional Derivatives and Gradients

Directional derivatives and gradients for functions of three or more variables are critical concepts, both in this course and in future coursework in math, economics, and the physical sciences. Recall that the directional derivative measures the instantaneous rate of change of a multivariable function when the inputs are changed in a particular direction. None of the arguments in Section 11.8 were specific to functions of two variables, but rather than stating these results in terms of an abstract function of n variables, we will illustrate the various definitions and results in terms of functions of three or four variables: $f(x, y, z)$ or $g(x, y, z, w)$.

Let $f = f(x, y, z)$ be a function of three variables. The **derivative of f at the point (x, y, z) in the direction of the unit vector $\vec{u} = \langle u_1, u_2, u_3 \rangle$** is denoted $D_{\vec{u}} f(x, y, z)$ and is given by

$$D_{\vec{u}} f(x, y, z) = \lim_{t \to 0} \frac{f(x + u_1 t, y + u_2 t, z + u_3 t) - f(x, y, z)}{t} \tag{11.9.1}$$

for those values of x, y, and z for which the limit exists. We can make a similar limit definition for a function of more than three variables because we are able to separate the length of the step in a particular direction (t in the above statement) and the unit vector in that particular direction ($\vec{u}$ from above) for vectors with any number of components.

We can calculate the directional derivative in terms of partial derivatives of the function and $\vec{u}$. This result comes from using the chain rule on a composition of the multivariable function with the line in the direction of $\vec{u}$. If $f(x, y, z)$ and $g(x, y, z, w)$ are functions of three and four variables, respectively, then

$$D_{\vec{u}} f(x, y, z) = f_x(x, y, z)u_1 + f_y(x, y, z)u_2 + f_z(x, y, z)u_3$$

and

$$\begin{aligned} D_{\vec{u}} g(x, y, z, w) = g_x(x, y, z, w)u_1 + g_y(x, y, z, w)u_2 \\ + g_z(x, y, z, w)u_3 + g_w(x, y, z, w)u_4. \end{aligned}$$

Remember that the direction vector will have as many components as there are inputs to the function because the direction vector corresponds to a change in the inputs of the function.

The formulas above have the same form of a dot product of the gradient and the direction vector, so we more compactly write

$$D_{\vec{u}} f(x, y, z) = \nabla f \cdot \langle u_1, u_2, u_3 \rangle$$

and

$$D_{\vec{u}} g(x, y, z, w) = \nabla g \cdot \langle u_1, u_2, u_3, u_4 \rangle$$

where $\nabla f = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z) \rangle$ and

$$\nabla g = \langle g_x(x, y, z, w), g_y(x, y, z, w), g_z(x, y, z, w), g_w(x, y, z, w) \rangle.$$

This generalization shows that in any dimension, the directional derivative can be calculated as the dot product of the gradient and the direction vector.

This also means that all of our work to understand the meaning of the gradient will generalize to any dimension as well. In particular, we update our summary of the meaning of the gradient below.

The Meaning of the Gradient as a Vector.

Let f be a differentiable function and P a point for which $\nabla f(P) \neq \vec{0}$.

- The gradient points in a direction perpendicular to the level set $f(P) = k$.
- The gradient $\nabla f(P)$ points in the direction of greatest rate of increase for f at P, and the instantaneous rate of change of f in that direction is the length of the gradient vector.
- If $\vec{u} = \frac{1}{\|\nabla f(P)\|} \nabla f(P)$, then $\vec{u}$ is a unit vector in the direction of

greatest increase of f at P, and $D_{\vec{u}}f(P) = \|\nabla f(P)\|$.

- The gradient $\nabla f(P)$ points in the opposite direction of greatest rate of decrease for f at (P), and the instantaneous rate of change of f in that direction is the length of the gradient vector times -1.
- If $\vec{u} = -\frac{1}{\|\nabla f(P)\|}\nabla f(P)$, then $\vec{u}$ is a unit vector in the direction of greatest decrease of f at (P), and $D_{\vec{u}}f(P) = -\|\nabla f(P)\|$.

The first idea above is useful when finding an equation for the plane tangent to an implicit surface. Let P_1 be a point on S_1, an implicit surface given by $F(P) = k$. The tools of Section 11.6 will not work because our surface is not of the form $z = f(x, y)$. Because $\nabla F(P_1)$ is perpendicular to the surface S_1, $\nabla F(P_1)$ is a normal vector for the tangent plane at P_1. This leads to the following key idea.

Key Idea 11.9.6 *Let S_1 be an implicit surface given by $F(x, y, z) = k$ and let $P_0 = (x_0, y_0, z_0)$ be a point on S_1. The tangent plane to S_1 at P_0 is given by*

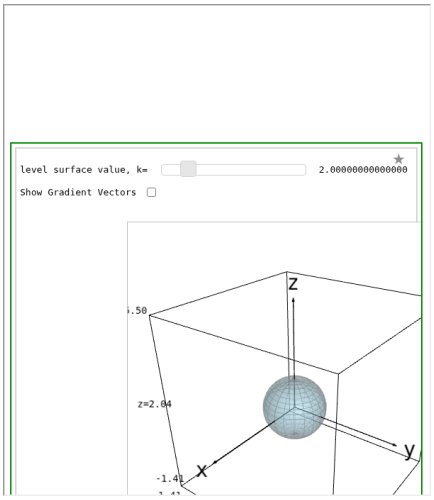
$$\frac{\partial F}{\partial x}(P_0)(x - x_0) + \frac{\partial F}{\partial y}(P_0)(y - y_0) + \frac{\partial F}{\partial z}(P_0)(z - z_0) = 0.$$

Proof. By Equations of a plane we need to find a normal vector for the plane and a point on the plane. The point $P_0 = (x_0, y_0, z_0)$ is the location where the tangent plane will be tangent to S_1. Thus, we only need to find a normal vector for the plane. By the interpretation of the gradient above, $\nabla F(P_0) = \langle \frac{\partial F}{\partial x}(P_0), \frac{\partial F}{\partial y}(P_0), \frac{\partial F}{\partial z}(P_0)\rangle$ is perpendicular to S_1 because this is a level surface. Therefore,

$$\frac{\partial F}{\partial x}(P_0)(x - x_0) + \frac{\partial F}{\partial y}(P_0)(y - y_0) + \frac{\partial F}{\partial z}(P_0)(z - z_0) = 0$$

is an equation for this tangent plane in scalar form. ■

Example 11.9.7 This example examines the function $f(x, y, z) = x^2 + y^2 + z^2$ and the meaning of its gradient $\nabla f = \langle 2x, 2y, 2z\rangle$. In order to visualize the graph of f we can look at a level surface given by $f(x, y, z) = k$. Figure 11.9.8 shows a plot of the level surface for a value of k. Note that the level surfaces of f are spheres of radius $R = \sqrt{k}$ centered at the origin. Move the slider at the top of Figure 11.9.8 to change the value of k. This illustrates how the scale of the level surface changes but the shape does not.



Standalone
In Context
Embed

Figure 11.9.8 A plot of a level surfaces for $g(x, y, z) = x^2 + y^2 + z^2$ with the option to plot gradient vectors for some points on the level surface

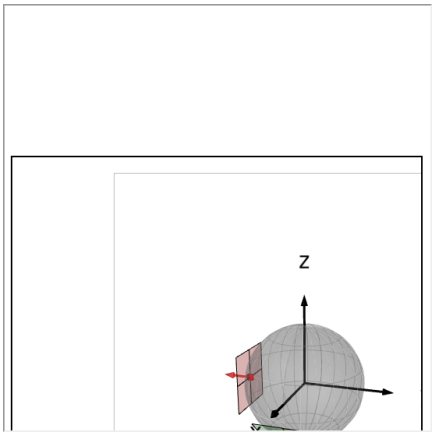
Use the checkbox in Figure 11.9.8 to show the gradient vector plotted at a collection of points on the level surface. If you turn on the gradient vectors, you can see that the gradient vector is perpendicular to the level surface and this relationship is true regardless of what value of k is used for the level surface. Two other properties should be noticed as well. First, the gradient always points *out* of the sphere (in the direction going away from the origin). Second, the length of the gradient increases as k increases.

The function f takes a point (x, y, z) as input and outputs the square of the distance to this point from the origin. This means that f increases as the point considered moves away from the origin, but the direction of *greatest* increase is to move directly away from the origin. The rate of increase of f also increases as the point gets farther from the origin. Both of these ideas are demonstrated in Figure 11.9.8 by the red gradient vectors. Specifically, you can see that length of the gradient vectors increases as you increase k and the direction of the gradient vectors is always perpendicular to the level surface/ sphere. You can also reason that the direction for greatest decrease in f will be in the direction going toward the origin.

We can verify Key Idea 11.9.6 by looking at tangent points for a few points on these spheres as level surfaces. Let S_1 be the level surface with value 1, $1 = f(x, y, z) = x^2 + y^2 + z^2$. Note that $P_0 = (0, -1, 0)$ is on S_1 and $\nabla f(P_0) = \langle 0, -2, 0 \rangle$. By Key Idea 11.9.6, the tangent plane to S_1 at P_0 is given by

$$(0)(x - 0) + (-2)(y + 1) + (0)(z - 0) = 0 \to y = -1.$$

This point, normal vector, and tangent plane is shown in red in Figure 11.9.9.



Standalone
In Context
Embed

Figure 11.9.9

If we consider the point $P_1 = (\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2})$, which his also on S_1, we can apply the same argument. Note that $\nabla f(P_1) = \langle \sqrt{2}, 0, -\sqrt{2} \rangle$, which means that the tangent plane has equation

$$(\sqrt{2})\left(x - \frac{\sqrt{2}}{2}\right) + (0)(y-0) + (-\sqrt{2})\left(z - \frac{\sqrt{2}}{2}\right) = 0 \Rightarrow x - z = \sqrt{2}.$$

This point, normal vector, and tangent plane is shown in green in Figure 11.9.9. ▲

Example 11.9.10

(a) Suppose we wish to find an equation for the tangent plane to the hyperboloid of one sheet with equation $x^2 + y^2 - z^2 = 1$ at the point $(-1, 0, 1)$. Because this is an implicit surface for which we cannot solve for z as a function of x and y, the techniques we developed in Section 11.6 do not apply directly. However, we can use the gradient vector as we did in the previous example. To do so, we recognize that the surface is a level surface of the three-variable function $g(x, y, z) = x^2 + y^2 - z^2$ for the value 1. Thus, we know that the gradient vector for g at the point $(-1, 0, 1)$ is normal to the surface at that point and can be used as the normal vector for the tangent plane.

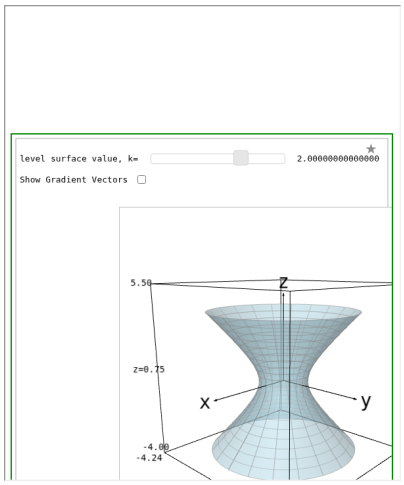
We can compute $\nabla g = \langle 2x, 2y, -2z \rangle$, which tells us that $\nabla g(-1, 0, 1) = \langle -2, 0, -2 \rangle$ is normal to the hyperboloid of one sheet at $(-1, 0, 1)$. Thus, we can write an equation for the tangent plane to this surface at this point as

$$-2(x+1) + 0(y-0) - 2(z-1) = 0.$$

(b) A particularly nice aspect of the approach of viewing surfaces in three-dimensional space as level surfaces of functions of three variables is that the method does not change even as the surface changes. To continue with the function g from the previous part, we saw in Example 11.9.2 that different values of k for the level surfaces gives rise to different types of surfaces. In particular,

- if $k < 0$, the level surface is a hyperboloid of two sheets (as we saw in the previous part),
- if $k = 0$, the level surface is a cone, and
- if $k > 0$, the level surface is a hyperboloid of one sheet.

Even as the type of surface changes, the gradient's algebraic form $\nabla g = \langle 2x, 2y, -2z \rangle$ does not change. If you select the checkbox in Figure 11.9.11 to plot gradient vectors at points along the level surface, you will see that the gradient vectors are perpendicular to the level surface. Use the slider to change the value of the level surface, which changes the shape of the level surface when the sign of k changes.



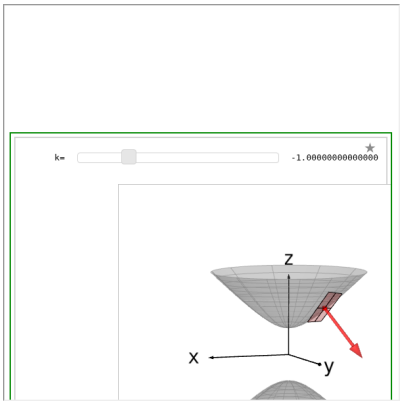
Standalone
In Context
Embed

Figure 11.9.11 A plot to view level surfaces for $g(x, y, z) = x^2 + y^2 - z^2$ with the option to plot gradient vectors for some points on the level surface

To see this in action, we can find a general way to express the tangent plane at a point on these level surfaces. Let S_k be the level surface with value k. In other words, $g(x, y, z) = k = x^2 + y^2 - z^2$. Let $P_k = (-1, 1, \sqrt{2-k})$ be a point, which you can verify for yourself is on S_k. This means that $\nabla g(P_k) = \langle -2, 2, -2\sqrt{2-k} \rangle$. Hence by Key Idea 11.9.6, the tangent plane to S_k at P_k can be written as

$$(-2)(x+1) + (2)(y-1) + (-2\sqrt{2-k})(z - \sqrt{2-k}) = 0.$$

This point, the gradient vector (drawn at half length), and tangent plane are shown in red in Figure 11.9.12. Use the slider to change the value of k using the slider at the top of Figure 11.9.12 to see how even though the shape of the level surface changes, the gradient vector at P_k remains orthogonal to the surface and will serves as a normal vector for our tangent plane.



Standalone
In Context
Embed

Figure 11.9.12 A plot of a level surface for $g(x, y, z) = x^2 + y^2 - z^2$ along with the normal vector and tangent plane at the point $(-1, 1, \sqrt{2-k})$

▲

Activity 11.9.3 In this activity, we will look at the calculation and interpretation of the gradient for $f(x, y, z) = x + y - z$ and $g(x, y, z) = z - x^2 - y^2$.

(a) Calculate ∇f and ∇g.

(b) Sketch the level surfaces of f for the values $k = \{-2, -1, 0, 1, 2\}$. Write a few sentences about the shape of each of these level surfaces and describe how the level surfaces change in terms of the value of k.

(c) Write a few sentences about how the direction and magnitude of ∇f is related to the level surfaces from the previous part.

(d) Sketch the level surfaces of g for the values $k = \{-2, -1, 0, 1, 2\}$. Write a few sentences about the shape of each of these level surfaces and describe how the level surfaces change in terms of the value of k. You may find it helpful to notice that each of the level surfaces can be expressed with z as a function of x and y.

(e) Write a few sentences about how the direction and magnitude of ∇g is related to the level surfaces from the previous part.

The level surfaces of the function $g(x) = z - x^2 - y^2$ from the previous activity were of the form $z = x^2 + y^2 + C$ for some constant C. This means that the gradient of g gives a convenient way to calculate a vector perpendicular to the surface $z = x^2 + y^2 + C$. In fact, for any surface of the form $z = h(x, y)$, the same surface will be a level surface of the three variable function $g(x, y, z) = z - h(x, y)$.

Summary

- The set of points (x, y, z) for which a function of three variables, $f(x, y, z)$, has a constant output, k, is called the level surface and a implicit equation of the level surface is $f(x, y, z) = k$.
- The partial derivatives for functions of functions three or more variables are calculated by measuring the rate of change in one input variable while holding all other inputs constant. The gradient of a function of three or more variables is a vector-valued, multivariable function where each component is the first partial derivative of with respect to a different coordinate. If $g(x_1, x_2, \ldots, x_n)$, then the gradient of g has n components

and is given by $\nabla g = \left\langle \frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}, \ldots, \frac{\partial g}{\partial x_m} \right\rangle$.

- The directional derivative of a function of three or more variables measures the rate of change of the output when the input variables are changed in a particular direction and can be calculated by $Dg_{\vec{u}}(P) = \nabla g(P) \cdot \vec{u}$.
- The gradient of a multivariable function g at an input P points in the direction of greatest rate of increase for the output of g, which also means that the gradient at P, $\nabla g(P)$ will be perpendicular to the level surface corresponding to the value $k = g(P)$. The direction opposite of the gradient, $-\nabla g(P)$, will be the direction of greatest rate of decrease at P. The greatest rate of increase and decrease are $\|\nabla g(P)\|$ and $-\|\nabla g(P)\|$, respectively.

11.9.6 Exercises

1. Evaluate the function at the specified points.
$h(x, y, z) = xy\frac{1}{z^2}, (2, 5, 5), (-6, -2, 4)$
At $(2, 5, 5)$: _____
At $(-6, -2, 4)$: _____

2. Find the limit, if it exists, or type N if it does not exist.
$$\lim_{(x,y,z)\to(2,5,2)} \frac{5ze^{x^2+y^2}}{2x^2 + 5y^2 + 2z^2} = _____$$

3. Find the limit, if it exists, or type N if it does not exist.
$$\lim_{(x,y,z)\to(0,0,0)} \frac{2xy + 3yz + 4xz}{4x^2 + 9y^2 + 16z^2} = _____$$

4. Find the limit, if it exists, or type N if it does not exist.
$$\lim_{(x,y,z)\to(0,0,0)} \frac{3xy + 2yz + xz}{9x^2 + 4y^2 + z^2} = _____$$

5. Find the limit, if it exists, or type N if it does not exist.
$$\lim_{(x,y,z)\to(4,5,3)} \frac{2ze^{x^2+y^2}}{4x^2 + 5y^2 + 3z^2} = _____$$

6. Let $f(x, y, z) = \frac{x^2-5y^2}{y^2+4z^2}$. Then

$f_x(x, y, z) =$ _____
$f_y(x, y, z) =$ _____
$f_z(x, y, z) =$ _____

7. Find the first partial derivatives of $f(x, y, z) = z \arctan(\frac{y}{x})$ at the point (5, 5, -4).
A. $\frac{\partial f}{\partial x}(5, 5, -4) =$ _____
B. $\frac{\partial f}{\partial y}(5, 5, -4) =$ _____
C. $\frac{\partial f}{\partial z}(5, 5, -4) =$ _____

8. If $\sin(5x + 4y + z) = 0$, use implicit differentiation to find the first partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at the point (0, 0, 0).
A. $\frac{\partial z}{\partial x}(0, 0, 0) =$ _____
B. $\frac{\partial z}{\partial y}(0, 0, 0) =$ _____

9. Find the first partial derivatives of the function $f(x, y, z) = 2x \sin(y - z)$.

1. $\dfrac{\partial f}{\partial x} =$ _____
2. $\dfrac{\partial f}{\partial y} =$ _____
3. $\dfrac{\partial f}{\partial z} =$ _____

10. Find the first partial derivatives of the function $w = 3ze^{xyz}$.

1. $\dfrac{\partial w}{\partial x} =$ _____
2. $\dfrac{\partial w}{\partial y} =$ _____
3. $\dfrac{\partial w}{\partial z} =$ _____

11. Given $F(r, s, t) = r\left(5t^3 - 4s^2\right)$, compute:
$F_{rst} =$_____

12. Given $F(r, s, t) = -r\left(2s^4 + t^2\right)$, compute:
$F_{rst} =$_____

13. For each function below, consider its level surfaces, where $w = \cdots -2, -1, 0, 1, 2 \cdots$ if negative w's are allowed, or $w = 1, 2, 3, \cdots$ if negative w's are not allowed. Match each function with the verbal description of its level surfaces by placing the letter of the verbal description to the left of the number of the function.

(a) $w = \sqrt{(x^2 + y^2 + z^2)}$

(b) $w = \sqrt{(x + 2y + 3z)}$

(c) $w = x^2 + 2y^2 + 3z^2$

(d) $w = x^2 + y^2 + z^2$

(e) $w = x^2 - y^2 - z^2$

(f) $w = x + 2y + 3z$

(g) $w = \sqrt{(x^2 + 2y^2 + 3z^2)}$

A. a collection of equally spaced concentric spheres

B. a collection of equally spaced parallel planes

C. a collection of concentric ellipsoids

D. a collection of unequally spaced concentric spheres

E. two cones and two collections of hyperboloids

F. a collection of unequally spaced parallel planes

14. Find a formula for a function $g(x, y, z)$ whose level surfaces are planes parallel to the plane $z = 6x + 7y - 3$.
$g(x, y, z) =$ _____

15. Find a function $f(x, y, z)$ whose level surface $f = 6$ is the graph of the function $g(x, y) = 6x + 4y$.
$f(x, y, z) =$ _____

16. Describe in words the level surfaces of $g(x, y, z) = e^{-\left(x^2 + (y+3)^2 + z^2\right)}$. Think what the intersections of these surfaces with the three coordinate

planes look like.

(a) In symbolic form, these level surfaces are given by $e^{-\left(x^2+(y+3)^2+z^2\right)} = c$ where c is a constant. For what values of c are level surfaces defined?

$c \in$ _____

(Give your answer as an interval or list of intervals; for example, if c is less than 1 or greater than or equal to 2, enter ***(-inf,1),[2,inf)****.)*

(b) Pick a level surface that intersects the xz-plane. What is your value of c?

$c =$ _____

Write the equation for the intersection, as an expression set equal to a constant:

_____ = _____

(c) Pick a level surface that intersects the yz-plane. What is your value of c (which might be different from that you used in ***(a)***?

$c =$ _____

Write the equation for the intersection, as an expression set equal to a constant:

_____ = _____

(d) Pick a level surface that intersects the xy-plane. What is your value of c?

$c =$ _____

Write the equation for the intersection, as an expression set equal to a constant:

_____ = _____

17. The concentration of salt in a fluid at (x, y, z) is given by $F(x, y, z) = 4x^2 + 3y^4 + 3x^2z^2$ mg/cm^3. You are at the point $(1, -1, 1)$.

(a) In which direction should you move if you want the concentration to increase the fastest?

direction: _____

(Give your answer as a vector.)

(b) You start to move in the direction you found in part (a) at a speed of 2 cm/sec. How fast is the concentration changing?

rate of change = _____

18. Consider the surface

$$16x^2 + 4y^2 + 1z^2 = 21$$

and the point $P = (1, 1, 1)$ on this surface.

A. Starting with the equation $x = 1 + 32t$, find equations for y and z which combine with this equation to give parametric equations for the normal line through P.

$y =$ _____

$z =$ _____

Note: Your answers should be expressions of t; e.g. “3x - 4y”

B. Find an equation for the tangent plane through P.

$z =$ _____

Note: Your answers should be expressions of x and y; e.g. ”3xy + 2y”

19. If $f(x, y, z) = 2zy^2$, then the gradient at the point $(2, 3, 2)$ is

$\nabla f(2, 3, 2) =$ _____

20. Consider the function $f(x, y, z) = xy + yz^2 + xz^3$.

Find the gradient of f:

$\langle$_____, _____, _____$\rangle$

Find the gradient of f at the point (-2, 2, 3).

$\langle$_____, _____, _____$\rangle$

Find the rate of change of the function f at the point (-2, 2,3) in the direction $\mathbf{u} = \langle 4/\sqrt{36}, -2/\sqrt{36}, -4/\sqrt{36}\rangle$.

21. Compute the gradient vector fields of the following functions:

A. $f(x, y) = 4x^2 + 4y^2$

$\nabla f(x, y) =$ _____ **i**+ _____ **j**

B. $f(x, y) = x^2y^6$,

$\nabla f(x, y) =$ _____ **i**+ _____ **j**

C. $f(x, y) = 4x + 4y$

$\nabla f(x, y) =$ _____ **i**+ _____ **j**

D. $f(x, y, z) = 4x + 4y + 2z$

$\nabla f(x, y) =$ _____ **i**+ _____ **j**+ _____ **k**

E. $f(x, y, z) = 4x^2 + 4y^2 + 2z^2$

$\nabla f(x, y, z) =$ _____ **i**+ _____ **j**+ _____ **k**

22. Calculate the gradient of $h(x, y, z) = x^3y\frac{1}{z}$

$\nabla h =$ _____

23. The temperature at a point (x, y, z) is given by

$$T(x, y, z) = 1300e^{-x^2-2y^2-z^2}$$

where T is measured in °C and x, y, and z in meters.

1. Find the rate of change of the temperature at the point $P(2, -2, 2)$ in the direction toward the point $Q(3, -4, 3)$.

Answer: $D_{\overrightarrow{PQ}} f(2, -2, 2) =$ _____

2. In what direction does the temperature increase fastest at P?

Answer: _____

3. Find the maximum rate of increase at P.

Answer: _____

24. Find the gradient of the function $f(x, y, z) = 4xe^{\frac{y}{2}} \sin(3z)$.

grad $f =$ _____

25. Find the gradient of the function

$f(p, q, r) = e^{\frac{r}{5}} + \ln(q) + e^p$.

grad$f =$ _____

26. Find the gradient of the function $f(x, y, z) = z^2 \ln(yx)$, at the point $(e, 1, -1)$

$\nabla f(e, 1, -1) =$ _____

27. Find the gradient of the function $f(x, y, z) = xy^4$, at the point $(-1, 1, 0)$

$\nabla f(-1, 1, 0) =$ _____

28. Find the directional derivative of $f(x, y, z) = xz + y^3$, at $(1, 2, 3)$ in the direction of $\vec{v} = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

$f_{\vec{u}} =$ _____

29. Find the directional derivative of $f(x, y, z) = xz + y^3$ at the point $(1, 3, 2)$ in the direction of a vector making an angle of $\frac{2\pi}{3}$ with $\nabla f(1, 3, 2)$.

$f_{\vec{u}}(1, 3, 2) =$ _____

30. Check that the point $(-1, -1, 1)$ lies on the given surface. Then, viewing the surface as a level surface for a function $f(x, y, z)$, find a vector normal to the surface and an equation for the tangent plane to the surface at $(-1, -1, 1)$.

$$2x^2 - y^2 + 4z^2 = 5$$

vector normal = _____
tangent plane:
$z =$ _____

31. If the gradient of f is $\nabla f = 2x\,\mathbf{i} + z^2\,\mathbf{j} + 3yz\,\mathbf{k}$ and the point $P = (-2, 7, 2)$ lies on the level surface $f(x, y, z) = 0$, find an equation for the tangent plane to the surface at the point P.
$z =$ _____

32. Consider the ellipsoid $x^2 + 4y^2 + z^2 = 21$.
The implicit form of the tangent plane to this ellipsoid at $(-1, 2, -2)$ is _____.
The parametric form of the line through this point that is perpendicular to that tangent plane is $L(t) =$ _____.

33. Let $f(x, y) = 2x^2 - 4xy - 3y^2$.
Then an implicit equation for the tangent plane to the graph of f at the point $(-2, -2)$ is _____.

34. Find the differential of the function $w = x^6 \sin(y^6 z^5)$
dw =_____dx+ _____dy+ _____dz

35. Consider the ellipsoid $x^2 + 5y^2 + z^2 = 18$. Find all the points where the tangent plane to this ellipsoid is parallel to the plane $2x - 5y - 3z = 0$.
(If there are several points, separate them by commas.)

Hint. Hint:*and*

36. Find the differential of the function $w = x \sin\left(1yz^1\right)$.
dw = _____ dx + _____ dy + _____ dz
Note: Your answers should be expressions of x, y and z; e.g. "3xy + 4z"

37. Find the differential of the function $w = x \sin\left(1yz^3\right)$.
dw = _____ dx + _____ dy + _____ dz
Note: Your answers should be expressions of x, y and z; e.g. "3xy + 4z"

38. Find the linear approximation to $f(x, y, z) = \frac{xy}{z}$ at the point $(-2, -1, -1)$:
$f(x, y, z) \approx$_____

39. Find the directional derivative of the function $f(x, y, z) = xe^y + ye^z + ze^x$ at the point $(0, 0, 0)$ in the direction of the vector $\mathbf{v} = \langle -1, -2, 1 \rangle$.
Answer: $D_{\mathbf{v}} f(0, 0, 0) =$ _____

40. Find the directional derivative of the function $f(x, y, z) = \sqrt{xyz}$ at the point $(3, 2, 6)$ in the direction of the vector $\mathbf{v} = \langle 1, -2, -1 \rangle$.
Answer: $D_{\mathbf{v}} f(3, 2, 6) =$ _____

41. The temperature at a point (x,y,z) is given by $T(x, y, z) = 200e^{-x^2 - y^2/4 - z^2/9}$, where T is measured in degrees Celsius and x,y, and z in meters. There are lots of places to make silly errors in this problem; just try to keep track of what needs to be a unit vector.
Find the rate of change of the temperature at the point (-1, -1, 1) in the direction toward the point (-3, 4, -4).
In which direction (unit vector) does the temperature increase the fastest at (-1, -1, 1)?
$\langle$_____, _____,_____$\rangle$

What is the maximum rate of increase of T at (-1, -1, 1)?

42. Find the first partial derivatives of $f(x, y, z) = z \arctan(\frac{y}{x})$ at the point (4, 4, -5).

A. $\frac{\partial f}{\partial x}(4, 4, -5) =$ _____

B. $\frac{\partial f}{\partial y}(4, 4, -5) =$ _____

C. $\frac{\partial f}{\partial z}(4, 4, -5) =$ _____

43. Your monthly car payment in dollars is $P = f(P_0, t, r)$, where $\$P_0$ is the amount you borrowed, t is the number of months it takes to pay off the loan, and r percent is the interest rate.

(a) Is $\partial P/\partial t$ positive or negative? (☐ positive ☐ negative)

Suppose that your bank tells you that the magnitude of $\partial P/\partial t$ is 45.

What are the units of this value? _____

*(For this problem, write our your units in full, writing **dollars** for \$, **months** for months, **percent** for %, etc. Note that fractional units generally have a plural numerator and singular denominator.)*

(b) Is $\partial P/\partial r$ positive or negative? (☐ positive ☐ negative)

Suppose that your bank tells you that the magnitude of $\partial P/\partial r$ is 20.

What are the units of this value? _____

*(For this problem, write our your units in full, writing **dollars** for \$, **months** for months, **percent** for %, etc. Note that fractional units generally have a plural numerator and singular denominator.)*

For both parts of this problem, be sure you can explain what the practical meanings of the partial derivatives are.

44. Suppose that the temperature in a region of space is described by

$$T(x, y, z) = 100e^{-x^2-y^2-z^2}$$

and that you are standing at the point $(1, 2, -1)$.

(a) Find the instantaneous rate of change of the temperature in the direction of $\vec{v} = \langle 0, 1, 2\rangle$ at the point $(1, 2, -1)$. Remember that you should first find a *unit* vector in the direction of $\vec{v}$.

(b) In what direction from the point $(1, 2, -1)$ would you move to cause the temperature to decrease as quickly as possible?

(c) How fast does the temperature decrease in this direction?

(d) Find a direction in which the temperature does not change at $(1, 2, -1)$.

45. Consider the surface $x^2 - y^2 + z^2 = 4$.

(a) View the surface as a level surface for a function $f(x, y, z)$ and find a normal vector to the surface at the point $(-1, 1, 2)$.

(b) Now, find an equation for the tangent plane to the surface at $(-1, 1.2)$.

46. Find the equation of the tangent plane to the surface given by $xz + 2x^2y + y^2z^3 = 11$ at the point $(2, 1, 1)$.

47. Suppose that $\nabla f_P = \langle 2, -4, 4\rangle$. Is f increasing or decreasing at P in the direction $\langle 2, 1, 3\rangle$?

48. Find the derivative of the function at the given point in the direction of the vector $\vec{v}$.

(a) $f(x, y) = e^x \sin(y)$ at the point $(0, \pi/3)$ in the direction $\vec{v} = \langle -6, 8 \rangle$

(b) $g(r, s) = \tan^{-1}(rs)$ at the point $(1, 2)$ in the direction $\vec{v} = \langle 5, 10 \rangle$

(c) $h(r, s, t) = \ln(3r + 6s + 9t)$ at the point $(1, 1, 1)$ in the direction $\vec{v} = \langle 4, 12, 6 \rangle$

11.10 Optimization

Motivating Questions

- What properties will a local maximum or minimum of $f(x, y)$ have?
- How can we determine whether critical points of $f(x, y)$ are local maxima or minima?
- How can we find the absolute maximum and minimum of $f(x, y)$ on a closed and bounded domain?

11.10.1 Introduction

In single-variable calculus we learned that the derivative is useful for finding the local maxima and minima of functions. We also saw how these ideas can be employed in applied settings. In particular, if a function f, such as the one shown in Figure 11.10.1 is differentiable everywhere, we know that the tangent line is horizontal at any point where f has a local maximum or minimum. Horizontal tangent lines occur when the derivative f' is zero. Thus, one way to find possible extreme values of a function is to first find where the function's derivative is zero. Not every point where $f'(a) = 0$ will correspond to either a local maximum or local minimum, however. You may also recall that the concavity at a location where the derivative is zero can be used to help determine if a point is a local maximum or local minimum.

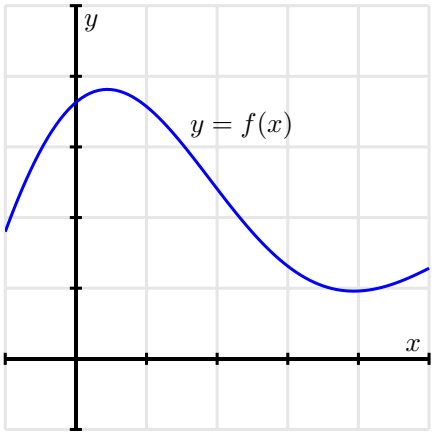


Figure 11.10.1 The graph of $y = f(x)$.

In multivariable calculus, we are often similarly interested in finding the greatest or least values that a function may achieve. Moreover, there are many applied settings in which a quantity of interest depends on several variables. In the preview activity below, we will use the application of elevation as a function of location to consider the characteristics that are necessary for an input to a function of two variables to correspond to a local maximum or local minimum.

An emphasis of this activity is that we can only look at measurements at our current location to determine if a point is a local maximum or minimum.

Preview Activity 11.10.1 In this activity, we will use a function that takes a two-dimensional location as its input and outputs the elevation of a surface at that point to explore what properties a point corresponding to a local maximum or local minimum of a function of two variables must have. Similar to Activity 11.8.6, imagine that you are hiking in a foggy park and cannot see anything more than a few feet in front of you. There is nothing blocking you from walking in any particular direction, but because of the fog, you cannot see where the highest point on the mountain is. You want to try to find the top of the mountain, but you don't have a map to consult or trail to follow or even a line of sight to other landmarks. You do have a compass that works in the fog, which you can use to determine directions. You also have a level that can be placed on the ground at your feet to measure how steep the change in elevation is in a particular direction. In other words, the level can measure the directional derivative at a location. Conceptually, our goal in this preview activity is to understand how the measurements you can take with your level allows you to identify when you have reached the top of the mountain.

(a) In Activity 11.8.6, you saw how you can move toward the top of the mountain by taking steps in the "uphill" direction, which is the direction of greatest rate of increase for elevation. Suppose you are standing at the top of the mountain, which corresponds to a local maximum for elevation, and you set the level at your feet while facing east. Will the level measure that the elevation is increasing, decreasing, or constant? Write a couple of sentences to justify your answer.

Hint. Remember that the level measures the *instantaneous* rate of change of elevation at your current location, but will not describe what happens for a small step to the east.

(b) If you are at the top of the mountain and set the level at your feed while facing south, will the level measure that the elevation is increasing, decreasing, or constant? Write a couple of sentences to justify your answer.

(c) Write a few sentences to explain why your level *must* show a zero rate of change in *every* direction at the top of the mountain.

(d) If you took a small step to the west from the top of the mountain, should your level measurement to the west at your new location indicate that the elevation is increasing, decreasing, or constant? Explain your reasoning.

(e) If you took a small step in *any direction* from the top of the mountain, should your level measurement in the direction you took the step at your new location indicate that the elevation is increasing, decreasing, or constant? Explain your reasoning.

(f) Now imagine that you are at the *lowest* point in the foggy park. Write a couple of sentences to explain why your level *must* show a zero rate of change in *every* direction.

(g) If you took at step in *any direction* from the lowest point in the park, should your level measurement at your new location in the the direction in which you took the step indicate that the elevation is increasing, decreasing, or constant? Explain your reasoning.

(h) Is it possible on your hike in the fog that you find a point where your level shows constant elevation in every direction but you are *not* at the top of the mountain or lowest point on the mountain? Explain what the the terrain would look like at this type of location.

As the Preview Activity demonstrates, points corresponding to local minima and local maxima must have a zero slope in *every* direction and the rate of the change after taking a small step in any direction must also be positive or negative, respectively. This will form the conceptual basis of our attempt to find and classify extreme points of functions of two variables.

11.10.2 Extrema and Critical Points

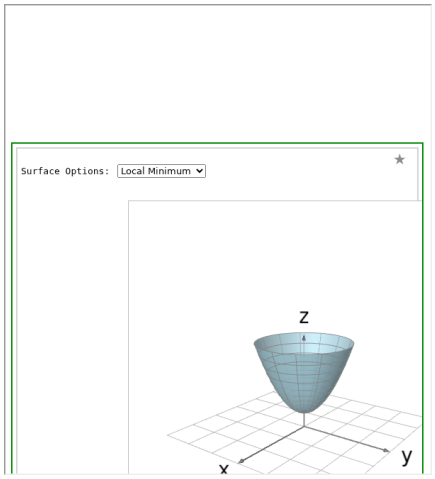
One of the important applications of single-variable calculus is the use of derivatives to identify local extremes of functions (that is, local maxima and local minima). Using the tools we have developed so far, we can extend the concept of local maxima and minima to multivariable functions.

Definition 11.10.2 Let f be a function of two variables x and y.

- The function f has a *local maximum* at a point (x_0, y_0) provided that $f(x, y) \leq f(x_0, y_0)$ for all points (x, y) near (x_0, y_0). In this situation we say that $f(x_0, y_0)$ is a *local maximum value*.
- The function f has a *local minimum* at a point (x_0, y_0) provided that $f(x, y) \geq f(x_0, y_0)$ for all points (x, y) near (x_0, y_0). In this situation we say that $f(x_0, y_0)$ is a *local minimum value*.
- An *absolute maximum point* is a point (x_0, y_0) for which $f(x, y) \leq f(x_0, y_0)$ for all points (x, y) in the domain of f. The value of f at an absolute maximum point is the *maximum value* of f.
- An *absolute minimum point* is a point such that $f(x, y) \geq f(x_0, y_0)$ for all points (x, y) in the domain of f. The value of f at an absolute minimum point is the *minimum value* of f.

◆

We use the term *extremum point* to refer to any point (x_0, y_0) at which f has a local maximum or minimum. In addition, the function value $f(x_0, y_0)$ at an extremum is called an *extremal value*. Figure 11.10.3 illustrates the graphs of two functions that have an absolute maximum and minimum, respectively, at the origin $(x_0, y_0) = (0, 0)$.



Standalone
In Context
Embed

Figure 11.10.3 An local maxima and a local minima

In single-variable calculus, we saw that the extrema of a continuous function f always occur at **critical numbers**, which are values of x where f fails to be differentiable or where $f'(x) = 0$. Said differently, critical numbers provide the locations where the extrema of a function may appear. Our work in Preview Activity 11.10.1 suggests that something similar happens with two-variable functions.

Suppose that a continuous function f has an extremum at (x_0, y_0). In this case, the trace $f(x, y_0)$ has an extremum at x_0, which means that x_0 is a critical number of $f(x, y_0)$. Therefore, either $f_x(x_0, y_0)$ does not exist or $f_x(x_0, y_0) = 0$. Similarly, either $f_y(x_0, y_0)$ does not exist or $f_y(x_0, y_0) = 0$. This implies that the extrema of a two-variable function occur at points that satisfy the following definition.

Definition 11.10.4 A **critical point** (x_0, y_0) of a function $f = f(x, y)$ is a point in the domain of f at which either $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ are both zero or at least one of $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ fails to exist. ◆

In the analogy of Preview Activity 11.10.1, these critical points correspond to locations where our level will measure no change in elevation for every direction. How can we ensure that the directional derivative will be zero in every direction? If $\nabla f = \vec{0}$, then $D_{\vec{u}} f = \vec{0} \cdot \vec{u} = 0$ for *every direction*. Geometrically, if $\nabla f = \vec{0}$ at a point (x_0, y_0), then the tangent plane at this point will be horizontal, meaning its equation has the form $z = c$ for a constant c.

This means that we can find critical points of a function f by computing partial derivatives and identifying any values of (x, y) for which one of the partials doesn't exist or for which both partial derivatives are simultaneously zero. For the latter, note that we have to solve the system of equations

$$\begin{aligned} f_x(x, y) &= 0 \\ f_y(x, y) &= 0. \end{aligned}$$

Example 11.10.5 In this example, we will look for the critical points of the two-variable function $f(x, y) = 4x^4 + y^2 - 16xy$. By the definition of critical points above, we will need to solve for the (x, y) points with both f_x and f_y equal to zero. There are no points for which the partial derivatives of f are undefined for this function because $f_x = 16x^3 - 16y$ and $f_y = 2y - 16x$ are polynomials. The critical points will be solutions to the system of equations

$$\begin{aligned} 16x^3 - 16y &= 0 \\ 2y - 16x &= 0. \end{aligned}$$

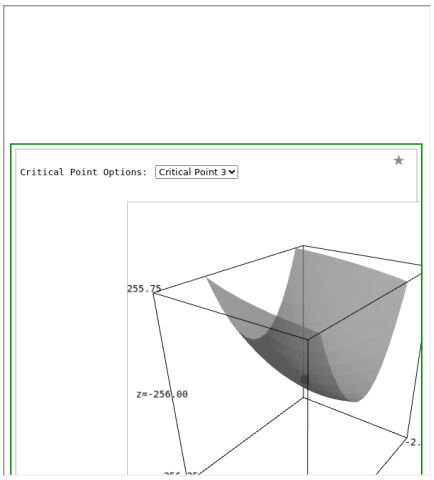
Solving the second equation for y, we get $y = 8x$, which we can substitute into the first equation: $16x^3 - 16(8x) = 0$. Factoring this equation, we have

$$16x(x - \sqrt{8})(x + \sqrt{8}) = 0,$$

which means our critical points must have $x = 0$, $x = \sqrt{8}$, or $x = -\sqrt{8}$. We can substitute each of these values into $y = 8x$ to obtain the following critical points:

1. $(x, y) = (0, 0)$
2. $(x, y) = (\sqrt{8}, 8\sqrt{8})$
3. $(x, y) = (-\sqrt{8}, -8\sqrt{8})$

For each point above, you can use Figure 11.10.6 to show a plot of $z = f(x, y)$ for a small region around the critical point by selecting from the selector. Pay attention to how the coordinate values on the axes change in each case!



Standalone
In Context
Embed

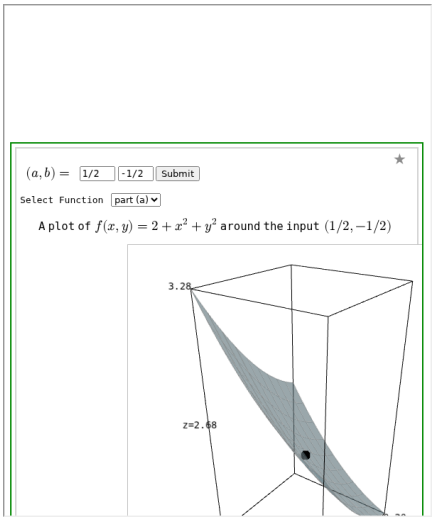
Figure 11.10.6 Plots of $z = f(x, y)$ around each of the critical points

You can see from the plots around each of the critical points in Figure 11.10.6, that the first critical point $(x, y) = (0, 0)$ will *not* be a local maximum or local minimum, but both of $(x, y) = (\sqrt{8}, 8\sqrt{8})$ and $(x, y) = (-\sqrt{8}, -8\sqrt{8})$ will be local minima. ▲

For continuous functions of one variable, you may recall that between two local minima, there must be a local maximum. The previous example shows how for functions of more than one variable, there does not need to be a local maximum between two local minima. We encourage you to think carefully about how possibilities can be different than for one-variable functions as we move forward.

Activity 11.10.2 For each of the functions below, do the following:

1. Find all critical points for the function.
2. Use Figure 11.10.7 to plot the surface near each critical point. Enter the coordinates of your critical point as (a, b) and the function you are using for $f(x, y)$.
3. Use your plots of the surfaces $z = f(x, y)$ near your critical point to decide whether each critical point is a local maximum, a local minimum, or not an extreme point.



Standalone
In Context
Embed

Figure 11.10.7 Plots of $z = f(x, y)$ around each of the critical points

(a) $f(x, y) = 2 + x^2 + y^2$

(b) $f(x, y) = 2 + x^2 - y^2$

(c) $f(x, y) = 2x - x^2 - \frac{1}{4}y^2$

(d) $f(x, y) = |x| + |y|$

(e) $f(x, y) = 2xy - 4x + 2y - 3$

11.10.3 Classifying Critical Points: The Second Derivative Test

While the extreme values of a continuous function f always occur at critical numbers, it is important to note that not every critical point leads to an extremum. Recall, for instance, $f(x) = x^3$ from single variable calculus. Because $f'(0) = 0$, we know that $x_0 = 0$ is a critical number. However, $x_0 = 0$ is not the location of a local maximum or a local minimum of f. Geometrically, this corresponds to the tangent line $x_0 = 0$ being horizontal as seen in Figure 11.10.8 but at a critical number where $f(x_0)$ is neither greater than or equal to or less than or equal to all values $f(x)$ with x near x_0.

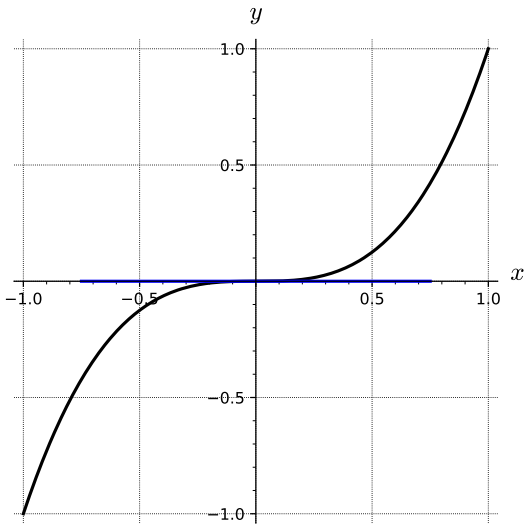
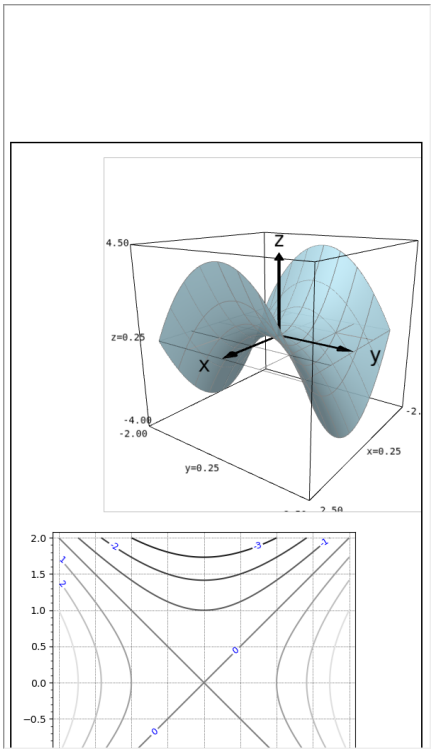


Figure 11.10.8 A plot of $y = x^3$ with the horizontal tangent line at $x_0 = 0$

As we have seen a few times already, a similar situation can arise in a multivariable setting. Let's look at the function $f(x, y) = x^2 - y^2$ whose graph and contour plot are shown in Figure 11.10.9. Because $\nabla f = \langle 2x, -2y \rangle$, we see that the origin $(x_0, y_0) = (0, 0)$ is a critical point. However, this critical point is neither a local maximum nor a local minimum. We can explain why this happens by noticing that the origin is a local *minimum* on the trace defined by $y = 0$, while the origin is a local *maximum* on the trace defined by $x = 0$. We call such a critical point a **saddle point** due to the shape of the graph near the critical point.



Standalone
In Context
Embed

Figure 11.10.9 The surface plot and contour plot of $f(x, y) = x^2 - y^2$

As in single-variable calculus, we would like to have an algebraic test to help us identify whether a critical point is a local maximum, local minimum, or neither. Recall that the **Second Derivative Test** for single-variable functions states that if x_0 is a critical number of a function f for which $f'(x_0) = 0$ and $f''(x_0)$ exists, then we can conclude that

- if $f''(x_0) < 0$, then x_0 is a local maximum,
- if $f''(x_0) > 0$, then x_0 is a local minimum, and
- if $f''(x_0) = 0$, then the test is inconclusive.

The first two cases are shown geometrically in Figure 11.10.10. The inconclusive case does not necessarily mean that you have critical number like the one shown in Figure 11.10.8, but that you need more information to determine the behavior of f at the critical number.

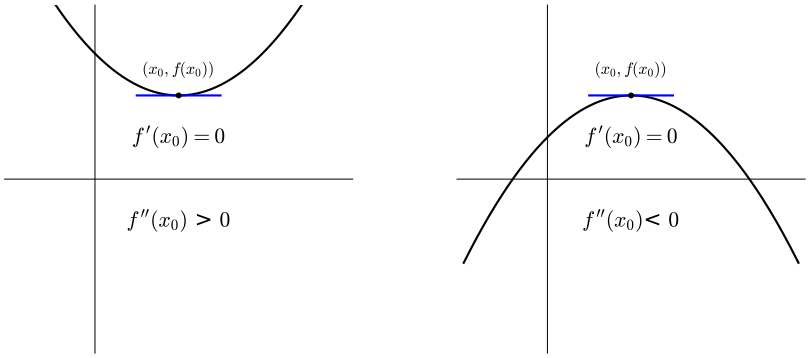
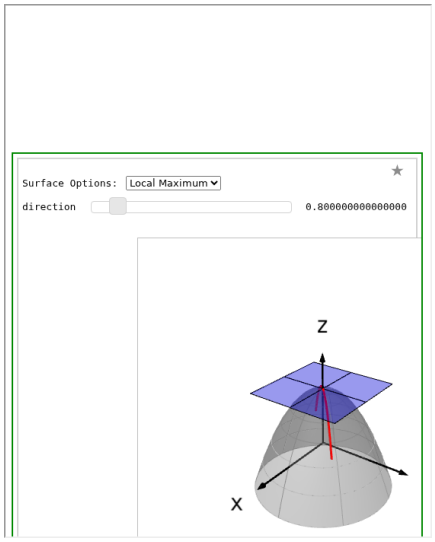


Figure 11.10.10 Conclusive outcomes of the single-variable Second Derivative Test

At a critical point for a function of two variables, we must look at the concavity in all directions around our point. Use the dropdown box in Figure 11.10.11 to select three different types of critical point to examine graphically. Use the slider to change the direction in which the red trace is plotted. Notice that no matter what kind of critical point behavior is selected, the blue tangent plane is horizontal and the red trace curve is tangent to the blue plane.



Standalone
In Context
Embed

Figure 11.10.11 Possible critical points and concavity

- When looking at the local maximum critical point in Figure 11.10.11, use the slider to look at all directions in which you can move away from the critical point. In *every* direction, the trace curves are concave down at the local maximum.
- When looking at the local minimum critical point in Figure 11.10.11, use the slider to look at all directions in which you can move away from the critical point. In *every* direction, the trace curves are concave up at the local minimum.
- When looking at the saddle point, however, changing the direction of movement away from the critical point leads to some directions where the trace curve is concave up and some directions where the trace curve is concave down.

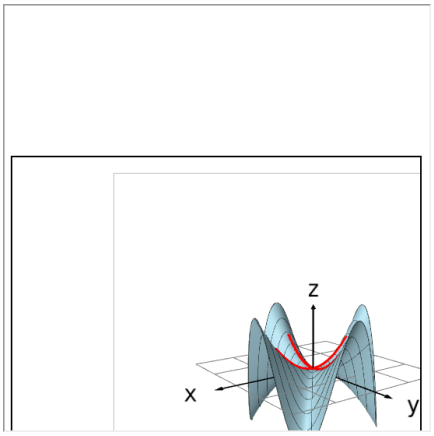
The converse of these patterns should also make sense to you geometrically:

- If every direction is concave up for a critical point, then the critical point is a local minimum.
- If every direction is concave down for a critical point, then the critical point is a local maximum.
- If the concavity changes from up to down in different directions, then the critical point is a saddle.

In order to classify our critical points, we will need to figure out if the behavior around a critical point is always concave up, always concave down, or if the concavity depends on the direction you look.

We begin with the convenient directions in which to look at concavity: parallel to the x- and y-axes. Recall from Section 11.5 that f_{xx} and f_{yy} measure the concavity of the traces in the x and y directions, respectively. If $f_{xx} > 0$ at a critical point, then in the direction parallel to the x-axis, the trace curve is concave up. If $f_{xx} < 0$ at a critical point, then in the direction parallel to the x-axis, the trace curve is concave down. Similar statements hold for f_{yy}.

Suppose that $f_{xx} < 0$ and $f_{yy} > 0$ at a critical point (a, b). Because the concavity is different for two different direction, the critical point (a, b) *must* be a saddle point. Unfortunately, if $f_{xx} > 0$ and $f_{yy} > 0$ at a critical point (a, b), then we do not necessarily have a local minimum, as shown in Figure 11.10.12.



Standalone
In Context
Embed

Figure 11.10.12 A surface plot where $f_{xx} > 0$ and $f_{yy} > 0$ at the origin.

As the plot in Figure 11.10.12 shows, we must look at the concavity in more than the directions corresponding to the coordinate axes. Recall from

Activities 11.5.3 and Activity 11.5.4 that the mixed second partial derivative f_{xy} measures the "twist" of the graph as y varies along a particular trace where x is held constant. In the same way, f_{yx} measures how the graph twists as x varies along a trace with y fixed. For most functions of interest, remember that Clairaut's theorem tells us that $f_{xy} = f_{yx}$. Hence, the amount of twisting is the same in both directions. These mixed second partial derivatives are used to see if there is enough twist to change the concavity in the directions between the coordinate axes.

The **discriminant** of a function of two variables is

$$D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.$$

If the discriminant is negative, then the concavity will change depending on direction. On the other hand, if the discriminant is positive, then all directions have the same concavity. To illustrate this, we consider a few cases:

- If $f_{xx}(x_0, y_0)$ and $f_{yy}(x_0, y_0)$ have different signs or at least one unmixed second partial is zero, then the discriminant is negative because the second term in D is always negative. This is the case of the saddle shown above: one coordinate direction is concave up and one is concave down. In this situation, we do not have an extreme point.
- If $f_{xx}(x_0, y_0)$ and $f_{yy}(x_0, y_0)$ are both positive or both negative, then the first term is positive. If our coordinate directions have the same concavity, then we will need to see if there is enough twist in the surface to change the concavity for directions between the coordinate axes. In other words, if $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) < f_{xy}(x_0, y_0)^2$, there is enough twist to have some direction to have different concavity from the coordinate directions.
- If $f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) > f_{xy}(x_0, y_0)^2$, then there is not enough twist between the coordinate directions and every direction has the same concavity.

These conceptual arguments are summarized below.

The Second Partial Derivatives Test.

Suppose (x_0, y_0) is a critical point of a function f for which $f_x(x_0, y_0) = 0$, $f_y(x_0, y_0) = 0$, and all four second-order partial derivatives exist. Let

$$D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.$$

- If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).
- If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).
- If $D < 0$, then f has a saddle point at (x_0, y_0).
- If $D = 0$, then this test is inconclusive about the behavior of f at (x_0, y_0).

The quantity D is called the **discriminant** of the function f at (x_0, y_0).

A key idea of the Second Partial Derivatives Test is that if the discriminant is positive, all directions have the same concavity, while if the discriminant is negative, then the concavity changes depending on the direction. The last case of the Second Partial Derivatives Test is similar to the inconclusive case for the Second Derivative Test from single-variable calculus.

To properly understand the origin of the Second Partial Derivatives Test, we would need to introduce a "second-order directional derivative." If this second-order directional derivative were negative in every direction, for instance, we could guarantee that the critical point is a local maximum. A complete justification of the Second Partial Derivatives Test requires key ideas from linear algebra that are beyond the scope of this book.

In Activity 11.10.3, we apply the test to more complicated examples.

Activity 11.10.3 For each of the functions given below:

1. Find the critical points of the function
2. Compute the discriminant D and evaluate D for each critical point.
3. Use the Second Partial Derivatives Test to classify the critical points

(a) $f(x, y) = 3x^3 + y^2 - 9x + 4y$

(b) $f(x, y) = xy + \frac{2}{x} + \frac{4}{y}$

(c) $f(x, y) = x^3 + y^3 - 3xy$

As we learned in single-variable calculus, finding extremal values of functions can be particularly useful in applied settings. For instance, we can often use calculus to determine the least expensive way to construct something or to find the most efficient route between two locations. The same possibility holds in settings with two or more variables.

Activity 11.10.4 While the quantity of a product demanded by consumers is often a function of the price of the product, the demand for a product may also depend on the price of other products. For instance, the demand for blue jeans at a clothing store may be affected not only by the price of the jeans themselves, but also by the price of khakis.

Suppose we have two goods whose respective prices are p_1 and p_2. The demand for these goods, q_1 and q_2, depend on the prices by the following relations:

$$q_1 = 150 - 2p_1 - p_2 \tag{11.10.1}$$
$$q_2 = 200 - p_1 - 3p_2 \tag{11.10.2}$$

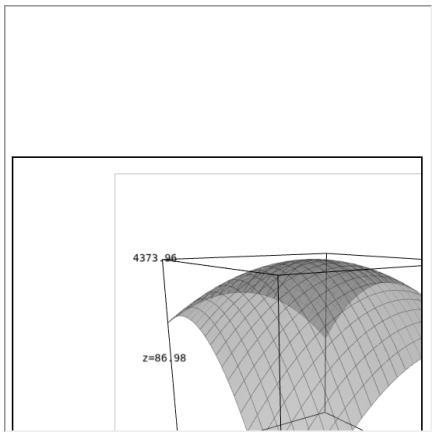
The seller wants to set the prices p_1 and p_2 in order to maximize revenue. We will assume that the seller meets the full demand for each product. Thus, if we let R be the revenue obtained by selling q_1 items of the first good at price p_1 per item and q_2 items of the second good at price p_2 per item, we have

$$R = p_1 q_1 + p_2 q_2.$$

We can then write the revenue as a function of just the two variables p_1 and p_2 by using Equations (11.10.1) and (11.10.2), giving us

$$\begin{aligned} R(p_1, p_2) &= p_1(150 - 2p_1 - p_2) + p_2(200 - p_1 - 3p_2) \\ &= 150p_1 + 200p_2 - 2p_1p_2 - 2p_1^2 - 3p_2^2. \end{aligned}$$

A graph of R as a function of p_1 and p_2 is shown in below.



Standalone
In Context
Embed

(a) Find all critical points of the revenue function R.

Hint. You should obtain a system of two equations in two unknowns which can be solved by elimination or substitution.

(b) Apply the Second Partial Derivatives Test to determine the type of any critical point(s).

(c) Where should the seller set the prices p_1 and p_2 to maximize the revenue?

11.10.4 Optimization on a Restricted Domain

The Second Partial Derivatives Test helps classify critical points for a function of two variables, but it does not tell us if the function has an *absolute* maximum or minimum at each such point. For single-variable functions, the Extreme Value Theorem says that a continuous function on a closed interval $[a, b]$ must have both an absolute maximum and minimum on that interval, and that these absolute extremes must occur at either an endpoint or at a critical number inside the interval. Thus, finding the absolute maximum and absolute minimum of a single-variable function follows a straightforward list of steps: determine the critical points in the interval, evaluate the function at these critical points and at the endpoints of the interval, and identify the largest and smallest output values as the absolute maximum and absolute minimum respectively.

A similar approach works for functions of two variables, once we have determined the appropriate replacement for the idea of the endpoints of an interval. For functions of two variables, regions in the xy-plane that are both closed and bounded play the role that closed intervals did for functions of a single variable. A closed region is a region that contains its boundary. The unit disk $x^2 + y^2 \leq 1$ is closed, for example, while its interior $x^2 + y^2 < 1$ is not. A bounded region is one that does not stretch to infinity in any direction. Just as for functions of a single variable, continuous functions of several variables that are defined on closed, bounded regions must have absolute maxima and minima in those regions. The endpoints of $[a, b]$ are the boundary of the interval, and the boundary of a region in the xy-plane will provide the additional locations beyond critical points at which we must check for absolute extremes.

The Extreme Value Theorem.

If $f = f(x, y)$ is a continuous function on a closed and bounded region R in the xy-plane, then f has an absolute maximum and an absolute minimum in R.

To find the highest point in a state, you would need to look up the elevations for the tops of the hills/mountains and compare those values to the elevation along the state's borders. Note that the elevation along a border of a state is a one-dimensional maximization problem, which means that tools from single-variable calculus can be applied.

The absolute extremes must occur at either a critical point in the interior of R or at a boundary point of R. We therefore must test both possibilities, as we demonstrate in the following example.

Example 11.10.13 Consider a heated circular plate of radius 1 meter with location on the plate described relative to the center. The domain $R = \{(x, y) : x^2 + y^2 \leq 1\}$ of our temperature function is a closed and bounded region, as shown in Figure 11.10.14.

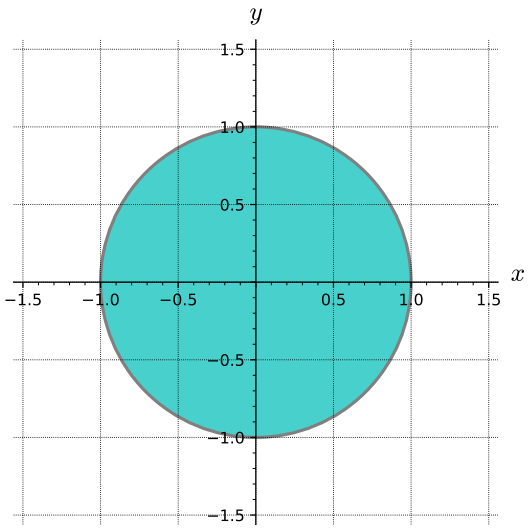


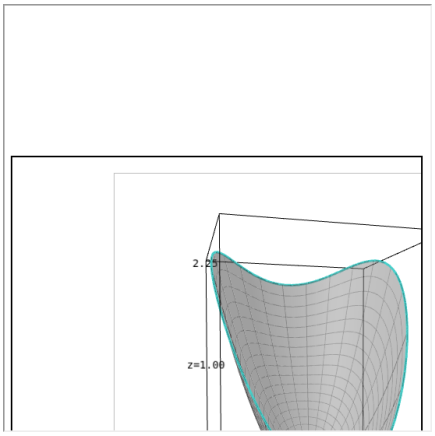
Figure 11.10.14 A plot of the circular plate as the domain of our temperature function

Suppose the temperature T at each point on the circular plate $x^2 + y^2 \leq 1$ is given by

$$T(x, y) = 2x^2 + y^2 - y.$$

We would like to find the absolute maximum and absolute minimum temperature of the circular plate.

Because the domain of T is closed and bounded, the Extreme Value Theorem guarantees that T has an absolute maximum and minimum on the plate. The graph of T over its domain R is shown in Figure 11.10.15. We will find the hottest and coldest points on the plate.



Standalone
In Context
Embed

Figure 11.10.15 A plot of the temperature across the circular metal plate

If the absolute maximum or absolute minimum occurs inside the disk, then the absolute extreme point will be at a critical point. We begin by looking for critical points inside the disk. To do this, notice that critical points are given by the conditions $T_x = 4x = 0$ and $T_y = 2y - 1 = 0$. This means that there is one critical point of the function at the point $(x_0, y_0) = (0, 1/2)$, which does lie inside the disk.

We now find the hottest and coldest points on the boundary of the disk, which is the circle of radius 1. To turn this into a single-variable calculus problem, we use the familiar parameterization of the unit circle as

$$x(t) = \cos(t), \ y(t) = \sin(t),$$

where $0 \leq t \leq 2\pi$. The temperature at a point on the circle is then described by

$$T(x(t), y(t)) = 2\cos^2(t) + \sin^2(t) - \sin(t).$$

To find the hottest and coldest points on the boundary, we look for the critical points of this single-variable function on the interval $0 \leq t \leq 2\pi$. We have

$$\begin{aligned} \frac{dT}{dt} &= -4\cos(t)\sin(t) + 2\cos(t)\sin(t) - \cos(t) \\ &= -2\cos(t)\sin(t) - \cos(t) = \cos(t)(-2\sin(t) - 1) \end{aligned}$$

Therefore, we need to identify the points where $\cos(t)(-2\sin(t)-1) = 0$. These happen when $\cos(t) = 0$ or $\sin(t) = -1/2$ on the interval $0 \leq t \leq 2\pi$. This occurs when $t = \pi/2$, $3\pi/2$, $7\pi/6$, and $11\pi/6$. Since $x(t) = \cos(t)$ and $y(t) = \sin(t)$, the corresponding points are

- $(x, y) = (0, 1)$ when $t = \frac{\pi}{2}$
- $(x, y) = (0, -1)$ when $t = \frac{3\pi}{2}$
- $(x, y) = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ when $t = \frac{11\pi}{6}$
- $(x, y) = \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ when $t = \frac{7\pi}{6}$

These are the critical points of T on the boundary which means that this collection of points includes the hottest and coldest points on the boundary.

We now have a list of candidates for the hottest and coldest points: the critical point in the interior of the disk and the critical points on the boundary. We find the hottest and coldest points by evaluating the temperature at each of these points, and find that

- $T\left(0, \frac{1}{2}\right) = -\frac{1}{4}$
- $T\left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$
- $T(0,1) = 0$
- $T\left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$
- $T(0,-1) = 2$

Hence, the maximum temperature on the disk $x^2 + y^2 \leq 1$ is $\frac{9}{4}$, which occurs at the two points $\left(\pm\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ on the boundary, and the minimum value of T on the disk is $-\frac{1}{4}$ which occurs at the critical point $\left(0, \frac{1}{2}\right)$ in the interior of R. ▲

From this example, we see that we use the following procedure for determining the absolute maximum and absolute minimum of a function on a closed and bounded domain:

Step 1 Find all critical points of the function in the interior of the domain.

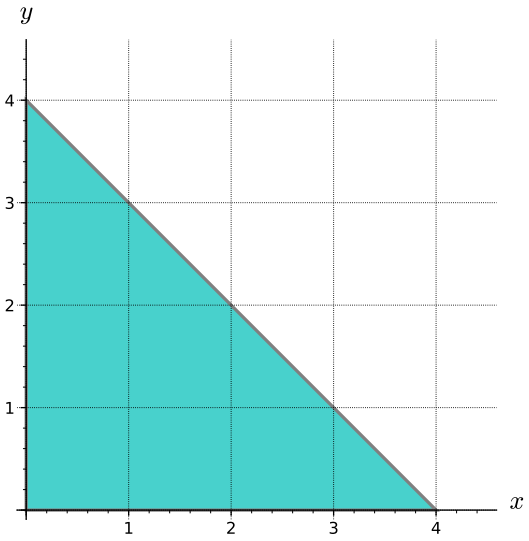
Step 2 Find all the critical points of the function on the boundary of the domain.

- This involves one or more single-variable optimization problems
- There may be endpoints on portions of the boundary that need to be considered.

Step 3 Evaluate the function at each of the points found in Steps 1 and 2.

Step 4 The maximum value of the function is the largest value obtained in Step 3, and the minimum value of the function is the smallest value obtained in Step 3.

Activity 11.10.5 Let $f(x,y) = x^2 - 3y^2 - 4x + 6y$ with triangular domain R whose vertices are at $(0,0)$, $(4,0)$, and $(0,4)$. The domain R is illustrated below. In this activity, you will carry out the steps to find the absolute maximum and absolute minimum of f on R.



(a) Find all of the critical points of f in R.

(b) Parameterize the edge of R that is on the x-axis and find the critical points of f on that edge.

Hint. You may need to consider endpoints.

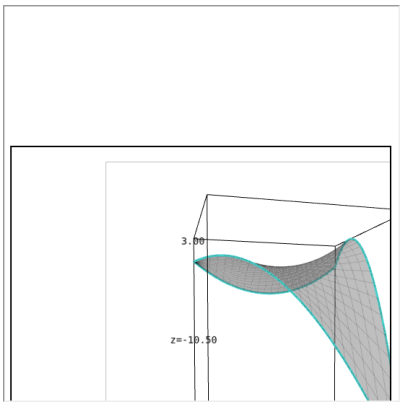
(c) Parameterize the edge of R that is on the y-axis and find the critical points of f on that edge.

Hint. You may need to consider endpoints.

(d) Parameterize the diagonal edge of R and find the critical points of f on that edge.

Hint. You may need to consider endpoints.

(e) Find the absolute maximum and absolute minimum values of f on R. Write a couple of sentences to describe how the surface plot of f shown below illustrates your results.



Standalone
In Context
Embed

Summary

- To find the extrema of a function $f = f(x, y)$, we first find the critical points, which are points where one of the partials of f fails to exist or where both $f_x = 0$ and $f_y = 0$.
- The Second Partial Derivatives Test helps determine whether a critical point is a local maximum, local minimum, or saddle point.
- If f is defined on a closed and bounded domain, we find the absolute maxima and absolute minima by finding the critical points in the interior of the domain, finding the critical points on the boundary, and testing the value of f at both sets of critical points.

11.10.5 Exercises

1. Find the critical points of the function $f(x, y) = x^2+y^2+10x-4y+4$.***List your answers as points in the form*** (a, b)***.***

Answer (separate by commas): _____

2. Find the critical points of the function $f(x, y) = x^3 - 3x^2 - 2y^3 + 9y^2 + 2$.***List your answers as points in the form*** (a, b)***.***

Answer (separate by commas): _____

3. Locate all the critical points of the function $f(x, y) = 2x - x^2 - xy^2$.
(If there are several points, separate them by commas.)

4. The function
$$k(x, y) = e^{-y^2} \cos(2x)$$
has a critical point at $(0, 0)$.
What is the value of D at this critical point? $D =$ _____
What type of critical point is it? (☐ maximum ☐ minimum ☐ saddle point ☐ point with unknown behavior)

5. Find the maximum and minimum values of $f(x, y) = xy$ on the ellipse $5x^2 + y^2 = 9$.
maximum value = _____
minimum value = _____

6. Consider the three points $(3, 3)$, $(5, 5)$, and $(7, 6)$.
(a) Supposed that at $(3, 3)$, we know that $f_x = f_y = 0$ and $f_{xx} = 0$, $f_{yy} = 0$, and $f_{xy} > 0$. What can we conclude about the behavior of this function near the point $(3, 3)$? (☐ (3,3) is a local maximum ☐ (3,3) is a local minimum ☐ (3,3) is a saddle point ☐ (3,3) is a none of these)
(b) Supposed that at $(5, 5)$, we know that $f_x = f_y = 0$ and $f_{xx} = 0$, $f_{yy} > 0$, and $f_{xy} < 0$. What can we conclude about the behavior of this function near the point $(5, 5)$? (☐ (5,5) is a local maximum ☐ (5,5) is a local minimum ☐ (5,5) is a saddle point ☐ (5,5) is a none of these)
(c) Supposed that at $(7, 6)$, we know that $f_x = f_y = 0$ and $f_{xx} = 0$, $f_{yy} < 0$, and $f_{xy} > 0$. What can we conclude about the behavior of this function near the point $(7, 6)$? (☐ (7,6) is a local maximum ☐ (7,6) is a local minimum ☐ (7,6) is a saddle point ☐ (7,6) is a none of these)
Using this information, on a separate sheet of paper sketch a possible contour diagram for f.

7. Consider the function $f(x, y) = (12x - x^2)(6y - y^2)$.
Find the first- and second-order partial derivatives of f.

- $f_x =$_____
- $f_y =$_____
- $f_{xx} =$_____
- $f_{xy} =$_____
- $f_{yy} =$_____

Find and classify all critical points (x, y) of the function. If there are more blanks than critical points, leave the remaining entries blank.
There are several critical points to be listed. List them lexicographically, that is in ascending order by x-coordinates, and for equal x-coordinates in ascending order by y-coordinates (e.g., (1,1),(1,10), (2, -1), (2, 3) is a correct order)
In lexicographic order:

- The critical point with the smallest x and y coordinates is _____. Classification: (☐ local minimum ☐ local maximum ☐ saddle point ☐ can not be determined)
- The next critical point is _____. Classification: (☐ local minimum ☐ local maximum ☐ saddle point ☐ can not be determined)

- The next critical point is _____. Classification: (☐ local minimum ☐ local maximum ☐ saddle point ☐ can not be determined)
- The next critical point is _____. Classification: (☐ local minimum ☐ local maximum ☐ saddle point ☐ can not be determined)
- The next critical point is _____. Classification: (☐ local minimum ☐ local maximum ☐ saddle point ☐ can not be determined)

8. Suppose $f(x, y) = xy - ax - by$.

(A) How many local minimum points does f have in $\mathbf{R}^2$? (The answer is an integer).

(B) How many local maximum points does f have in $\mathbf{R}^2$?

(C) How many saddle points does f have in $\mathbf{R}^2$?

9. ***(a)*** Find the critical point of $f(x, y) = \left(4x^2 - y\right)\left(x^2 + 3y\right)$.

$(x, y) =$ _____

(b) Find the discriminant, $D(x, y)$ for the function.

$D(x, y) =$ _____

what is the value of the discriminant at the critical point?

$D =$ _____

What does this tell us about the critical point?

(☐ The critical point is a local maximum ☐ The critical point is a local minimum ☐ The critical point is a saddle point ☐ The discriminant doesn't tell us anything)

(c) Sketch contours near the critical point to determine, or confirm, whether it is a local maximum, a local minimum, a saddle point, or none of these.

(☐ The critical point is a local maximum ☐ The critical point is a local minimum ☐ The critical point is a saddle point ☐ The discriminant doesn't tell us anything)

10. Suppose $g(x, y) = x^3 - 12xy + 2y^2$. Answer the following.

In each case ***give all three coordinates of each point:*** (a, b, c)***. If there are two or more points separate them by commas. If there are no points answer*** None.

1. Find the local maxima of g.

Answer: _____

2. Find the local minima of g.

Answer: _____

3. Find the saddle points of g.

Answer: _____

11. Each of the following functions has at most one critical point. Graph a few level curves and a few gradients and, on this basis alone, decide whether the critical point is a local maximum, a local minimum, a saddle point, or that there is no critical point.

For $f(x, y) = e^{-x^2-4y^2}$, type of critical point: (☐ Local Minimum ☐ No Critical Point ☐ Local Maximum ☐ Saddle Point)

For $f(x, y) = e^{x^2-4y^2}$, type of critical point: (☐ Saddle Point ☐ Local Minimum ☐ Local Maximum ☐ No Critical Point)

For $f(x, y) = x^2 + 4y^2 + 4$, type of critical point: (☐ Saddle Point ☐ No Critical Point ☐ Local Maximum ☐ Local Minimum)

For $f(x, y) = x^2 + 4y + 4$, type of critical point: (☐ Saddle Point

☐ Local Minimum ☐ Local Maximum ☐ No Critical Point)

12. Suppose $f(x, y) = (x - y)(1 - xy)$. Answer the following. ***Each answer should be a list of points*** (a, b, c) ***separated by commas, or, if there are no points, the answer should be NONE.***

1. Find the local maxima of f.
Answer: _____

2. Find the local minima of f.
Answer: _____

3. Find the saddle points of f.
Answer: _____

13. Find the absolute minimum and absolute maximum of

$$f(x, y) = 7 - 3x + 7y$$

on the closed triangular region with vertices $(0, 0), (7, 0)$ and $(7, 12)$.

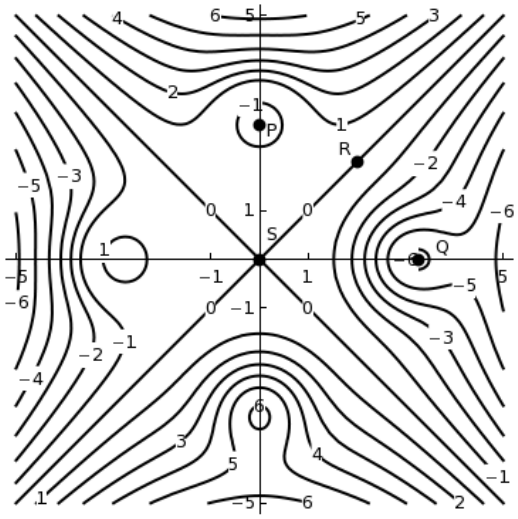
List the minimum/maximum values as well as the point(s) at which they occur. If a min or max occurs at multiple points separate the points with commas.

Minimum value: _____
Occurs at _____
Maximum value: _____
Occurs at _____

14. Find A and B so that $f(x, y) = x^2 + Ay + y^2 + B$ has a local minimum at the point $(0, 7)$, with z-coordinate 15.

$A =$ _____
$B =$ _____

15. The contours of a function f are shown in the figure below.



For each of the points shown, indicate whether you think it is a local maximum, local minimum, saddle point, or none of these.

(a) Point P is (☐ a local maximum ☐ a local minimum ☐ a saddle point ☐ none of these)

(b) Point Q is (☐ a local maximum ☐ a local minimum ☐ a saddle point ☐ none of these)

(c) Point R is (☐ a local maximum ☐ a local minimum ☐ a saddle point ☐ none of these)

(d) Point S is (☐ a local maximum ☐ a local minimum ☐ a saddle

point ☐ none of these)

16. Find the critical points for the function

$$f(x,y) = 3x^2 - 6xy + 4y^2 - 2y$$

and classify each as a local maximum, local minimum, saddle point, or none of these.

critical points: _____

(give your points as a comma separated list of (x,y) coordinates.)

classifications: _____

*(give your answers in a comma separated list, specifying **maximum**, **minimum**, **saddle point**, or **none** for each, in the same order as you entered your critical points)*

17. Find the critical points for the function

$$f(x,y) = x^3 + y^3 - 12x^2 - 27y - 1$$

and classify each as a local maximum, local minimum, saddle point, or none of these.

critical points: _____

(give your points as a comma separated list of (x,y) coordinates.)

classifications: _____

*(give your answers in a comma separated list, specifying **maximum**, **minimum**, **saddle point**, or **none** for each, in the same order as you entered your critical points)*

18. Find the critical points for the function

$$f(x,y) = 96xy - 0.75(x+y)^4$$

and classify each as a local maximum, local minimum, saddle point, or none of these.

critical points: _____

(give your points as a comma separated list of (x,y) coordinates.)

classifications: _____

*(give your answers in a comma separated list, specifying **maximum**, **minimum**, **saddle point**, or **none** for each, in the same order as you entered your critical points)*

19. Let

$$f(x,y) = e^{-x^2} + 1y^4.$$

Find critical points and classify them as local maxima, local minima, saddle points, or none of these.

critical points: _____

(give your points as a comma separated list of (x,y) coordinates.)

classifications: _____

*(give your answers in a comma separated list, specifying **maximum**, **minimum**, **saddle point**, or **none** for each, in the same order as you entered your critical points)*

20. Let

$$f(x,y) = 1 + 3x^2 - \cos(4y)\,.$$

Find all critical points and classify them as local maxima, local minima, saddle points, or none of these.

critical points: _____

(give your points as a comma separated list of (x,y) coordinates. if

your answer includes points that occur at a sequence of values, e.g., at every odd integer, or at any constant multiple of another value, use ***m*** *for any non-zero even integer,* ***n*** *for any non-zero odd integer, and/or* ***k*** *for other arbitrary constants.)*

classifications: _____

(give your answers in a comma separated list, specifying ***maximum****,* ***minimum****,* ***saddle point****, or* ***none*** *for each, in the same order as you entered your critical points)*

21. Let $f(x,y) = 1/x + 2/y + 3x + 4y$ in the region R where $x, y > 0$.

Explain why f must have a global minimum at some point in R (note that R is unbounded---how does this influence your explanation?). Then find the global minimum.

minimum = _____

22. Each of the following functions has at most one critical point. Graph a few level curves and a few gradients and, on this basis alone, decide whether the critical point is a local maximum, a local minimum, a saddle point, or that there is no critical point.

For $f(x,y) = e^{-2x^2-y^2}$, type of critical point: (☐ Local Maximum ☐ No Critical Point ☐ Local Minimum ☐ Saddle Point)

For $f(x,y) = e^{2x^2-y^2}$, type of critical point: (☐ No Critical Point ☐ Local Maximum ☐ Local Minimum ☐ Saddle Point)

For $f(x,y) = 2x^2 + y^2 + 3$, type of critical point: (☐ Saddle Point ☐ Local Minimum ☐ Local Maximum ☐ No Critical Point)

For $f(x,y) = 2x^2 + y + 3$, type of critical point: (☐ Saddle Point ☐ Local Minimum ☐ Local Maximum ☐ No Critical Point)

23. Find three positive real numbers whose sum is 38 and whose product is a maximum.

Enter the three numbers separated by commas: _____

24. A closed rectangular box has volume $40\,\text{cm}^3$. What are the lengths of the edges giving the minimum surface area?

lengths = _____

(Give the three lengths as a comma separated list.)

25. An open rectangular box has volume 44 cm^3. What are the lengths of the edges giving the minimum surface area?

lengths = _____

(Give the three lengths as a comma separated list.)

26. What is the shortest distance from the surface $xy + 12x + z^2 = 137$ to the origin?

distance = _____

27. Find the volume of the largest rectangular box with edges parallel to the axes that can be inscribed in the ellipsoid

$$\frac{x^2}{64} + \frac{y^2}{9} + \frac{z^2}{36} = 1$$

Hint: By symmetry, you can restrict your attention to the first octant (where $x, y, z \geq 0$), and assume your volume has the form $V = 8xyz$. Then arguing by symmetry, you need only look for points which achieve the maximum which lie in the first octant.

Maximum volume: _____

28. Design a rectangular milk carton box of width w, length l, and height h which holds 544 cm^3 of milk. The sides of the box cost 3 cents/cm^2 and the top and bottom cost 4 cents/cm^2. Find the dimensions of the box that minimize the total cost of materials used.

width = _____, length = _____, height = _____

(Include units in your answers.) ( )

29. Find the absolute maximum and absolute minimum of the function $f(x, y) = 2x^3 + y^4$ on the region $(x, y) \mid x^2 + y^2 \leq 1$.

If the absolute max or min occurs at multiple points list them all, separated by commas.

Absolute minimum value: _____

attained at _____.

Absolute maximum value: _____

attained at _____.

30. Does the function

$$f(x, y) = \frac{x^2}{2} + 3y^3 + 3y^2 - 7x$$

have a global maximum and global minimum? If it does, identify the value of the maximum and minimum. If it does not, be sure that you are able to explain why.

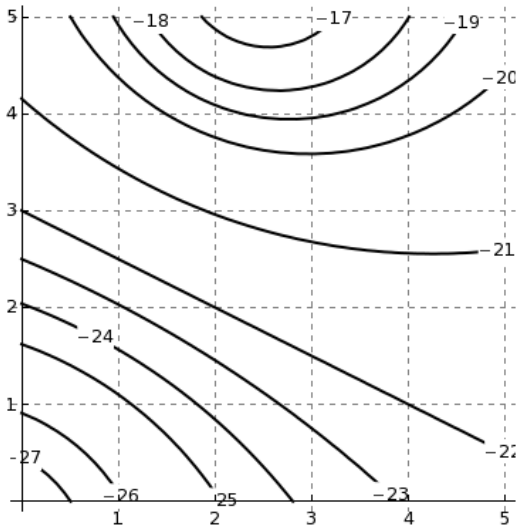
Global maximum? _____

*(Enter the value of the global maximum, or **none** if there is no global maximum.)*

Global minimum? _____

*(Enter the value of the global minimum, or **none** if there is no global minimum.)*

31. A contour diagram for a function $f(x, y)$ is shown below.



Estimate the position and approximate value of the global maximum and global minimum on the region shown.

Global maximum at _____ of _____

Global minimum at _____ of _____

32. Find the parabola of the form $y = ax^2 + b$ which best fits the points

$(-1, 0)$, $(4, 4)$, $(6, 8)$ by minimizing the sum of squares, S, given by

$$S = (a + b)^2 + (16a + b - 4)^2 + (36a + b - 8)^2.$$

$y =$ _____ $x^2 +$ _____

33. A closed rectangular box with faces parallel to the coordinate planes has one bottom corner at the origin and the opposite top corner in the first octant on the plane $4x + 2y + z = 1$. What is the maximum volume of such a box?

volume = _____

34. Respond to each of the following prompts to solve the given optimization problems.

(a) Let $f(x, y) = \sin(x) + \cos(y)$. Determine the absolute maximum and minimum values of f. At what points do these extreme values occur?

(b) For a certain differentiable function F of two variables x and y, its partial derivatives are

$$F_x(x, y) = x^2 - y - 4 \text{ and } F_y(x, y) = -x + y - 2.$$

Find each of the critical points of F, and classify each as a local maximum, local minimum, or a saddle point.

(c) Determine all critical points of $T(x, y) = 48 + 3xy - x^2y - xy^2$ and classify each as a local maximum, local minimum, or saddle point.

(d) Find and classify all critical points of $g(x, y) = \frac{x^2}{2} + 3y^3 + 9y^2 - 3xy + 9y - 9x$

(e) Find and classify all critical points of $z = f(x, y) = ye^{-x^2-2y^2}$.

(f) Determine the absolute maximum and absolute minimum of $f(x, y) = 2 + 2x + 2y - x^2 - y^2$ on the triangular plate in the first quadrant bounded by the lines $x = 0$, $y = 0$, and $y = 9 - x$.

(g) Determine the absolute maximum and absolute minimum of $f(x, y) = 2 + 2x + 2y - x^2 - y^2$ over the closed disk of points (x, y) such that $(x - 1)^2 + (y - 1)^2 \leq 1$.

(h) Find the point on the plane $z = 6 - 3x - 2y$ that lies closest to the origin.

35. If a continuous function f of a single variable has two critical numbers c_1 and c_2 at which f has relative maximum values, then f must have another critical number c_3, because "it is impossible to have two mountains without some sort of valley in between. The other critical point can be a saddle point (a pass between the mountains) or a local minimum (a true valley)." (From *Calculus in Vector Spaces* by Lawrence J. Corwin and Robert H. Szczarb.) Consider the function f defined by $f(x, y) = 4x^2e^y - 2x^4 - e^{4y}$. (From Ira Rosenholz in the Problems Section of the *Mathematics Magazine*, Vol. 60 NO. 1, February 1987.) Show that f has exactly two critical points, and that f has relative maximum values at each of these critical points. Explain how this function f illustrates that it really is possible to have two mountains without some sort of valley in between. Use appropriate technology to draw the surface defined by f to see graphically how

this happens.

36. If a continuous function f of a single variable has exactly one critical number with a relative maximum at that critical point, then the value of f at that critical point is an absolute maximum. In this exercise we see that the same is not always true for functions of two variables. Let $f(x, y) = 3xe^y - x^3 - e^{3y}$ (from "The Only Critical Point in Town" Test by Ira Rosenholz and Lowell Smylie in the *Mathematics Magazine*, VOL 58 NO 3 May 1985.). Show that f has exactly one critical point, has a relative maximum value at that critical point, but that f has no absolute maximum value. Use appropriate technology to draw the surface defined by f to see graphically how this happens.

37. A manufacturer wants to procure rectangular boxes to ship its product. The boxes must contain 20 cubic feet of space. To be durable enough to ensure the safety of the product, the material for the sides of the boxes will cost \$0.10 per square foot, while the material for the top and bottom will cost \$0.25 per square foot. In this activity we will help the manufacturer determine the box of minimal cost.

(a) What quantities are constant in this problem? What are the variables in this problem? Provide appropriate variable labels. What, if any, restrictions are there on the variables?

(b) Using your variables from (a), determine a formula for the total cost C of a box.

(c) Your formula in part (b) might be in terms of three variables. If so, find a relationship between the variables, and then use this relationship to write C as a function of only two independent variables.

(d) Find the dimensions that minimize the cost of a box. Be sure to verify that you have a minimum cost.

38. A rectangular box with length x, width y, and height z is being built. The box is positioned so that one corner is stationed at the origin and the box lies in the first octant where x, y, and z are all positive. There is an added constraint on how the box is constructed: it must fit underneath the plane with equation $x + 2y + 3z = 6$. In fact, we will assume that the corner of the box "opposite" the origin must actually lie on this plane. The basic problem is to find the maximum volume of the box.

(a) Sketch the plane $x + 2y + 3z = 6$, as well as a picture of a potential box. Label everything appropriately.

(b) Explain how you can use the fact that one corner of the box lies on the plane to write the volume of the box as a function of x and y only. Do so, and clearly show the formula you find for $V(x, y)$.

(c) Find all critical points of V. (Note that when finding the critical points, it is essential that you factor first to make the algebra easier.)

(d) Without considering the current applied nature of the function V, classify each critical point you found above as a local maximum, local minimum, or saddle point of V.

(e) Determine the maximum volume of the box, justifying your answer completely with an appropriate discussion of the critical points of the function.

(f) Now suppose that we instead stipulated that, while the vertex of the box opposite the origin still had to lie on the plane, we were only going to permit the sides of the box, x and y, to have values in a specified range (given below). That is, we now want to find the maximum value of V on the closed, bounded region

$$\frac{1}{2} \leq x \leq 1, \quad 1 \leq y \leq 2.$$

Find the maximum volume of the box under this condition, justifying your answer fully.

39. The airlines place restrictions on luggage that can be carried onto planes.

- A carry-on bag can weigh no more than 40 lbs.
- The length plus width plus height of a bag cannot exceed 45 inches.
- The bag must fit in an overhead bin.

Let x, y, and z be the length, width, and height (in inches) of a carry on bag. In this problem we find the dimensions of the bag of largest volume, $V = xyz$, that satisfies the second restriction. Assume that we use all 45 inches to get a maximum volume. (Note that this bag of maximum volume might not satisfy the third restriction.)

(a) Write the volume $V = V(x, y)$ as a function of just the two variables x and y.

(b) Explain why the domain over which V is defined is the triangular region R with vertices (0,0), (45,0), and (0,45).

(c) Find the critical points, if any, of V in the interior of the region R.

(d) Find the maximum value of V on the boundary of the region R, and then determine the dimensions of a bag with maximum volume on the entire region R. (Note that most carry-on bags sold today measure 22 by 14 by 9 inches with a volume of 2772 cubic inches, so that the bags will fit into the overhead bins.)

40.

According to *The Song of Insects* by G.W. Pierce (Harvard College Press, 1948) the sound of striped ground crickets chirping, in number of chirps per second, is related to the temperature. So the number of chirps per second could be a predictor of temperature. The data Pierce collected is shown in Table 11.10.16., where x is the (average) number of chirps per second and y is the temperature in degrees Fahrenheit.

A scatterplot of the data would show that, while the relationship between x and y is not exactly linear, it looks to have a linear pattern. It could be that the relationship is really linear but experimental error causes the data to be slightly inaccurate. Or perhaps the data is not linear, but only approximately linear.

x	y
20.0	88.6
16.0	71.6
19.8	93.3
18.4	84.3
17.1	80.6
15.5	75.2
14.7	69.7
17.1	82.0
15.4	69.4
16.2	83.3
15.0	79.6
17.2	82.6
16.0	80.6
17.0	83.5
14.4	76.3

Table 11.10.16 Crickets chirping.

If we want to use the data to make predications, then we need to fit a curve of some kind to the data. Since the cricket data appears roughly linear, we will fit a linear function f of the form $f(x) = mx+b$ to the data. We will do this in such a way that we minimize the sums of the squares of the distances between the y values of the data and the corresponding y values of the line defined by f. This type of fit is called a *least squares* approximation. If the data is represented by the points (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), then the square of the distance between y_i and $f(x_i)$ is $(f(x_i) - y_i)^2 = (mx_i + b - y_i)^2$. So our goal is to minimize the sum of these squares, of minimize the function S defined by

$$S(m, b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2.$$

(a) Calculate S_m and S_b.

(b) Solve the system $S_m(m, b) = 0$ and $S_b(m, b) = 0$ to show that the critical point satisfies

$$m = \frac{n\left(\sum_{i=1}^n x_i y_i\right) - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

$$b = \frac{\left(\sum_{i=1}^n y_i\right)\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n x_i y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}.$$

Hint. Don't be daunted by these expressions, the system $S_m(m, b) = 0$ and $S_b(m, b) = 0$ is a system of two linear equations in the unknowns m and b. It might be easier to let $r = \sum_{i=1}^n x_i^2$, $s = \sum_{i=1}^n x_i$, $t = \sum_{i=1}^n y_i$, and $u = \sum_{i=1}^n x_i y_i$ and write your equations using these constants.

(c) Use the Second Derivative Test to explain why the critical point gives a local minimum. Can you then explain why the critical point gives an absolute minimum?

(d) Use the formula from part (b) to find the values of m and b that give the line of best fit in the least squares sense to the cricket data. Draw your line on the scatter plot to convince yourself that you have a well-fitting line.

11.11 Constrained Optimization: Lagrange Multipliers

Motivating Questions

- What geometric condition enables us to optimize a function $f = f(x, y)$ subject to a constraint given by $g(x, y) = k$, where k is a constant?
- How can we exploit this geometric condition to find the extreme values of a function subject to a constraint?

11.11.1 Introduction

We previously considered how to find the extreme values of functions on both unrestricted domains and closed, bounded domains. Other types of optimization problems involve maximizing or minimizing a quantity subject to an external constraint. In these cases the extreme values frequently won't occur at the points where the gradient is zero, but rather at other points that satisfy an important geometric condition. These problems are called **constrained optimization** problems and can be solved with the method of Lagrange Multipliers, which we study in this section. This method also provides an alternative approach to looking for maxima or minima on the boundary of a closed, bounded domain and can be useful in cases where parameterizing the boundary is more difficult than the situation we saw in Example 11.10.13.

Preview Activity 11.11.1 A shipping company places a restriction on packages so that the sum of the package's length and girth must be at most 12 feet. (The girth is the perimeter of the smallest end of the package.) Our goal is to find the largest possible volume of a rectangular package with a *square* end that can be sent by this shipping company.

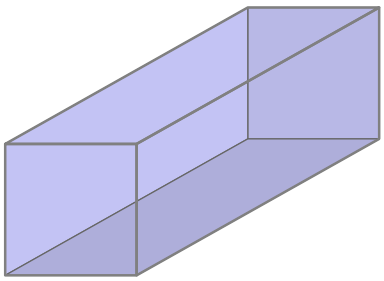


Figure 11.11.1 A shipping box with square end

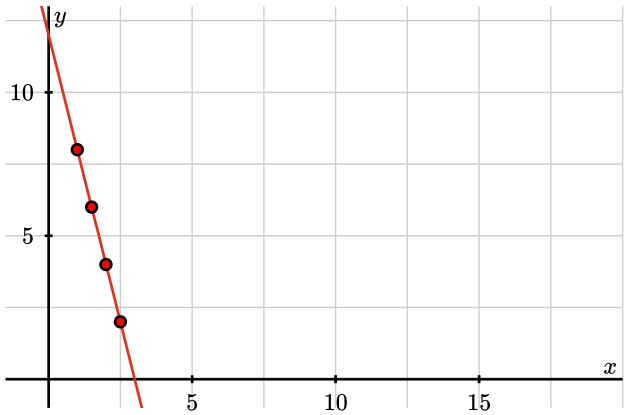
Let x be the length in feet of the side of one square end of the package and y be the length in feet of the package. This allows us to define a function that gives the volume of the package as $f(x, y) = x^2 y$ cubic feet. The girth of the package is $4x$, so we can define $g(x, y) = 4x + y$ to be the sum of the girth and length. The domain on which we wish to maximize f is all points for which $x \geq 0$, $y \geq 0$, and $g(x, y) \leq 12$. A moment of reflection should confirm that any package with $g(x, y) < 12$ cannot maximize the volume, as increasing either x

or y so that $g(x, y)$ remains at most 12 will only lead to a larger volume.

Thus, we say that we wish to maximize $f(x, y) = x^2 y$ subject to the constraint $g(x, y) = 4x + y = 12$. We call f the objective function since it is our goal to find extreme values for this function. We call g our contraint function because we are constrainted (limited) to a specific value of this function.

(a) Compute ∇f and evaluate the gradient at the points $P_1 : (1, 8)$, $P_2 : (1.5, 6)$, $P_3 : (2, 4)$, and $P_4 : (2.5, 2)$.

(b) The plot below shows the graph of $g(x) = 12$ as well as the four points for which you evaluated ∇f in the previous part. Sketch ∇f at each of the four points. You may not have enough space to sketch get the magnitude of each gradient vector correct, but you should accurately depict the direction of each of the four vectors.



(c) Compute ∇g. Describe the geometric relationship between ∇g and the line $g(x, y) = 12$ plotted in the previous part and why they have this relationship.

(d) Recall that the gradient always points in the direction of steepest ascent. However, in this *constrained* optimization setting, the domain that we allowed to consider for f is only those points for which $g(x, y) = 12$. Thus, we may not be able to move in the direction of the gradient while staying within the domain. Look back at the four points at which you sketched ∇f in part (b). For each point, identify which direction you could move *along the line* to increase the value of f. What is the relationship between ∇f and ∇g at a point where you cannot move along the line while increasing the value of f?

Hint. For each of the points in the plot, you may find it helfpul to sketch ∇g with its tail at the point.

11.11.2 Constrained Optimization and Lagrange Multipliers

We will further examine our work in Preview Activity 11.11.1 to motivate and understand the general method for solving constrained optimization problems. In Preview Activity 11.11.1, we considered an optimization problem where there is an external constraint on the variables, namely that the girth plus the length of the package cannot exceed 12 feet.

You saw in your plots that the direction to move in the (x, y)-plane to increase the volume would take you off of the constraint line. For points P_1

and P_2 you saw how moving down on the constraint line would move you at least a little bit in the direction of greatest increase (as measured by ∇f). Similarly, at points P_3 and P_4 you could move up the constraint line to increase the value of the objective function. So somewhere between P_2 and P_3 there has to be a place where moving down the constraint line will no longer increase the output of our objective function.

Ultimately, we observed that the gradient of the **objective function** f and the gradient of the **constraint** g must be parallel when we maximize f subject to the constraint. This is because the only way to increase the value of the objective function at such a point is to move off of the constraint curve, which is not allowed. In Figure 11.11.2, we show the contour $f(x, y) = x^2y = 16$, the constraint $g(x, y) = 4x + y = 12$, and the point $(2, 4)$ at which these two intersect. We have also included the gradients at this point, illustrating that the gradients are parallel.

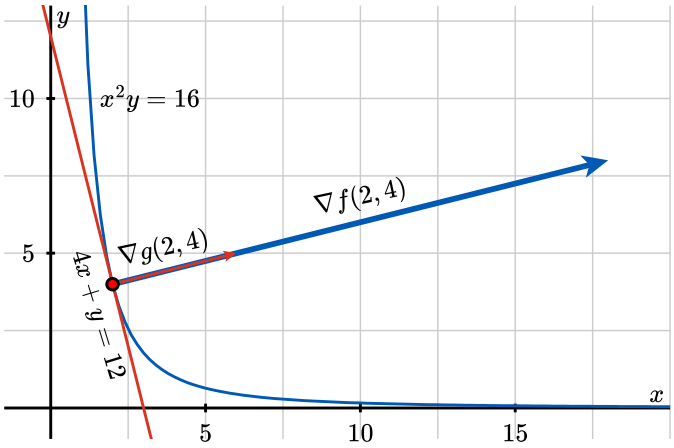


Figure 11.11.2 A contour of f along with the constraint line and the gradients to each at the point where they intersect.

After doing the preview activity and looking at Figure 11.11.2, you may be wondering if there are *other* contours that intersect the constraint line and have a point where the gradient to the objective function and the gradient to the constraint function are parallel. In the next example, we will set up an algebraic approach to finding all points on the constraint where the gradients of the objective and constraint functions are parallel and then find where maximum and minimum values occur for the objective function.

Example 11.11.3 Summarizing the information from the Preview Activity, we define

$$f(x, y) = x^2y$$
$$g(x, y) = 4x + y$$

and wish to maximize f subject to the constraint $g(x, y) = 12$. Computing the gradients of both f and g gives us

$$\nabla f(x, y) = \langle 2xy, x^2 \rangle$$
$$\nabla g(x, y) = \langle 4, 1 \rangle.$$

Now recall that two vectors are parallel exactly when one is a scalar multiple of the other. Algebraically, this happens when one vector is a scalar multiple of the other:

$$\nabla f = \lambda g. \tag{11.11.1}$$

We refer to the scalar λ as a **Lagrange multiplier**. For our objective and constraint functions, this gives rise to the system of equations:

$$2xy = 4\lambda \tag{11.11.2}$$

$$x^2 = \lambda \tag{11.11.3}$$

$$4x + y = 12 \tag{11.11.4}$$

The first two equations come from equating the first and second components of (11.11.1), respectively. The third equation is the constraint equation.

Now we must solve the system of equations to find all points where the gradients of f and g will be parallel. We can use (11.11.3) to give us an expression for λ that can be substituted into (11.11.2). This gives the equation $2xy = 4x^2$, which is equivalent to

$$0 = 4x^2 - 2xy = 2x(2x - y).$$

From this, we can see that either $x = 0$ or $y = 2x$. It is important that we not "cancel" the x from $2xy = 4x^2$, as that would lead to us losing the solution $x = 0$. We can now substitute these solutions into (11.11.4), which gives us $y = 12$ when $x = 0$ and $4x + 2x = 12$ or $x = 2$ when $y = 2x$. Substituting $x = 2$ into (11.11.4) gives us $y = 4$. Thus we need to consider the points $(0, 12)$ and $(2, 4)$.

Conceptually, the point $(0, 12)$ would correspond to a minimum volume of zero because the side length on the square side of our box would be zero. Algebraically, we get $f(0, 12) = 0$. On the other hand, $f(2, 4) = 16$, so we have a maximum value of 16 cubic feet for a rectangular package with a square end subject to the given constraint. You should note that this point occurs between P_2 and P_3, as we previously reasoned. ▲

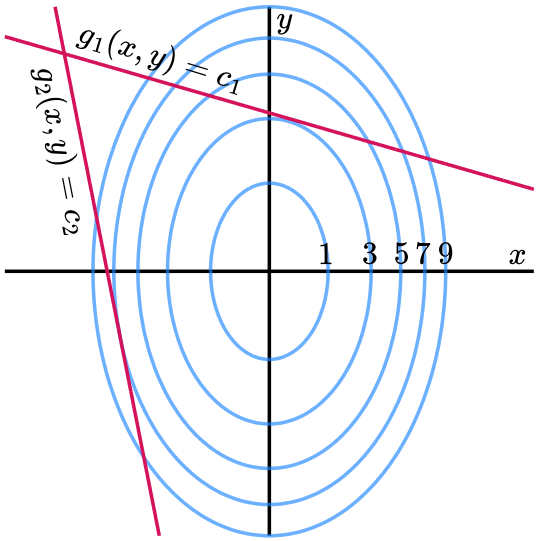
We summarize the process of Lagrange multipliers as follows.

Key Idea 11.11.4 Method of Lagrange Multipliers. *The general technique for optimizing a function $f = f(x, y)$ subject to a constraint $g(x, y) = c$ is to solve the system $\nabla f = \lambda \nabla g$ and $g(x, y) = c$ for x, y, and λ. We then evaluate the function f at each point (x, y) that results from a solution to the system in order to find the optimum values of f subject to the constraint.*

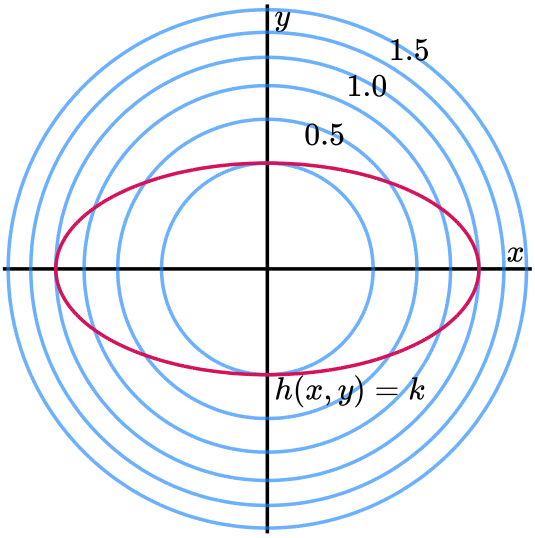
Before getting a chance to practice with the algebraic aspects of Lagrange multipliers, the next activity gives you a chance to review your geometric understanding of the method.

Activity 11.11.2

(a) Below is the contour plot of a function $f(x, y)$. Also shown are two lines, which are constraint curves. Use the plot to identify the extreme value(s) of $f(x, y)$ subject to the constraint $g_1(x, y) = c_1$ and the point(s) at which these extrema occur. Repeat this for the constraint $g_2(x, y) = c_2$. Write a sentence to explain your reasoning.



(b) The circular curves in the graph below are contours of a function $T(x, y)$. Use the plot to identify the extreme value(s) of $T(x, y)$ subject to the constraint $h(x, y) = k$, which is also shown in the plot. Also identify the points at which the extreme value(s) occur and write a sentence to explain your reasoning.



The next activity gives you an opportunity to practice the algebraic aspects of solving a constrained optimization problem using the method of Lagrange multipliers.

Activity 11.11.3 A cylindrical soda can holds about 355 cc of liquid. In this activity, we want to find the dimensions of such a can that will minimize the surface area, since this will minimize the material used to construct this can. For the sake of simplicity, assume the can is a perfect cylinder.

(a) What are the variables in this problem? Based on the context, what restriction(s), if any, are there on these variables? You will likely find it useful to draw a picture of your cylindrical can and label your variables.

(b) What quantity do we want to optimize in this problem? What equation describes the constraint? You need to decide which of these functions

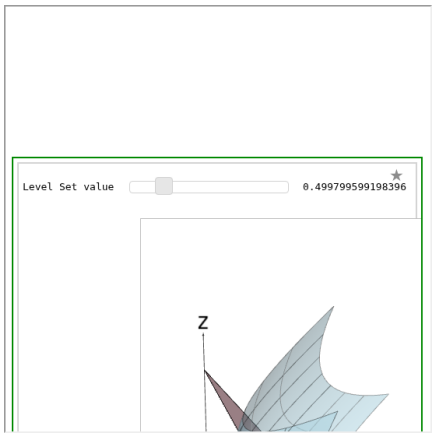
plays the role of the objective function, f, and which plays the role of the constraint function, g, in our discussion of Lagrange multipliers.

(c) Use your choices for the objective and constraint equations to set up a system of equations in terms of your variables and λ corresponding to the Lagrange Multipliers. Find λ and the values of your variables that satisfy Equation (11.11.1) in the context of this problem.

(d) Determine the dimensions of the soda can that give the desired solution to this constrained optimization problem.

The method of Lagrange multipliers also works for functions of three variables. That is, if we have a function $f = f(x, y, z)$ that we want to optimize subject to a constraint $g(x, y, z) = c$, the optimal point (x, y, z) lies on the level surface S defined by the constraint $g(x, y, z) = k$. Because gradient vectors are orthogonal to level *surfaces* as well, the argument we made to justify the need for the gradient of the objective function to be parallel to the gradient of the constraint at the location of an exreme value still holds. In other words, just as in the two variable case, we can optimize $f(x, y, z)$ subject to the constraint $g(x, y, z) = c$ by finding all points (x, y, z) that satisfy $\nabla f = \lambda \nabla g$ and $g(x, y, z) = c$. In fact, this method also works for functions of an arbitrary number of variables.

In Figure 11.11.5, you can see a level set of a three variable objective function plotted in light blue and a constraint given by the pink plane. This is three dimensional analog to our plot earlier. You can use the slider at the top of Figure 11.11.5 to find the value for our objective function where the level surface is tangent to the pink constraint. When these surfaces are tangent, their gradients, visualized as normal vectors to the surfaces, will be parallel.



Standalone
In Context
Embed

Figure 11.11.5 A three dimensional plot of a level surface for $f(x, y, z)$ shown in blue and a constraint plane shown in pink.

Example 11.11.6 illustrates how to use the method of Lagrange multipliers to optimize a function of three variables. The problem in Activity 11.11.3 is one that you could have solved using only single-variable calculus techniques such as solving the constraint equation for one variable and substituting into the objective function to reduce the number of variable. Example 11.11.6 illustrates how Lagrange multipliers can be used to help solve problems that are not amenable to that strategy.

Example 11.11.6 Suppose that we have created a three-dimensional coordinate system for a heated chamber with the origin at the center of the chamber. The temperature, in degrees Celsius, at the point (x, y, z) inside the chamber

is given by $T(x, y, z) = 6x - 3y^2 + 8z + 400$. We wish to find the hottest and coldest points on the sphere $x^2 + y^2 + z^2 = 100$.

To fit our usual framework of a Lagrange multipliers problem, let $g(x, y, z) = x^2 + y^2 + z^2$. Our goal here then becomes to optimize T subject to the constraint $g(x, y, z) = 100$. We begin by computing the gradients:

$$\nabla T = \langle 6, -3y, 8 \rangle$$
$$\nabla g = \langle 2x, 2y, 2z \rangle$$

Since the method of Lagrange multipliers tells us that the extreme values will occur where the gradients are parallel, we need $\nabla T = \lambda \nabla g$ as well as $g(x, y, z) = 100$. This leads to the following equations:

$$6 = 2\lambda x \tag{11.11.5}$$
$$-6y = 2\lambda y \tag{11.11.6}$$
$$8 = 2\lambda z \tag{11.11.7}$$
$$x^2 + y^2 + z^2 = 100 \tag{11.11.8}$$

The second equation can be rewritten as

$$0 = 2\lambda y + 6y = 2y(\lambda + 3).$$

Thus, we know that solutions of the system must have $y = 0$ or $\lambda = -3$. We consider each of these cases separately:

$y = 0$ In this case, we first isolate x and z in (11.11.5) and (11.11.7), respectively. We then substitute $x = 3/\lambda$, $y = 0$, and $z = 4/\lambda$ into (11.11.8), which gives us

$$\left(\frac{3}{\lambda}\right)^2 + 0^2 + \left(\frac{4}{\lambda}\right)^2 = 100$$
$$9 + 16 = 100\lambda^2$$
$$\frac{1}{4} = \lambda^2$$
$$\pm\frac{1}{2} = \lambda.$$

We can now substitute the values of λ into (11.11.5) and (11.11.7). This gives us two points: $(6, 0, 8)$ and $(-6, 0, -8)$.

$\lambda = -3$ By using (11.11.5) and (11.11.7), this leads us to find directly that $x = -1$ and $z = -4/3$. Now we can substitute into (11.11.8) and solve for y, which gives $y = \pm 5\sqrt{35}/3$. Therefore, we have two points to consider: $(-1, 5\sqrt{35}/3, -4/3)$ and $(-1, -5\sqrt{35}/3, -4/3)$.

Now that we have found the four potential locations of our extreme values, we need so evaluate T at each of these points. The points and the temperature values are shown below.

(x, y, z)	$T(x, y, z)$
$(6, 0, 8)$	500
$(-6, 0, -8)$	300
$(-1, 5\sqrt{35}/3, -4/3)$	275/3
$(-1, -5\sqrt{35}/3, -4/3)$	275/3

Therefore, we can see that the minimum temperature of a point on the sphere is $275/3 \approx 91.67°$C, which occurs at the points $(-1, 5\sqrt{35}/3, -4/3)$ and $(-1, -5\sqrt{35}/3, -4/3)$. The maximum temperature of a point on the sphere is 500°C, which occurs at the point $(6, 0, 8)$. ▲

If it has not yet become apparent to you, solving the system of equations is generally the most difficult part of using the method of Lagrange multipliers. In most cases, this system is *nonlinear*, meaning that at least one variable will appear to a power other than 1 or variables are multiplied (or perhaps trigonometric, exponential, or logarithmic functions appear). We cannot provide a general method that will work to solve every nonlinear system of equations, so you may need to experiment a bit to find a the solutions. The final activity of this section gives you an opportunity to practice this for another function of three variables.

Activity 11.11.4 Use the method of Lagrange multipliers to find the dimensions of the least expensive rectangular packing crate with a volume of 240 cubic feet when the material for the top costs \$2 per square foot, the bottom costs \$3 per square foot, and the sides cost \$1.50 per square foot.

Summary

- The extrema of a function $f = f(x, y)$ subject to a constraint $g(x, y) = c$ occur at points for which the contour of f is tangent to the curve that represents the constraint equation. This occurs when

$$\nabla f = \lambda \nabla g.$$

- We use the condition $\nabla f = \lambda \nabla g$ to generate a system of equations, together with the constraint $g(x, y) = c$, that may be solved for x, y, and λ. Once we have all the solutions, we evaluate f at each of the (x, y) points to determine the extrema.

11.11.3 Exercises

1. Use Lagrange multipliers to find the maximum and minimum values of $f(x, y) = 3x - 5y$ subject to the constraint $x^2 + 2y^2 = 86$, if such values exist.

maximum = _____
minimum = _____
*(For either value, enter **DNE** if there is no such value.)*

2. Use Lagrange multipliers to find the maximum and minimum values of $f(x, y) = x^2y + 3y^2 - y$, subject to the constraint $x^2 + y^2 \leq 38.3333333333333$

maximum = _____
minimum = _____
*(For either value, enter **DNE** if there is no such value.)*

3. Find the absolute maximum and minimum of the function $f(x, y) = x^2 + y^2$ subject to the constraint $x^4 + y^4 = 6561$.

As usual, ignore unneeded answer blanks, and list points in lexicographic order.

Absolute minimum value: _____
attained at (_____, _____), (_____, _____), (_____, _____), (_____, _____).
Absolute maximum value: _____

attained at (_____, _____), (_____, _____),
(_____, _____), (_____, _____).

4. Find the absolute maximum and minimum of the function $f(x, y) = x^2 - y^2$ subject to the constraint $x^2 + y^2 = 361$.

As usual, ignore unneeded answer blanks, and list points in lexicographic order.

Absolute minimum value: _____
attained at (_____, _____) and (_____, _____).
Absolute maximum value: _____
attained at (_____, _____) and (_____, _____).

5. Find the minimum distance from the point $(1, 1, 12)$ to the paraboloid given by the equation $z = x^2 + y^2$.

Minimum distance = _____

Note: If you need to find roots of a polynomial of degree ≥ 3, you may want to use a calculator of computer to do so numerically. Also be sure that you can give a geometric justification for your answer.

6. For each value of λ the function $h(x, y) = x^2 + y^2 - \lambda(2x + 4y - 18)$ has a minimum value $m(\lambda)$.

(a) Find $m(\lambda)$

$m(\lambda) =$ _____

*(Use the letter **L** for λ in your expression.)*

(b) For which value of λ is $m(\lambda)$ the largest, and what is that maximum value?

$\lambda =$ _____

maximum $m(\lambda) =$ _____

(c) Find the minimum value of $f(x, y) = x^2 + y^2$ subject to the constraint $2x+4y = 18$ using the method of Lagrange multipliers and evaluate λ.

minimum $f =$ _____

$\lambda =$ _____

(How are these results related to your result in part (b)?)

7. The plane $x+y+2z = 6$ intersects the paraboloid $z = x^2+y^2$ in an ellipse. Find the points on this ellipse that are nearest to and farthest from the origin.

Point farthest away occurs at
(_____, _____,_____).
Point nearest occurs at
(_____, _____,_____).

8. Find the maximum and minimum values of the function $f(x, y, z) = x^2y^2z^2$ subject to the constraint $x^2 + y^2 + z^2 = 64$.

Maximum value is _____, occurring at _____ points (positive integer or "infinitely many").

Minimum value is _____, occurring at _____ points (positive integer or "infinitely many").

9. Find the maximum and minimum values of the function $f(x, y, z, t) = x + y + z + t$ subject to the constraint $x^2 + y^2 + z^2 + t^2 = 100$.

Maximum value is _____, occurring at _____ points (positive integer or "infinitely many").

Minimum value is _____, occurring at _____ points (positive integer or "infinitely many").

10. Find the maximum and minimum volumes of a rectangular box whose surface area equals 7000 square cm and whose edge length (sum of lengths of all edges) is 440 cm.

Hint: It can be deduced that the box is not a cube, so if x, y, and z are the lengths of the sides, you may want to let x represent a side with $x \neq y$ and $x \neq z$.

Maximum value is _____,
occurring at (_____, _____, _____).

Minimum value is _____,
occurring at (_____, _____, _____).

11. ***(a)*** If $\sum_{i=1}^{3} x_i = 4$, find the values of x_1, x_2, x_3 making $\sum_{i=1}^{3} {x_i}^2$ minimum.

$x_1, x_2, x_3 =$ _____

(Give your values as a comma separated list.)

(b) Generalize the result of part (a) to find the minimum value of $\sum_{i=1}^{n} {x_i}^2$ subject to $\sum_{i=1}^{n} x_i = 4$.

minimum value = _____

12. The Cobb-Douglas production function is used in economics to model production levels based on labor and equipment. Suppose we have a specific Cobb-Douglas function of the form

$$f(x, y) = 50x^{0.4}y^{0.6},$$

where x is the dollar amount spent on labor and y the dollar amount spent on equipment. Use the method of Lagrange multipliers to determine how much should be spent on labor and how much on equipment to maximize productivity if we have a total of 1.5 million dollars to invest in labor and equipment.

13. Use the method of Lagrange multipliers to find the point on the line $x - 2y = 5$ that is closest to the point $(1, 3)$. To do so, respond to the following prompts.

a. Write the function $f = f(x, y)$ that measures the *square* of the distance from (x, y) to $(1, 3)$. (The extrema of this function are the same as the extrema of the distance function, but $f(x, y)$ is simpler to work with.)

b. What is the constraint $g(x, y) = c$?

c. Write the equations resulting from $\nabla f = \lambda \nabla g$ and the constraint. Find all the points (x, y) satisfying these equations.

d. Test all the points you found to determine the extrema.

14. Apply the Method of Lagrange Multipliers to solve each of the following constrained optimization problems.

a. Determine the absolute maximum and absolute minimum values of $f(x, y) = (x-1)^2 + (y-2)^2$ subject to the constraint that $x^2 + y^2 = 16$.

b. Determine the points on the sphere $x^2+y^2+z^2 = 4$ that are closest to and farthest from the point $(3, 1, -1)$. (As in the preceding exercise, you may find it simpler to work with the square of the distance

formula, rather than the distance formula itself.)

c. Find the absolute maximum and minimum of $f(x,y,z) = x^2+y^2+z^2$ subject to the constraint that $(x-3)^2+(y+2)^2+(z-5)^2 \leq 16$. (Hint: here the constraint is a closed, bounded region. Use the boundary of that region for applying Lagrange Multipliers, but don't forget to also test any critical values of the function that lie in the interior of the region.)

15. In this exercise we consider how to apply the Method of Lagrange Multipliers to optimize functions of three variable subject to two constraints. Suppose we want to optimize $f = f(x,y,z)$ subject to the constraints $g(x,y,z) = c$ and $h(x,y,z) = k$. Also suppose that the two level surfaces $g(x,y,z) = c$ and $h(x,y,z) = k$ intersect at a curve C. The optimum point $P = (x_0, y_0, z_0)$ will then lie on C.

a. Assume that C can be represented parametrically by a vector-valued function $\vec{r} = \vec{r}(t)$. Let $\overrightarrow{OP} = \vec{r}(t_0)$. Use the Chain Rule applied to $f(\vec{r}(t))$, $g(\vec{r}(t))$, and $h(\vec{r}(t))$, to explain why

$$\begin{aligned} \nabla f(x_0,y_0,z_0) \cdot \vec{r}\,'(t_0) &= 0, \\ \nabla g(x_0,y_0,z_0) \cdot \vec{r}\,'(t_0) &= 0, \text{ and} \\ \nabla h(x_0,y_0,z_0) \cdot \vec{r}\,'(t_0) &= 0. \end{aligned}$$

Explain how this shows that $\nabla f(x_0,y_0,z_0)$, $\nabla g(x_0,y_0,z_0)$, and $\nabla h(x_0,y_0,z_0)$ are all orthogonal to C at P. This shows that $\nabla f(x_0,y_0,z_0)$, $\nabla g(x_0,y_0,z_0)$, and $\nabla h(x_0,y_0,z_0)$ all lie in the same plane.

b. Assuming that $\nabla g(x_0,y_0,z_0)$ and $\nabla h(x_0,y_0,z_0)$ are nonzero and not parallel, explain why every point in the plane determined by $\nabla g(x_0,y_0,z_0)$ and $\nabla h(x_0,y_0,z_0)$ has the form $s\nabla g(x_0,y_0,z_0) + t\nabla h(x_0,y_0,z_0)$ for some scalars s and t.

c. Parts (a.) and (b.) show that there must exist scalars λ and μ such that

$$\nabla f(x_0,y_0,z_0) = \lambda \nabla g(x_0,y_0,z_0) + \mu \nabla h(x_0,y_0,z_0).$$

So to optimize $f = f(x,y,z)$ subject to the constraints $g(x,y,z) = c$ and $h(x,y,z) = k$ we must solve the system of equations

$$\begin{aligned} \nabla f(x,y,z) &= \lambda \nabla g(x,y,z) + \mu \nabla h(x,y,z), \\ g(x,y,z) &= c, \text{ and} \\ h(x,y,z) &= k. \end{aligned}$$

for x, y, z, λ, and μ.

Use this idea to find the maximum and minimum values of $f(x,y,z) = x + 2y$ subject to the constraints $y^2 + z^2 = 8$ and $x + y + z = 10$.

16. There is a useful interpretation of the Lagrange multiplier λ. Assume that we want to optimize a function f with constraint $g(x,y) = c$. Recall that an optimal solution occurs at a point (x_0, y_0) where $\nabla f = \lambda \nabla g$. As the constraint changes, so does the point at which the optimal solution occurs. So we can think of the optimal point as a function of the parameter

c, that is $x_0 = x_0(c)$ and $y_0 = y_0(c)$. The optimal value of f subject to the constraint can then be considered as a function of c defined by $f(x_0(c), y_0(c))$. The Chain Rule shows that

$$\frac{df}{dc} = \frac{\partial f}{\partial x_0}\frac{dx_0}{dc} + \frac{\partial f}{\partial y_0}\frac{dy_0}{dc}.$$

a. Use the fact that $\nabla f = \lambda \nabla g$ at (x_0, y_0) to explain why

$$\frac{df}{dc} = \lambda \frac{dg}{dc}.$$

b. Use the fact that $g(x, y) = c$ to show that

$$\frac{df}{dc} = \lambda.$$

Conclude that λ tells us the rate of change of the function f as the parameter c increases. In other words, the value of λ is approximately how much the optimal value of the function f will change if we increase the value of c by 1 unit.

c. In Example 11.11.3, we did not explicity calculate the value of λ for our maximum value of the objective function. Use our work in Example 11.11.3 to find the value of λ and explain in context of our Preview Activity what the value λ tells us about the package.

d. Suppose that the maximum value of a function $f = f(x, y)$ subject to a constraint $g(x, y) = 100$ is 236. When using the method of Lagrange multipliers and solving $\nabla f = \lambda \nabla g$, we obtain a value of $\lambda = 15$ at this maximum. Find an approximation to the maximum value of f subject to the constraint $g(x, y) = 98$.

Chapter 12

Multiple Integrals

12.1 Integrating Functions of Several Variables

In single-variable calculus, recall that we used the classic calculus approach to define the definite integral as the area under a graph. Specifically, we approximated the area under the graph of a positive function f on an interval $[a, b]$ by adding areas of rectangles whose heights are determined by the curve. We then broke the interval $[a, b]$ into smaller subintervals, constructing rectangles on each of these smaller intervals to approximate the region under the curve on that subinterval, then summing the areas of these rectangles to approximate the area under the curve. We defined the definite integral of f using the limit of this Riemann sum as the size of all of the subintervals goes to zero.In this chapter we will expand our idea of integration and accumulation to apply to scalar valued, multivariable function.

12.2 Double Riemann Sums and Double Integrals over Rectangles

Motivating Questions

- How can we extend the idea of a Riemann sum from single-variable calculus to functions of two variables?
- How is the double integral of a continuous function $f = f(x, y)$ defined?
- How can we interpret the double integral of a function of two variables?

12.2.1 Introduction

In the Preview Activity, we will review a few ideas from integrals in single-variable calculus. The rest of this section will then be used to extend the ideas of integration and its interpretations to functions of two variables over a rectangular region.

Preview Activity 12.2.1

(a) A plot of f for inputs in the interval $[0, 2]$ is shown in Figure 12.2.1. Break the interval $[0, 2]$ into four equally sized subintervals and draw the rectangles that would be used to construct a Riemann sum to approximate the area under f on the interval $[0, 2]$. You can use whichever point you

want on each subinterval to evaluate the height the of the rectangles.

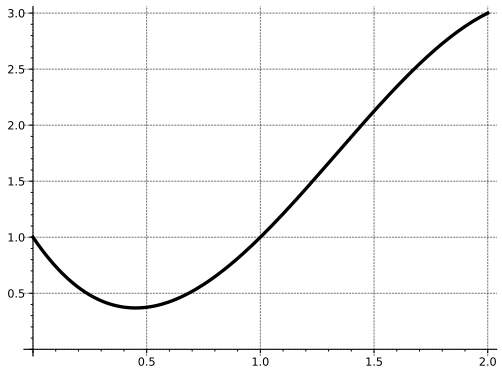


Figure 12.2.1 A continuous function on $[0, 2]$

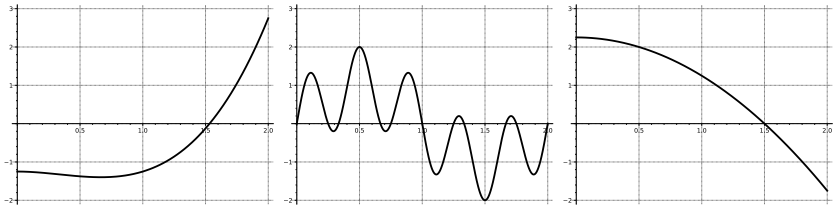
(b) Estimate the heights of the rectangles used in your Riemann sum above to estimate $\int_0^2 f(x)\ dx$. Write a couple of sentences about why you think your estimate for the definite integral of f is either an overestimate, an underestimate, or close to the true value.

(c) Recall that the definite integral is defined as

$$\int_a^b f(x)\ dx = \lim_{\Delta x_i \to 0} \sum_{i=1}^{n} f(x_i^*)\Delta x_i.$$

Explain why it doesn't matter what method (left endpoint, right endpoint, midpoint, etc.) you use for selecting which point is evaluated on each of the subintervals in the definition of the definite integral.

(d) For each of the functions plotted below, determine if the definite integral over the region shown will be positive, negative, or zero and write a sentence to justify your answer.

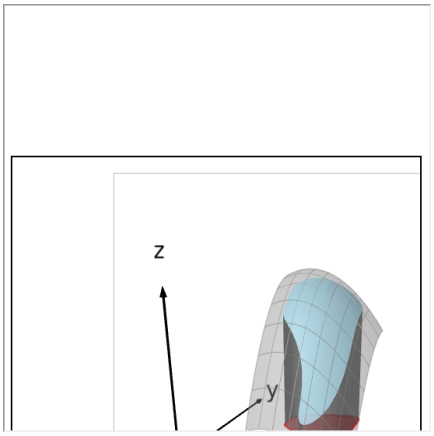


In this section, we will use the classic calculus approach to define and understand the definite integral for a function of two variables over a rectangular region of input values. Just as in single-variable calculus, we will develop some algebraic methods to help efficiently evaluate these integrals. Later in this chapter, we will will generalize to regions that are not rectangular and work with functions of three or more variables.

12.2.2 Double Riemann Sums over Rectangles

The motivating interpretation of a definite integral for a function of one variable was the area under a curve over a particular interval of input values. We will use the same motivating interpretation for the definite integral of a function of two variables. The definite integral of a nonnegative function of two variables will measure the volume of the region beneath the graph of f over a region of

input points. The definite integral for a function of two variables is called the double integral of the function f. We can show this geometrically as the volume of the solid below the surface given by $z = f(x, y)$ as shown in Figure 12.2.2. The surface given by $z = f(x, y)$ is shown in blue and the region of inputs above which we want to find the volume is shown in red in the xy-plane. The double integral of f over the red region should measure the volume of the solid shaded in gray.



Standalone
In Context
Embed

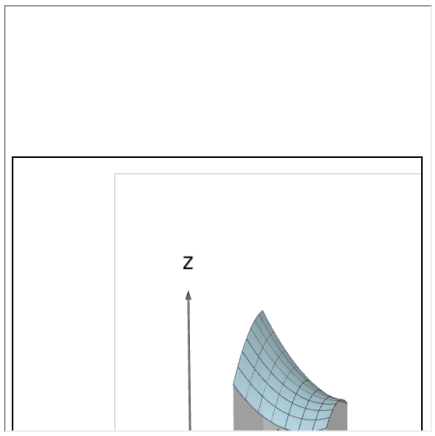
Figure 12.2.2 A plot of the volume under a surface $z = f(x, y)$ over a region R, in red

The region shaded red in Figure 12.2.2 is irregular in the coordinates (x and y). We will explore this kind of problem in Section 12.4 once we have some intuition and tools from examining rectangular regions of integration. To discuss how to approach rectangular regions, we apply the classic calculus approach to the region shown in Figure 12.2.3. In particular, for this rectangular region of input points, we want to

1. estimate the volume between the blue surface plot and the xy-plane,
2. quantify how this estimate for the volume of the solid under the graph of $f(x, y)$ changes when we use a smaller scale, and
3. use a limit to find the exact value of the volume under the surface.

To estimate the volume under the graph, we will use rectangular prisms because their volume is easy to compute:

$$\text{Volume of rectangular prism} = (\text{length})(\text{width})(\text{height})$$



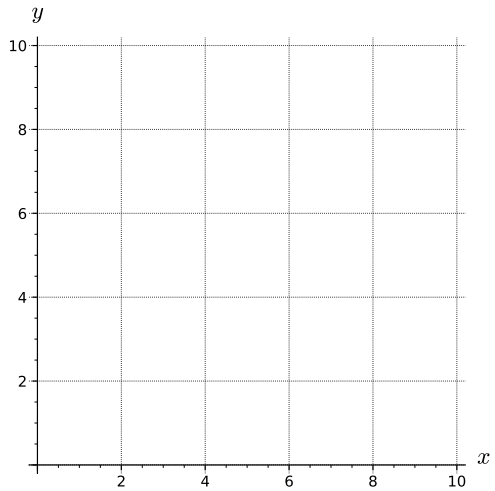
Standalone
In Context
Embed

Figure 12.2.3 A plot of the volume under a surface over a rectangular region

In the next activity, we will go through these steps for a particular function and rectangular region of integration as well as introduce some of the notation used in our development of double integrals.

Activity 12.2.2 In this activity, we will go through the first two steps of the classic calculus approach to compute the definite integral of $f(x, y) = 100 - x^2 - y^2$ on the rectangular domain $R = [0, 8] \times [2, 6]$. Remember that we are trying to measure the volume below the graph of f over the region R, so we will start with estimating this volume.

(a) To understand the numerical calculations involved in the classic calculus approach for a double integral, it is most important to understand the region of integration. Thus, we will not look at a graph of $z = f(x, y)$. Instead, we will stay focused in the xy-plane. On the axes below, outline the rectangular region R that corresponds to the region of integration.



(b) Because all of the regions and subregions we are considering in this section are rectangles, we can break up the x and y coordinates into pieces separately. For this activity, break the interval of x-coordinates into *four* equally-sized subintervals and break the interval of y-coordinates into *three* equally-sized subintervals.

What is the length of each subinterval in the x-direction? How about in the y-direction?

(c) Let S_i be the i-th subinterval for x. We want to state the endpoints of each of the S_i. The first subinterval, S_1, will go from x_0 to x_1, the second subinterval, S_2, will go from x_1 to x_2, the third subinterval, S_3, will go from x_2 to x_3, and the fourth subinterval, S_4, will go from x_3 to x_4.

Give the values for x_0, x_1, x_2, x_3, and x_4 and add these as tick marks on the x-axis of the graph in part (a) to make sure your subintervals are equally sized.

(d) Let T_i be the i-th subinterval for y. Give the values for y_0, y_1, y_2, and y_3 and add these as tick marks on the y-axis of the graph in part (a) to make sure your subintervals are equally sized.

(e) We will use the subintervals in the x- and y-coordinates to specify smaller rectangles into which R is divided to compute an approximation. Let R_{ij} be the rectangle corresponding to $S_i \times T_j$.

i. How many smaller rectangles are there in this partition?

ii. Outline each of the smaller rectangles on the graph in part (a) and label each rectangle as either $R_{11}, R_{12}, ...$

iii. Since each smaller rectangle R_{ij} has the same area, let ΔA denote the area of each of these smaller rectangles. What is ΔA?

(f) To find the volume of the rectangular prisms over each R_{ij}, we must pick a point in each subrectangle at which to evaluate f, which will be the height of the rectangular prism. For this activity, we will use the upper-right corner of each subrectangle as the designated point.

State the point at the upper right of each smaller rectangle and evaluate f at each of these points.

(g) Write a sentence about why the volume of each rectangular prism used for this approximation is

$$f(x_i, y_j)\Delta A.$$

Remember that x_i and y_j are the points from parts c and d of this activity. Write a couple sentences about how you would find an approximation of the volume under the surface $z = f(x, y)$ over the region R. (Do not do calculation, but rather explain what calculation is being done.)

(h) We chose the upper-right point of each subrectangle R_{ij} to find the height of each rectangular prism used in the approximation. Write a sentence or two about whether you think the upper right point provides an overestimate, an underestimate, or approximately the average value for f on each R_{ij}. Explain how this suggests that your estimate for the volume under the surface $z = f(x, y)$ over the region R is either an overestimate, an underestimate, or approximately the correct value.

12.2.3 Double Riemann Sums and Double Integrals

Every part of the previous activity can be generalized to work on any bounded rectangular region using any number of subintervals in either the x or y directions. This leads us to state these steps in general below and note that using more subintervals corresponds to estimating on a smaller scale, which is step 2 of the classic calculus approach.

Definition 12.2.4 Let f be a continuous function on a rectangle $R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$. A **double Riemann sum for f over R** is created as follows.

1. Partition the interval $[a, b]$ into m subintervals of equal length $\Delta x = \frac{b-a}{m}$. Let x_0, x_1, ..., x_m be the endpoints of these subintervals, where $a = x_0 < x_1 < x_2 < \cdots < x_m = b$.

2. Partition the interval $[c, d]$ into n subintervals of equal length $\Delta y = \frac{d-c}{n}$. Let y_0, y_1, ..., y_n be the endpoints of these subintervals, where $c = y_0 < y_1 < y_2 < \cdots < y_n = d$.

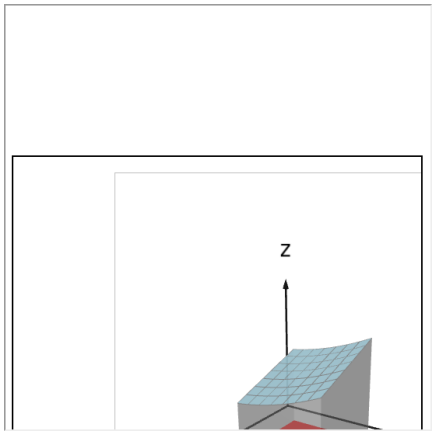
3. These two partitions create a partition of the rectangle R into mn subrectangles R_{ij} with opposite vertices (x_{i-1}, y_{j-1}) and (x_i, y_j) for i between 1 and m and j between 1 and n. These rectangles all have equal area $\Delta A = \Delta x \cdot \Delta y$.

4. Choose a point (x^*_{ij}, y^*_{ij}) in each rectangle R_{ij}. A double Riemann sum for f over R is given by

$$\sum_{j=1}^{n} \sum_{i=1}^{m} \left(f(x^*_{ij}, y^*_{ij}) \cdot \Delta A \right).$$

◆

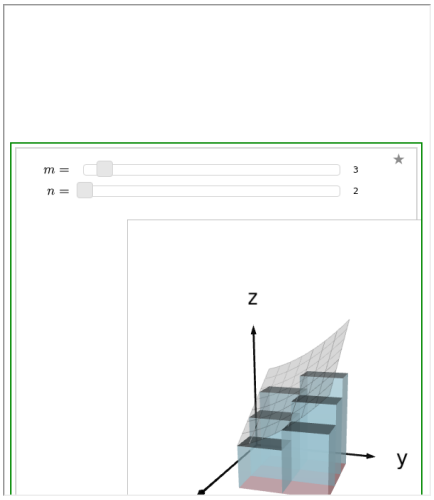
If $f(x, y) \geq 0$ on the rectangle R, we may ask to find the volume of the solid bounded above by f over R, as illustrated in Figure 12.2.5.



Standalone
In Context
Embed

Figure 12.2.5 The solid below a surface $z = f(x, y)$ over a rectangular region

This volume is approximated by a Riemann sum, which sums the volumes of the rectangular prisms shown in Figure 12.2.6. In a double Riemann sum, both m and n go to infinity. This means that the number of subrectangles increases without bound, as illustrated in Figure 12.2.6. As this happens, the sum of the volumes of the rectangular boxes approaches the true volume of the solid bounded above by $z = f(x, y)$ over the region R. Use the sliders at the top of Figure 12.2.6 to change the number of subintervals used in the Riemann sum and verify geometrically how the estimated volume will approach true volume under f as n and m become arbitrarily large.



Standalone
In Context
Embed

Figure 12.2.6 The estimated volume under a surface $z = f(x, y)$ when using m by n subrectangles

The third step in the classic calculus approach is to take the limit, and when this limit exists, its value is the double integral we have been seeking:

Definition 12.2.7 Let R be a rectangular region in the xy-plane and f a continuous function over R. With terms defined as in a double Riemann sum, the **double integral of f over R** is

$$\iint_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \cdot \Delta A.$$

◆

Some textbooks use the notation $\int_R f(x, y)\, dA$ for a double integral. You will see this in some of the WeBWorK problems.

For the solid shown in Figure 12.2.6, the exact volume is $231/64 = 3.609375$. Go back and see how close the estimate of the volume provided by the double Riemann sum is to the exact value for the largest values of n and m the sliders will allow.

We conclude this subsection by noting a couple of useful ideas in the development of double integrals that can simplify some aspects. Many times we will use equally-sized subintervals or pieces of our region of integration but that is not strictly necessary. The most general form of a double integral requires only that as we take the limit, the area of *every* smaller piece goes to zero. In Activity 12.2.2, we used the upper right point of each smaller rectangle to evaluate f but in Definition 12.2.7 there is not a particular point in each smaller rectangle used. Remember that as the area of *all* the smaller rectangles is going to zero, there will be smaller and smaller differences between the output of f on the points in the smaller rectangle.

12.2.4 Interpretation of Double Riemann Sums and Double integrals

We conclude this section with three important interpretations of the double integral that will appear repeatedly throughout the rest of this book. We summarize these three interpretations in the Key Idea below and then discuss each in greater depth.

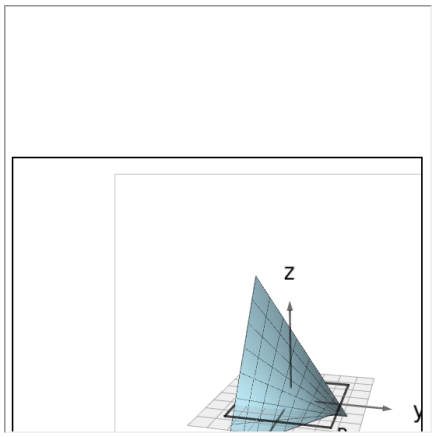
Key Idea 12.2.8 *If f is a function of two variables on a rectangular region R in the xy-plane, the double integral $\iint_R f(x,y)\,dA$ can be interpreted in the following ways:*

1. *The double integral can be interpreted as the "signed" volume for the solid enclosed by the surface and the xy-plane.*
2. *If f represents the density of some substance, then the double integral represents the mass of the substance.*
3. *Dividing the double integral by the area of R gives the average value of f over the region R.*

Signed Volume. Suppose that $f(x,y)$ assumes both positive and negative values on the rectangle R, as shown in Figure 12.2.9. When constructing a Riemann sum, for each i and j, the product $f(x_{ij}^*, y_{ij}^*)\ \Delta A$ can be interpreted as a "signed" volume of a box with base area ΔA and "signed" height $f(x_{ij}^*, y_{ij}^*)$. Since f can have negative values, this "height" could be negative. The sum

$$\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*)\ \Delta A$$

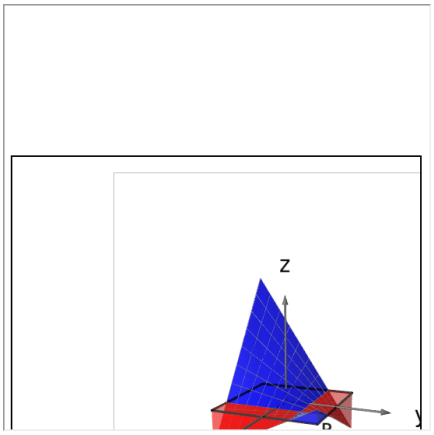
can then be interpreted as a sum of "signed" volumes of boxes, with a negative sign attached to those boxes whose heights are below the xy-plane.



Standalone
In Context
Embed

Figure 12.2.9 A plot of the surface $z = f(x,y)$ over the rectangular region R

We can then realize the double integral $\iint_R f(x,y)\,dA$ as a difference in volumes: $\iint_R f(x,y)\,dA$ tells us the volume of the solids the graph of f bounds above the xy-plane over the rectangle R minus the volume of the solids the graph of f bounds below the xy-plane under the rectangle R. This is shown in Figure 12.2.10.



Standalone
In Context
Embed

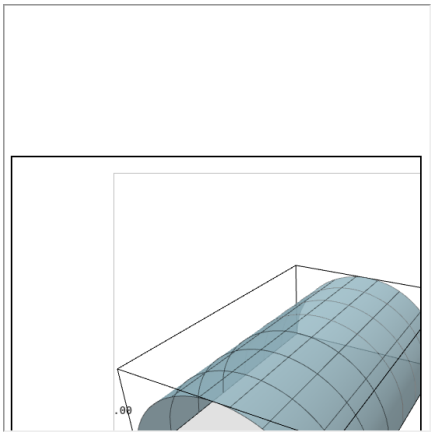
Figure 12.2.10 A plot of the signed volume with the volume above the xy-plane shown in blue and the volume below the xy-plane shown in red

The double integral of f over the rectangle R will be positive because there is more volume above the xy-plane (shown in blue) than volume below the xy-plane, which is counted as negative and shown in red.

Activity 12.2.3 Let $f(x,y) = \sqrt{4-y^2}$ on the rectangular region $R = [1,7] \times [-2,2]$. Suppose that we partition $[1,7]$ into 3 subintervals of equal length and $[-2,2]$ into 2 subintervals of equal length. A table of values of f at some points in R is given in Table 12.2.11, and a graph of f with the indicated partitions is shown in Figure 12.2.12.

Table 12.2.11 Table of values of $f(x,y) = \sqrt{4-y^2}$.

$x \downarrow \setminus y \rightarrow$	-2	-1	0	1	2
1	0	$\sqrt{3}$	2	$\sqrt{3}$	0
2	0	$\sqrt{3}$	2	$\sqrt{3}$	0
3	0	$\sqrt{3}$	2	$\sqrt{3}$	0
4	0	$\sqrt{3}$	2	$\sqrt{3}$	0
5	0	$\sqrt{3}$	2	$\sqrt{3}$	0
6	0	$\sqrt{3}$	2	$\sqrt{3}$	0
7	0	$\sqrt{3}$	2	$\sqrt{3}$	0



Standalone
In Context
Embed

Figure 12.2.12 A plot of the surface given by $f(x,y) = \sqrt{4-y^2}$

(a) Sketch the region R in the plane partitioned as described above.

(b) Calculate the double Riemann sum using the given partition of R and the values of f in the upper right corner of each subrectangle.

(c) Use geometry to calculate the exact value of $\iint_R f(x,y)\,dA$ and compare it to your approximation. Write a sentence to describe one way to obtain a better approximation using the given data.

Density Functions. If the function being integrated is a density function, then the double integral measures how much of the substance being measured is in the region of integration. You can see a single-variable calculus explanation of this in Recovering Position from Velocity, and we will explore this more in the multivariable setting in Section 12.5 and later applications.

As a brief example here, if f measures the density of some material on a rectangular plot of land, then the double integral of this function will measure the total amount of that material on the plot of land. The function $f(x,y)$ will have units of mass (of the material) per unit area and the Riemann sum

$$\sum_{j=1}^{n}\sum_{i=1}^{m}\left(f(x_{ij}^*, y_{ij}^*)\cdot \Delta A\right)$$

will have units of mass. This idea is used for the laminar masses given in Section 12.5 and applies to other settings such as electric charge, which has the possibility of negative values for both the density and total amount.

Average Value. The average of mn values $f(x_{ij}^*, y_{ij}^*)$ of a function of two variables can be expressed as a double sum:

$$\text{Avg}_{mn} = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*)$$

If we take the limit as m and n go to infinity, we obtain what we can define as the average value of f over the region R

Although the double sum expression for Avg_{mn} appears to be missing an essential component to recognize it as a double Riemann sum that corresponds to a double integral, we can do some algebra to adjust for that. First, note that

$$\Delta x = \frac{b-a}{m} \qquad \text{and} \qquad \Delta y = \frac{d-c}{n}.$$

Thus,

$$\frac{1}{mn} = \frac{\Delta x\cdot\Delta y}{(b-a)(d-c)} = \frac{\Delta A}{\text{area of } R},.$$

Then, the average value of the function f over R, which we denote by $f_{\text{AVG}(R)}$, is given by

$$\begin{aligned} f_{\text{AVG}(R)} &= \lim_{m,n\to\infty}\frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \\ &= \lim_{m,n\to\infty}\frac{1}{A(R)}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*)\cdot\Delta A \\ &= \frac{1}{\text{area of } R}\iint_R f(x,y)\,dA. \end{aligned}$$

Therefore, the double integral of f over R *divided by the area of* R is the average value of the function f on R.

As an additional note, if $f(x,y) \geq 0$ on R, then we can interpret this average value of f on R as the height of the box with base R that has the same volume as the volume of the surface $z = f(x,y)$ over R.

Activity 12.2.4 Let $f(x,y) = x + 2y$ and let $R = [0,2] \times [1,3]$.

(a) Draw a picture of R. Partition $[0,2]$ into 2 subintervals of equal length and the interval $[1,3]$ into two subintervals of equal length. Draw these partitions on your picture of R and label the resulting subrectangles using the labeling scheme we established in the definition of a double Riemann sum.

(b) For each i and j, let (x_{ij}^*, y_{ij}^*) be the midpoint of the rectangle R_{ij}. Identify the coordinates of each (x_{ij}^*, y_{ij}^*). Draw these points on your picture of R.

(c) Calculate the Riemann sum

$$\sum_{j=1}^{n} \sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$$

using the partitions we have described. If we let (x_{ij}^*, y_{ij}^*) be the midpoint of the rectangle R_{ij} for each i and j, then the resulting Riemann sum is called a *midpoint sum*.

(d) Explain the meaning of the sum you just calculated in terms of each of the interpretations in Key Idea 12.2.8

12.2.5 Properties of Double Integrals

We conclude this section with a list of properties of double integrals. Since similar properties are satisfied by single-variable integrals and the arguments for double integrals are essentially the same, we omit their justification.

Properties of Double Integrals.

Let f and g be continuous functions on a rectangle $R = \{(x,y) : a \leq x \leq b, c \leq y \leq d\}$, and let k be a constant. Then

1. $\displaystyle\iint_R (f(x,y) + g(x,y))\, dA = \iint_R f(x,y)\, dA + \iint_R g(x,y)\, dA.$
2. $\displaystyle\iint_R kf(x,y)\, dA = k \iint_R f(x,y)\, dA.$
3. If $f(x,y) \geq g(x,y)$ on R, then $\displaystyle\iint_R f(x,y)\, dA \geq \iint_R g(x,y)\, dA.$

Summary

- Let f be a continuous function on a rectangle $R = \{(x,y) : a \leq x \leq b, c \leq y \leq d\}$. The double Riemann sum for f over R is created as follows.
  1. Partition the interval $[a,b]$ into m subintervals of equal length $\Delta x = \frac{b-a}{m}$. Let x_0, x_1, ..., x_m be the endpoints of these subintervals, where $a = x_0 < x_1 < x_2 < \cdots < x_m = b$.
  2. Partition the interval $[c,d]$ into n subintervals of equal length $\Delta y =$

$\frac{d-c}{n}$. Let $y_0, y_1, \ldots, y_n$ be the endpoints of these subintervals, where $c = y_0 < y_1 < y_2 < \cdots < y_n = d$.

3. These two partitions create a partition of the rectangle R into mn subrectangles R_{ij} with opposite vertices (x_{i-1}, y_{j-1}) and (x_i, y_j) for i between 1 and m and j between 1 and n. These rectangles all have equal area $\Delta A = \Delta x \cdot \Delta y$.
4. Choose a point (x^*_{ij}, y^*_{ij}) in each rectangle R_{ij}.

Then a double Riemann sum for f over R is given by

$$\sum_{j=1}^{n}\sum_{i=1}^{m} f(x^*_{ij}, y^*_{ij}) \cdot \Delta A.$$

- With terms defined as in the Double Riemann Sum, the double integral of f over R is

$$\iint_R f(x,y)\,dA = \lim_{m,n\to\infty} \sum_{j=1}^{n}\sum_{i=1}^{m} f(x^*_{ij}, y^*_{ij}) \cdot \Delta A.$$

- Interpretations of the double integral $\iint_R f(x,y)\,dA$ are:
  - The volume of the solids the graph of f bounds above the xy-plane over the rectangle R minus the volume of the solids the graph of f bounds below the xy-plane under the rectangle R.
  - If $f(x,y)$ is the density of a substance over some area R, then the double integral of f over R will measure the total amount of the substance in R.
  - Dividing the double integral of f over R by the area of R gives us the average value of the function f on R. If $f(x,y) \geq 0$ on R, we can interpret this average value of f on R as the height of the box with base R that has the same volume as the volume of the surface defined by f over R.

12.2.6 Exercises

1. For each of the integrals below, decide (without calculation) whether the integrals are positive, negative, or zero. Let D be the region inside the unit circle centered on the origin, T be the top half of D, R be the right half of D.

(a) $\int_T 1\,dA$ is (☐ positive ☐ negative ☐ zero)
(b) $\int_R 5y\,dA$ is (☐ positive ☐ negative ☐ zero)
(c) $\int_R 5x\,dA$ is (☐ positive ☐ negative ☐ zero)
(d) $\int_D (5y^3 + y^5)\,dA$ is (☐ positive ☐ negative ☐ zero)

2. Use a Riemann sum to estimate the volume of the solid that lies below the surface $z = xy$ and above the region $R = \{(x,y) : 0 \leq x \leq 9, 0 \leq y \leq 6\}$. Create the Riemann sum by subdividing $[0,9]$ into $m = 3$ intervals, and $[0,6]$ into $n = 2$ intervals, and take the sample point to be the upper right corner of each rectangle.

Volume: _____

3. Let $R = [0,4] \times [-1,2]$. Create a Riemann sum by subdividing $[0,4]$ into $m = 2$ intervals, and $[-1,2]$ into $n = 3$ subintervals, then use it to estimate the value of

$$\iint_R (1 - xy^2)\, dA.$$

Take the sample points to be the upper left corner of each rectangle.

Answer: _____

4. Calculate a Riemann sum $S_{3,3}$ on the square $\mathcal{R} = [0,3] \times [0,3]$ for the function $g(x,y) = f(x,y) - 8$.

The contour plot of $f(x,y)$ is shown in Figure 4.

Choose sample points and use the plot to find the values of $f(x,y)$ at these points.

Use the values of $f(x,y)$ to evaluate $g(x,y)$ accordingly.

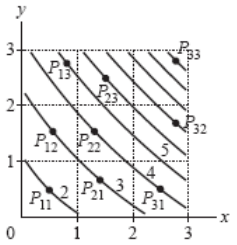


FIGURE 4 Contour plot of $f(x,y)$.

$S_{3,3} =$ _____

5. Calculate the Riemann sum for $f(x,y) = 4x - 4y$ and domain $\mathcal{D}$ in Figure 2 with two choices of sample points, $\bullet$ and $\circ$.

Which do you think is a better approximation to the integral of f over $\mathcal{D}$? Why?

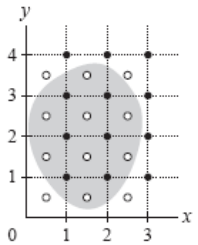


FIGURE 2

$S_{3,4}(\bullet) =$ _____
$S_{3,4}(\circ) =$ _____

6. Suppose $f(x,y) = 4 - x^2 - y^2$ and R is the rectangle with vertices (0,0), (6,0), (6,4), (0,4). In each part, estimate $\iint_R f(x,y)\, dA$ using Riemann sums. For underestimates or overestimates, consistently use either the lower left-hand corner or the upper right-hand corner of each rectangle in a subdivision, as appropriate.

(a) Without subdividing R,

Underestimate = _____

Overestimate = _____

(b) By partitioning R into four equal-sized rectangles.

Underestimate = _____

Overestimate = _____

7. Consider the solid that lies above the square (in the xy-plane) $R = [0,1] \times [0,1]$, and below the elliptic paraboloid $z = 49 - x^2 + 3xy - 4y^2$.

Estimate the volume by dividing R into 9 equal squares and choosing the sample points to lie in the midpoints of each square.

8. Let R be the rectangle with vertices $(0,0)$, $(6,0)$, $(6,6)$, and $(0,6)$ and let $f(x,y) = \sqrt{2xy}$.

(a) Find reasonable upper and lower bounds for $\int_R f\,dA$ without subdividing R.

upper bound = _____

lower bound = _____

(b) Estimate $\int_R f\,dA$ three ways: by partitioning R into four subrectangles and evaluating f at its maximum and minimum values on each subrectangle, and then by considering the average of these (over and under) estimates.

overestimate: $\int_R f\,dA \approx$ _____

underestimate: $\int_R f\,dA \approx$ _____

average: $\int_R f\,dA \approx$ _____

9. Using Riemann sums with four subdivisions in each direction, find upper and lower bounds for the volume under the graph of $f(x,y) = 4 + 3xy$ above the rectangle R with $0 \le x \le 3, \quad 0 \le y \le 4$.

upper bound = _____

lower bound = _____

10. Consider the solid that lies above the square (in the xy-plane) $R = [0,2] \times [0,2]$,

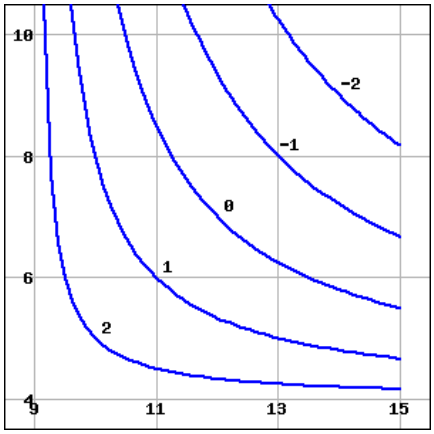
and below the elliptic paraboloid $z = 81 - x^2 - 4y^2$.

(A) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the lower left hand corners.

(B) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the upper right hand corners..

(C) What is the average of the two answers from (A) and (B)?

11. The figure below shows contours of $g(x,y)$ on the region R, with $9 \le x \le 15$ and $4 \le y \le 10$.

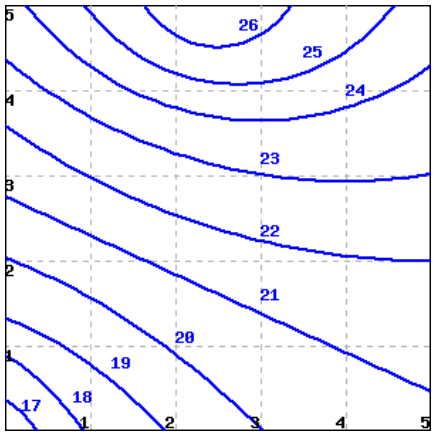


Using $\Delta x = \Delta y = 2$, find an overestimate and an underestimate for $\int_R g(x,y)\,dA$.

Overestimate = _____

Underestimate = _____

12. The figure below shows the distribution of temperature, in degrees C, in a 5 meter by 5 meter heated room.



Using Riemann sums, estimate the average temperature in the room.
average temperature = _____

13. Values of $f(x, y)$ are given in the table below. Let R be the rectangle $1 \le x \le 1.6, 2 \le y \le 3.2$. Find a Riemann sum which is a reasonable estimate for $\int_R f(x, y)\, da$ with $\Delta x = 0.2$ and $\Delta y = 0.4$. Note that the values given in the table correspond to midpoints.

$y \backslash x$	1.1	1.3	1.5
2.2	-3	8	-3
2.6	6	-1	10
3.0	2	2	7

$\int_R f(x, y)\, da \approx$ _____

14. Values of $f(x, y)$ are shown in the table below.

	$x = 1$	$x = 1.2$	$x = 1.4$
$y = 2$	6	7	10
$y = 2.3$	7	8	13
$y = 2.6$	8	9	4

Let R be the rectangle $1 \le x \le 1.4$, $2 \le y \le 2.6$. Find the values of Riemann sums which are reasonable over- and under-estimates for $\int_R f(x, y)\, dA$ with $\Delta x = 0.2$ and $\Delta y = 0.3$.
over-estimate: _____
under-estimate: _____

15. The temperature at any point on a metal plate in the xy plane is given by $T(x, y) = 100 - 4x^2 - y^2$, where x and y are measured in inches and T in degrees Celsius. Consider the portion of the plate that lies on the rectangular region $R = [1, 5] \times [3, 6]$.

(a) Estimate the value of $\iint_R T(x, y)\, dA$ by using a double Riemann sum with two subintervals in each direction and choosing (x_i^*, y_j^*) to be the point that lies in the upper right corner of each subrectangle.

(b) Determine the area of the rectangle R.

(c) Estimate the average temperature, $T_{\mathrm{AVG}(R)}$, over the region R.

(d) Do you think your estimate in (c) is an over- or under-estimate of the true temperature? Why?

16. Let f be a function of independent variables x and y that is increasing in both the positive x and y directions on a rectangular domain R. For each of the following situations, determine if the double Riemann sum of f over R is an overestimate or underestimate of the double integral $\iint_R f(x,y)\,dA$, or if it impossible to determine definitively. Provide justification for your responses.

(a) The double Riemann sum of f over R where f is evaluated at the lower left point of each subrectangle.

(b) The double Riemann sum of f over R where f is evaluated at the upper right point of each subrectangle.

(c) The double Riemann sum of f over R where f is evaluated at the midpoint of each subrectangle.

(d) The double Riemann sum of f over R where f is evaluated at the lower right point of each subrectangle.

17. The wind chill, as frequently reported, is a measure of how cold it feels outside when the wind is blowing. In Table 12.2.13, the wind chill $w = w(v,T)$, measured in degrees Fahrenheit, is a function of the wind speed v, measured in miles per hour, and the ambient air temperature T, also measured in degrees Fahrenheit. Approximate the average wind chill on the rectangle $[5,35]\times[-20,20]$ using 3 subintervals in the v direction, 4 subintervals in the T direction, and the point in the lower left corner in each subrectangle.

Table 12.2.13 Wind chill as a function of wind speed and temperature.

$v\backslash T$	-20	-15	-10	-5	0	5	10	15	20
5	-34	-28	-22	-16	-11	-5	1	7	13
10	-41	-35	-28	-22	-16	-10	-4	3	9
15	-45	-39	-32	-26	-19	-13	-7	0	6
20	-48	-42	-35	-29	-22	-15	-9	-2	4
25	-51	-44	-37	-31	-24	-17	-11	-4	3
30	-53	-46	-39	-33	-26	-19	-12	-5	1
35	-55	-48	-41	-34	-27	-21	-14	-7	0

18. Consider the box with a sloped top that is given by the following description: the base is the rectangle $R=[0,4]\times[0,3]$, while the top is given by the plane $z = p(x,y) = 20 - 2x - 3y$.

(a) Estimate the value of $\iint_R p(x,y)\,dA$ by using a double Riemann sum with four subintervals in the x direction and three subintervals in the y direction, and choosing (x_i^*, y_j^*) to be the point that is the midpoint of each subrectangle.

(b) What important quantity does your double Riemann sum in (a) estimate?

(c) Suppose it can be determined that $\iint_R p(x,y)\,dA = 138$. What is the exact average value of p over R?

(d) If you wanted to build a rectangular box (with the same base) that

has the same volume as the box with the sloped top described here, how tall would the rectangular box have to be?

12.3 Iterated Integrals

Motivating Questions

- How do we evaluate a double integral over a rectangle as an iterated integral, and why does this process work?

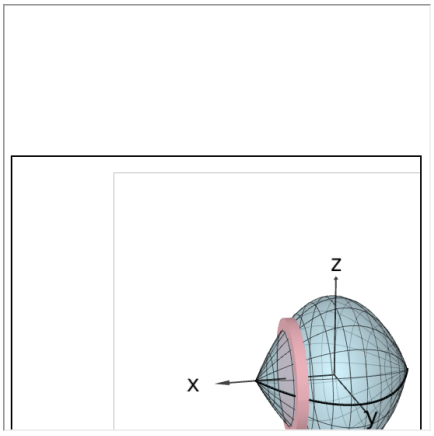
12.3.1 Introduction

In Section 12.2, we defined the double integral of a continuous function $f = f(x, y)$ over a rectangle $R = [a, b] \times [c, d]$ as

$$\iint_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f\left(x_{ij}^*, y_{ij}^*\right) \cdot \Delta A,$$

using the classic calculus approach. Thus $\iint_R f(x, y)\, dA$ is a limit of double Riemann sums. However, while this definition tells us exactly what a double integral *is*, it is not very helpful for computing the *value* of a double integral. Fortunately, there is a way to view a double integral as an **iterated integral**, which will make computations feasible in many cases.

The viewpoint of an iterated integral is closely connected to an important idea from single-variable calculus. When we studied solids of revolution, such as the one shown in Figure 12.3.1, we saw that in some circumstances we could slice the solid perpendicular to an axis and approximate the volume using circular disks or washers. From there, we were able to find the volume of each disk, after which we used an integral to add the volumes of the slices and found the total volume of the rotated solid.



Standalone
In Context
Embed

Figure 12.3.1 A plot of a volume of revolution with circular disk used in estimating the enclosed volume

We will consider an example of this approach in the following Preview Activity. Specifically, you examine how to how use single-variable integrals along traces in a similar way to the rotation volume described above. We will generalize this approach in this section for double integrals over rectangular regions. In Section 12.4, we will look at using this technique on non-rectangular regions of integration.

Preview Activity 12.3.1 In this activity we will explore the double integral of $f(x, y) = 25 - x^2 - y^2$ on the rectangular domain R containing the points that satisfy $-3 \leq x \leq 3$ and $-4 \leq y \leq 4$. As with partial derivatives, we may treat one of the variables in f as constant and think of the resulting function as a function of a single variable. We will now investigate what this means geometrically and algebraically when we integrate with one input variable fixed.

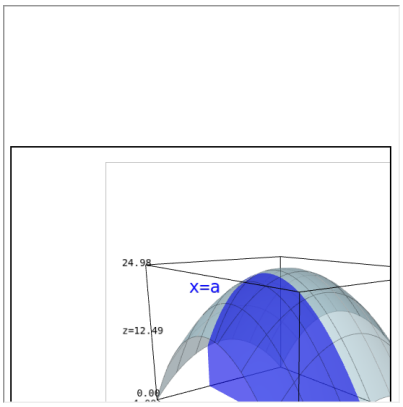
(a) Let a be a fixed value in the interval $[-3, 3]$. Compute

$$\int_{-4}^{4} f(a, y)dy.$$

(b) The answer of the previous part depended on the value of a, but the process would be the exact same set of steps for every a. This allows us to define a function $A(x)$ as

$$A(x) = \int_{-4}^{4} f(x, y)\, dy.$$

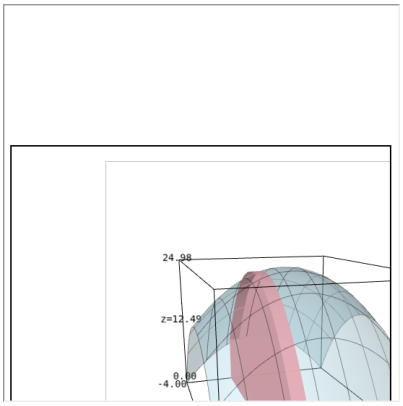
Write a couple of sentences to explain the geometric meaning of the value of $A(x)$. Your explanation should refer to different parts of Figure 12.3.2, including the surface defined by f and the trace determined by the fixed value of x.



Standalone
In Context
Embed

Figure 12.3.2 A plot of the surface $z = 25 - x^2 - y^2$ with a cross section shown at $x = a$

(c) Write a couple of sentences explaining the geometric meaning of $A(x_i^*)\ \Delta x$, where x_i^* is a fixed value of x. Your explanation should refer to the parts of Figure 12.3.3.



Standalone
In Context
Embed

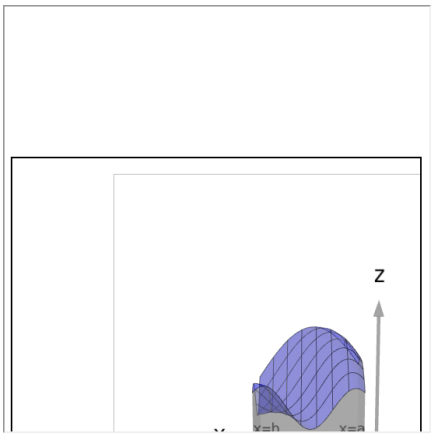
Figure 12.3.3

(d) Since f is continuous on R, we can define the function $A = A(x)$ at every value of x in $[-3, 3]$. Think about subdividing the x-interval $[-3, 3]$ into m subintervals and choosing a value x_i^* in each of those subintervals. Write a sentence describing the meaning of the sum $\sum_{i=1}^{m} A(x_i^*)\,\Delta x$.

(e) Write a couple of sentences to explain why $\int_{-3}^{3} A(x)\,dx$ determines the exact value of the volume under the surface $z = f(x, y)$ over the rectangle R.

12.3.2 Iterated Integrals

The ideas that we explored in Preview Activity 12.3.1 work more generally and lead to the idea of an iterated integral. We will explore how these ideas generalize by considering a surface $z = f(x, y)$ such as the one shown in Figure 12.3.4 over a rectangular region R satisfying $a \le x \le b$ and $c \le x \le d$. Our goal will be to use an interated integral to find the volume beneath this surface and above the xy-plane over the region R.



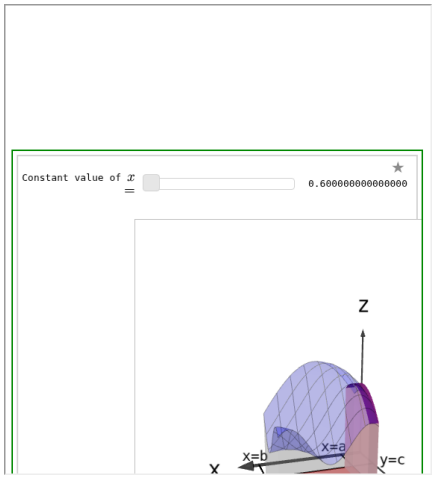
Standalone
In Context
Embed

Figure 12.3.4 A surface and the region below it

To avoid technical issues, we assume that f is continuous. As in the Preview Activity, we define a function

$$A(x) = \int_c^d f(x, y)\,dy.$$

The function $A(x)$ determines the net signed area of a slice of the solid. The slice has a fixed value of x. By adjusting the slider in Figure 12.3.5, you can see a tickened version of such a slice, which we will refer to as a slab. (Thus, a slice will is a *two*-dimensional object and a slab is a *three*-dimensional object.) You should also observe how the *shape* of the slice changes as the value of x changes. This illustrates how it is essential to use an integral to find the area of a cross section, as this area cannot be determined by using only the value of x and simple formulas from geometry.



Standalone
In Context
Embed

Figure 12.3.5 A slab with variable position underneath a surface

If we add up the volumes of the slabs in Figure 12.3.5, then we obtain an approximation for the volume of the solid between the surface and the xy-plane over the region R. If we let Δx denote the thickness of each slab, the volume of the slab with constant x-value x_i^* is $A(x_i^*)\Delta x$. Hence, the sum

$$\sum_{i=1}^{m} A(x_i^*)\Delta x$$

is an approximation of this volume. This sum is a Riemann summ of the single-variable function A. Therefore, as we let the number m of slabs go to infinity we obtain the integral

$$\int_a^b A(x)\,dx$$

as the volume of the solid.

Now remember how we originally defined $A(x)$ by using an integral. Thus, since the double integral represents the (net signed) volume of the solid enclosed by the xy-plane and the surface $z = f(x, y)$ over the region R, we have

$$\iint_R f(x,y)\,dA = \int_a^b A(x)\,dx = \int_a^b \left[\int_c^d f(x,y)\,dy\right] dx = \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

We call these nested integrals an **iterated integral**, which will prove to be a frequently-used tool for evaluating double integrals.

In our future discussions of double integrals, it will be helpful to recognize the roles of the different components of an interated integral. When we consider

$$\int_a^b \int_c^d f(x,y)\,dy\,dx,$$

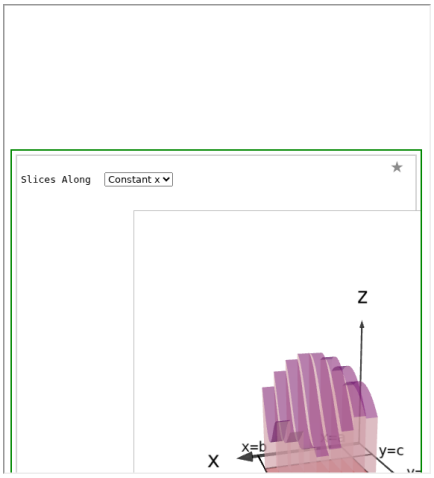
the differentials dy and dx carry both geometric and computational meaning. From a computational perspective, that dy is the *inner* differential tells us that the limits of integration on the inner integral sign are values of y: $c \le y \le d$ for this rectangular region. Similarly, the position of dx as the *outer* differential tells us that when computing the inner integral, we consider x to be constant. After that computation is complete, we then have a single-variable integral on the interval with $a \le x \le b$. Thinking geometrically, since the inner integral computes the area of a cross section of the solid, we can think of dy as being the width of the rectangles whose areas are summed. Recalling that the dx

arose from the limit where we had previously used Δx to find the volume of a slab, we see that the outer integral can be considered as summing up the volumes of slabs of thickness dx.

In our arguments above, we made the choice to integrate first with respect to y (holding x constant). We could just as easily have chosen to hold y constant first, which yields the double integral as an iterated integral that integrates first with respect to x:

$$\iint_R f(x,y)\,dA = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy = \int_c^d \int_a^b f(x,y)\,dx\,dy.$$

Figure 12.3.6 allows you to change the direction of slicing and provides a visualization of the slabs whose volumes are integrated by the outer integral in this iterated integral.



Standalone
In Context
Embed

Figure 12.3.6 Decomposition of the solid below a surface into slices along constant x-values or y-values

The fact that integrating in either order results in the same value is known as Fubini's Theorem.

Fubini's Theorem.

If $f = f(x,y)$ is a continuous function on a rectangle $R = [a,b] \times [c,d]$, then

$$\iint_R f(x,y)\,dA = \int_c^d \int_a^b f(x,y)\,dx\,dy = \int_a^b \int_c^d f(x,y)\,dy\,dx.$$

Fubini's theorem enables us to evaluate double integrals by using iterated integrals instead of resorting to the limit definition. Working with one integral at a time, we can use the Fundamental Theorem of Calculus from single-variable calculus to find the exact value of each integral, starting with the inner integral.

This is another time in which we trade a multi-dimensional problem (double integral) for multiple one-dimensional problems that we have much more experience with.

Activity 12.3.2 Let $f(x, y) = 25 - x^2 - y^2$ on the rectangular domain R with $-3 \leq x \leq 3$ and $-4 \leq y \leq 4$.

(a) Viewing x as a fixed constant, use the Fundamental Theorem of Calculus to evaluate the integral

$$A(x) = \int_{-4}^{4} f(x, y)\, dy.$$

Note that you will be integrating with respect to y, and holding x constant. Your result should be a function of x only.

(b) Next, use your result from (a) along with the Fundamental Theorem of Calculus to determine the value of $\int_{-3}^{3} A(x)\, dx$.

(c) What is the value of $\iint_R f(x, y)\, dA$? Write a sentence to interpret the meaning of this value for two out of the three different ways mentioned in Key Idea 12.2.8.

Activity 12.3.3 Let $f(x, y) = x + y^2$ on the rectangle R with $0 \leq x \leq 2$ and $0 \leq y \leq 3$.

(a) Evaluate $\iint_R f(x, y)\, dA$ using an iterated integral. Choose an order for integration by deciding whether you want to integrate first with respect to x or y.

(b) Evaluate $\iint_R f(x, y)\, dA$ using the iterated integral whose order of integration is the opposite of the order you chose in (a).

Summary

- We can evaluate the double integral $\iint_R f(x, y)\, dA$ over a rectangle $R = [a, b] \times [c, d]$ as an iterated integral in one of two ways:
  - $\displaystyle\int_a^b \left[\int_c^d f(x, y)\, dy \right] dx$, or
  - $\displaystyle\int_c^d \left[\int_a^b f(x, y)\, dx \right] dy$.

This process works because each inner integral represents a cross-sectional (signed) area and the outer integral then sums all of the cross-sectional (signed) areas. Fubini's Theorem guarantees that the resulting value is the same, regardless of the order in which we integrate.

12.3.3 Exercises

1. Evaluate the iterated integral $\int_0^2 \int_0^1 3x^2y^3\,dxdy$

2. Evaluate the iterated integral $\int_2^3 \int_4^5 (2x+y)^{-2}\,dydx$

3. Find $\int_3^6 \int_8^{12} (x+\ln y)\,dydx$

4. Find $\int_4^5 \int_1^3 xye^{x+y}\,dydx$

5. Calculate the double integral $\int\int_{\mathbf{R}} (10x+8y+80)\,dA$ where $\mathbf{R}$ is the region: $0 \le x \le 4, 0 \le y \le 5$.

6. Calculate the double integral $\int\int_{\mathbf{R}} x\cos(2x+y)\,dA$ where $\mathbf{R}$ is the region: $0 \le x \le \frac{\pi}{6}, 0 \le y \le \frac{\pi}{4}$

7. Evaluate the iterated integral $\int_0^1 \int_0^1 6x^2y^3\,dxdy$

8. Evaluate the iterated integral $\int_1^2 \int_3^4 (4x+y)^{-2}\,dydx$

9. Calculate the double integral $\int\int_{\mathbf{R}} (4x+4y+16)\,dA$ where $\mathbf{R}$ is the region: $0 \le x \le 2, 0 \le y \le 2$.

10. Consider the solid that lies above the square (in the xy-plane) $R = [0,2] \times [0,2]$,

and below the elliptic paraboloid $z = 49 - x^2 - y^2$.

(A) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the lower left hand corners.

(B) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the upper right hand corners..

(C) What is the average of the two answers from (A) and (B)?

(D) Using iterated integrals, compute the exact value of the volume.

11. Evaluate $\displaystyle\int_0^1 \int_1^2 \left(8x^3 - 27x^2y^2\right) dy\,dx.$

Answer: _____

12. Evaluate $\displaystyle\int_1^5 \int_1^4 \frac{\ln y}{xy} dy\,dx.$

Answer: _____

13. Evaluate $\displaystyle\int_0^4 \int_0^5 xy\sqrt{x^2+y^2}\,dy\,dx.$

Answer: _____

14. Calculate $\displaystyle\iint_R ye^{-xy}\,dA$, where $R = [0,7] \times [0,4]$.

Answer: _____

15. Find the volume of the solid that lies under the hyperbolic paraboloid $z = 3y^2 - x^2 + 6$ and above the rectangle $R = [-1,1] \times [1,2]$.

Answer: _____

16. Find the average value of $f(x,y) = x^2y$ over the rectangle R with vertices $(-4,0), (-4,8), (2,8), (2,0)$.

Answer: _____

17. Evaluate the iterated integral:

$$\int_0^2 \int_6^7 \sqrt{x+4y}\, dx\, dy$$

Answer : _____

18. Evaluate the double integral of the function over the rectangle :

$$\iint_{\mathcal{R}} \frac{y}{x+1}\, dA, \quad \mathcal{R} = [0,8] \times [0,8]$$

Answer : _____

19. Evaluate the integral:

$$\iint_{\mathcal{R}} e^{3x+4y}\, dA, \quad \mathcal{R} = [0,4] \times [0,9]$$

Answer : _____

20. Consider the solid that lies above the square (in the xy-plane) $R = [0,2] \times [0,2]$,

and below the elliptic paraboloid $z = 49 - x^2 - 3y^2$.

(A) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the lower left hand corners.

(B) Estimate the volume by dividing R into 4 equal squares and choosing the sample points to lie in the upper right hand corners..

(C) What is the average of the two answers from (A) and (B)?

(D) Using iterated integrals, compute the exact value of the volume.

21. If $\int_1^2 f(x)\ dx = -5$ and $\int_{-2}^1 g(x)\ dx = 4$, what is the value of $\iint_D f(x)g(y)\, dA$ where D is the rectangle: $1 \le x \le 2, \quad -2 \le y \le 1$?

22. Find the average value of $f(x,y) = 4x^6y^4$ over the rectangle R with vertices $(-6,0), (-6,2), (6,0), (6,2)$.

Average value = _____

23. Find the average value of $f(x,y) = 7e^y\sqrt{x+e^y}$ over the rectangle $R = [0,2] \times [0,5]$.

Average value = _____

24. Evaluate each of the following double or iterated integrals exactly.

(a) $\int_1^3 \left(\int_2^5 xy\, dy\right)\, dx$

(b) $\int_0^{\pi/4} \left(\int_0^{\pi/3} \sin(x)\cos(y)\, dx\right)\, dy$

(c) $\int_0^1 \left(\int_0^1 e^{-2x-3y}\, dy\right)\, dx$

(d) $\iint_R \sqrt{2x+5y}\, dA$, where $R = [0,2] \times [0,3]$.

25. The temperature at any point on a metal plate in the xy plane is given by $T(x,y) = 100 - 4x^2 - y^2$, where x and y are measured in inches and T in degrees Celsius. Consider the portion of the plate that lies on the rectangular region $R = [1,5] \times [3,6]$.

(a) Write an iterated integral whose value represents the volume under the surface T over the rectangle R.

(b) Evaluate the iterated integral you determined in (a).

(c) Find the area of the rectangle, R.

(d) Determine the exact average temperature, $T_{\text{AVG}(R)}$, over the region R.

26. Consider the box with a sloped top that is given by the following description: the base is the rectangle $R = [1, 4] \times [2, 5]$, while the top is given by the plane $z = p(x, y) = 30 - x - 2y$.

(a) Write an iterated integral whose value represents the volume under p over the rectangle R.

(b) Evaluate the iterated integral you determined in (a).

(c) What is the exact average value of p over R?

(d) If you wanted to build a rectangular box (with an identical base) that has the same volume as the box with the sloped top described here, how tall would the rectangular box have to be?

12.4 Double Integrals over General Regions

Motivating Questions

- How do we define a double integral over a non-rectangular region?
- What general forms do iterated integrals over non-rectangular regions have?

12.4.1 Introduction

Recall from Section 12.2 that we defined the double integral of a continuous function $f = f(x, y)$ to measure the volume of the region beneath the graph of f over a region of input points. If we consider the set of points in a rectangle R with $a \le x \le b$ and $c \le y \le d$, then we can calculate the double integral using a Riemann sum

$$\iint_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \cdot \Delta A.$$

Furthermore, we have seen in Fubini's Theorem that we can evaluate a double integral $\iint_R f(x, y)\, dA$ over R as an iterated integral of either of the forms

$$\int_a^b \int_c^d f(x, y)\, dy\, dx \qquad \text{or} \qquad \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

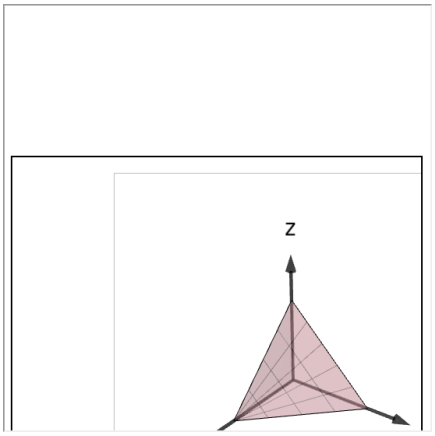
Most applied or theoretical problems involve regions of inputs that are *not* rectangles. Therefore, we want to extend the idea of using iterated integrals to allow us to evaluate double integrals over nonrectangular regions. We explore one such example in the following preview activity.

Preview Activity 12.4.1 A tetrahedron is a three-dimensional figure with four faces, each of which is a triangle. A picture of the tetrahedron T with

vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ is shown in Figure 12.4.1. By using techniques from Section 9.7, we can find that an equation for this plane is

$$z = 1 - (x + y).$$

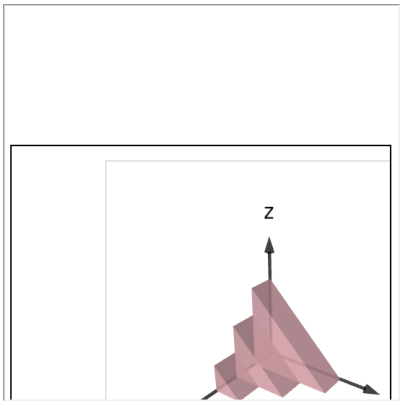
However, we cannot directly use the methods we have developed so far involving double integrals to find the volume of a solid above the xy-plane and below a surface $z = f(x, y)$, as the base of the tetrahedron is not rectangular. This Preview Activity generalizes the work we have done so far with double integrals to understand how we can use a double integral to compute the volume of this tetrahedron. To do this, we will apply the Classic Calculus Approach.



Standalone
In Context
Embed

Figure 12.4.1 The tetrahedron T with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$

(a) As our first step in applying the Classic Calculus Approach, we approximate the tetrahedron by using three triangular prisms with bases parallel to the yz-plane. Specifically, the bases lie in the planes $x = 0$, $x = 1/3$, and $x = 2/3$. We show these three triangular prisms in Figure 12.4.2.



Standalone
In Context
Embed

Figure 12.4.2 A decomposition of the tetrahedron into three triangular prisms with base parallel to the yz-plane

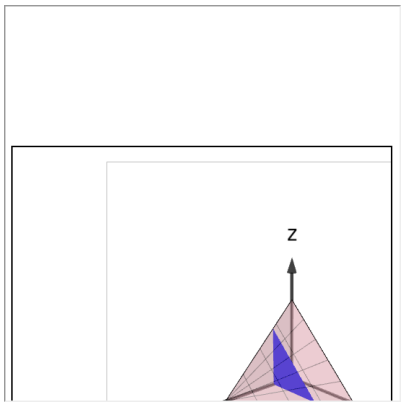
The volume of each triangular prism is the cross sectional area of the prism's base times the thickness of the prism. Find the cross sectional area for the three prisms shown in the figure.

(b) Write a sum in the form below that completes Step 1 of the Classic

Calculus Approach.

$$\sum_{i=1}^{3} (\text{cross sectional area}) \cdot (\text{prism width})$$

(c) For Step 2 of the Classic Calculus Approach, we generalize the approximation above to work for n triangular prisms of equal width. To do this, we must find the cross sectional area at an arbitrary location $x = a$. Create a carefully-labeled two-dimensional drawing of the highlighted cross section in Figure 12.4.3 that can be interpreted for any value of a where the plane $x = a$ intersects the tetrahedron. Your drawing should be in the yz-plane, and you should label components of your diagram in terms of the parameter a. Make sure to write an equation for the line that forms the hypotenuse of the triangle, which should be expressed in terms of y, z, and a.



Standalone
In Context
Embed

Figure 12.4.3 The tetrahedron T with cross section along $x = a$ shown

(d) Write a dy integral for the area of the cross section you drew in the previous part. Your integral will include a in multiple ways. Verify that your integral gives the same value you found earlier for the $a = 1/3$ cross sectional area.

(e) To complete Step 2 of the Classic Calculus Approach, we approximate the volume by a sum. Specifically, if we denote the cross sectional area of the ith triangular prism by CSA_i, then the volume is approximated by

$$\sum_{i=1}^{n} \text{CSA}_i \cdot \Delta x.$$

By the Classic Calculus approach, the volume of the tetrahedron is

$$\lim_{n\to\infty} \sum_{i=1}^{n} \text{CSA}_i \cdot \Delta x = \int_0^1 \text{CSA(x)}\, dx, \tag{12.4.1}$$

where $\text{CSA}(x)$ denotes the cross sectional area for position x.

In the previous part of the preview activity, you wrote a dy-integral for the cross sectional area at position $x = a$. Replace all occurrences of a in that integral by x and substitute the result integral into the integral in (12.4.1). This produces an iterated integral that can be used to evaluate $\iint_R 1 - (x + y)\, dA$, where R is the base of the tetrahedron.

(f) Compute $\int_0^{1-x} 1-(x+y)\,dy$. Remember that to do this, you treat x as a constant, much like you would when computing a partial derivative, to find an antiderivative of $1-(x+y)$ *with respect to* y.

(g) Write your answer from the previous part in the blank in in the following integral and then evaluate the integral. Your answer will be a constant.

$$\int_0^1 \qquad\qquad\quad dx$$

In this section, we will explore the ideas introduced by the Preview Activity in greater generality. We begin by understanding how to describe regions in the plane, as this will be essential to writing iterated integrals that can be used to evaluate double integrals.

12.4.2 Double Integrals over General Regions

So far, we have learned that a double integral over a rectangular region may be interpreted in three ways detailed in Key Idea 12.2.8: net signed volume, mass, and when divided by the area of the base region, average value.

As we saw in Preview Activity 12.2.1, a function $f = f(x, y)$ may be considered over regions other than rectangular ones, and thus we want to understand how to set up double integrals as iterated integrals over non-rectangular regions. Note that if we can set up double integrals over non-rectangular regions as iterated integrals, then the three interpretations of the double integral noted above will extend to solid regions with non-rectangular bases.

Remember that the double integral over a a region can be thought of as the signed volume of the solid bounded by the graph of f and the xy-plane. The volume above the xy-plane is counted as positive and the volume below the xy-plane is counted as negative. As shown in Figure 12.2.10, the total volume above the xy-plane is greater than the total volume below the xy-plane, so the double integral of f over the region R is positive. This argument comes from both the heights of the rectangular prisms used in the Riemann sum (determined by output values of f) and the area of the regions where the output of f is positive versus negative.

In the next activity, you are asked to assess and justify your thoughts about whether the double integrals over non-rectangular regions are positive, negative, or zero. Your justifications should consider both the area of the given region of integration, as well as the output values of the function being integrated on these regions.

Activity 12.4.2 Let D be the region inside the unit circle centered at the origin, let R be the right half of D, and let B be the bottom half of D.

(a) On three separate plots, graph and label the regions D, R, and B.

(b) For each double integral below, decide *without calculation* whether the double integral is positive, negative, or zero. Write a sentence or two to explain your answer for each part.

i. $\iint_D 1\,dA$

ii. $\iint_D x\,dA$

iii. $\iint_B x\,dA$

iv. $\iint_R x\,dA$

v. $\iint_D x^2\,dA$

vi. $\iint_B 2-x^2\,dA$

vii. $\iint_R -e^x\,dA$

viii. $\iint_B 1-e^y\,dA$

ix. $\iint_D xy\,dA$

12.4.3 Horizontally and Vertically Simple Regions

The key idea to calculating double integrals using iterated integrals as shown in Section 12.3 and Preview Activity 12.4.1 is to slice the region of integration along a trace where one of the coordinates is held constant. We then express the range of the other coordinate in terms of an inequality. In this subsection, we will step away from double integrals for a short while to practice slicing two dimensional regions and stating corresponding inequalities in terms of each coordinate.

In Preview Activity 12.4.1, we considered a tetrahedron, which had as its base a triangular region R in the xy-plane. This region is shown in Figure 12.4.4.

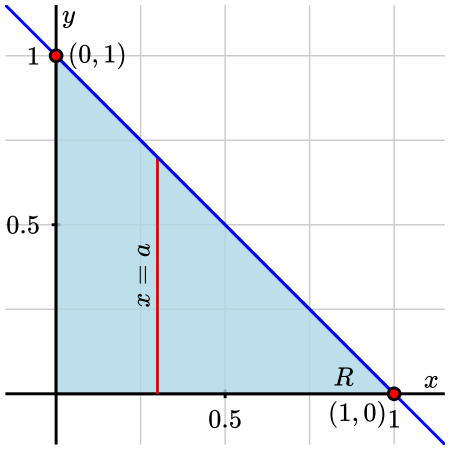


Figure 12.4.4 The projection of T onto the xy-plane

Notice how different values of a provide different ranges of y-values for the portion of the $x = a$ line segment that lies inside R. This is intrinsically linked with the fact that R is not a rectangular region, in contrast to the regions of integration from Section 12.3 which were all of the form $a \leq x \leq b$ and $c \leq y \leq d$ (with a, b, c, d constants).

These forms for the inequalities lead to a nice way to write the corresponding double integral as an iterated integral by considering the inner integral to evaluate the cross section area and the outer integral to sum the corresponding cross section volumes as the thickness of the slabs goes to zero.

Not all two-dimensional regions can be immediately expressed using this idea of converting geometric slices to inequalities. The key characteristic that allows us to do this is a region must have the same upper and lower bound expressions for each slice.

Key Idea 12.4.5 *A region is called* ***vertically simple*** *if it can be described by a set of inequalities of the form*

$$a \leq x \leq b$$
$$g_1(x) \leq y \leq g_2(x).$$

These inequalities correspond geometrically to the region in Figure 12.4.6.

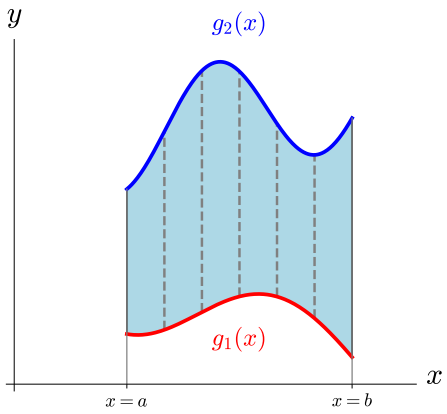


Figure 12.4.6 *A vertically simple region*

A region is called ***horizontally simple*** *if it can be described by a set of inequalities of the form*

$$c \le y \le d$$
$$h_1(y) \le x \le h_2(y)$$

which corresponds geometrically to the region in Figure 12.4.7.

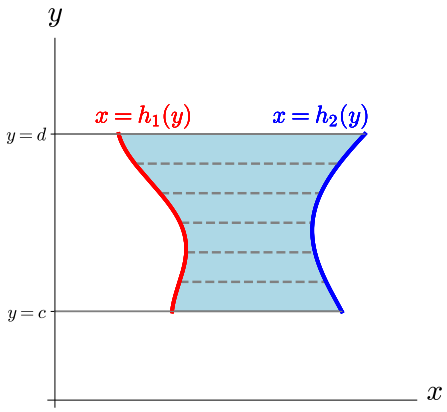


Figure 12.4.7 *A horizontally simple region*

A vertically simple region is one that can be split into vertical slices of the form $x =$ constant for x in the interval $[a, b]$ where every vertical slice has the same upper and lower bound function: $g_1(x) \le y \le g_2(x)$. A horizontally simple region is one that can be split into horizontal slices of the form $y =$ constant for y in the interval $[c, d]$ where every horizontal slice has the same upper (right) and lower (left) bound functions: $h_1(x) \le y \le h_2(x)$.

Example 12.4.8 Let D be the triangular region in the xy-plane with vertices $(0, 0)$, $(2, 0)$, and $(2, 3)$, shown in Figure 12.4.9. We will determine if D is horizontally simple, vertically simple, or both. For simplicity, we refer to the three sides as follows:

- the horizontal side from $(0, 0)$ to $(2, 0)$,
- the vertical side from $(2, 0)$ to $(2, 3)$, and
- the diagonal side from $(0, 0)$ to $(2, 3)$.

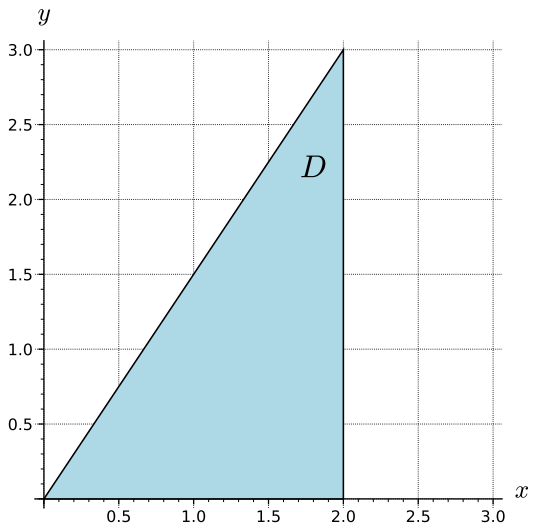


Figure 12.4.9 The triangular region D with vertices $(0,0)$, $(2,0)$, and $(2,3)$

(a) A vertical line has equation $x =$ constant. We must determine if every vertical slice in the region D has the same upper and lower bounds, which may depend on the value of x. Geometrically, we are asking if the region D has the same functions as the top and bottom of the region for each value of x.

Several vertical slices in the region D are shown in Figure 12.4.10 where you can see that the lower bound of the y-value for every vertical slice is given by $y = 0$. The upper bound of the y-value for every vertical slice is given by $y = \frac{3}{2}x$. Algebraically, for any constant value k in the interval $[0,2]$, the following statement describes the vertical slices of the region D:

$$\text{Along } x = k \text{ the } y\text{-coordinates go from } 0 \text{ to } \frac{3}{2}k$$

which corresponds to the inequalities

$$0 \le x \le 2 \qquad 0 \le y \le \frac{3}{2}x$$

as a description for D. Thus, the region D is vertically simple.

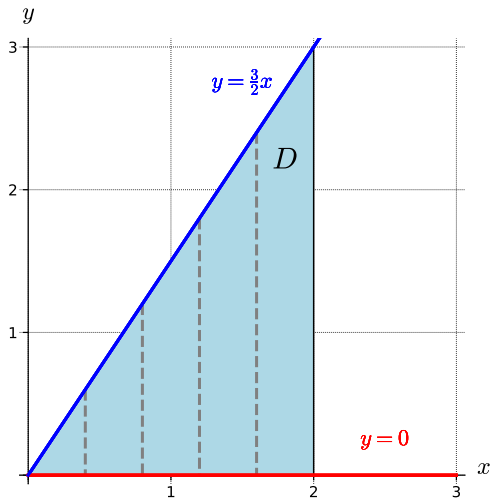


Figure 12.4.10 The triangular region D with vertical slices shown

(b) A horizontal line has equation $y = k$ where k is a constant. We must determine if every horizontal slice in the region D has the same upper and lower bounds, which may depend on the value of y. Geometrically, this corresponds to having the same functions of y as the right and left boundaries of the region.

Several horizontal slices in the region D are shown in Figure 12.4.11 where you can see that the lower bound of the x-value for every horizontal slice is given by $x = \frac{2}{3}y$. The upper bound of the x-value for every horizontal slice is given by $x = 2$. Algebraically, for any constant value k in the interval $[0, 3]$, the following statement describes the horizontal slices of the region D:

$$\text{Along } y = k \text{ the } x\text{-coordinates go from } \frac{2}{3}k \text{ to } 2$$

which corresponds to the inequalities

$$0 \leq y \leq 3 \qquad \frac{2}{3} \leq y \leq 2x$$

as a description for D. Thus, the region D is horizontally simple.

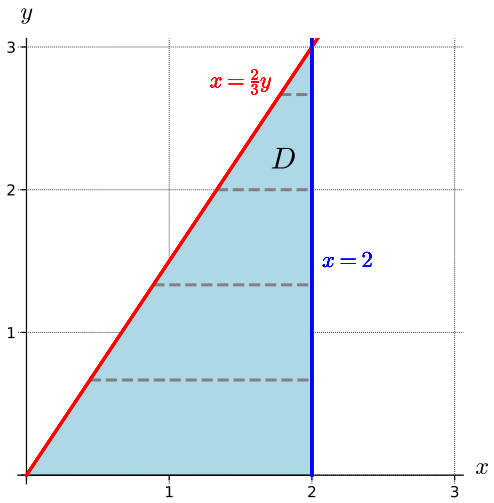


Figure 12.4.11 The triangular region D with horizontal slices shown

This example shows that just because a region is vertically simple does not mean that it is *not* horizontally simple as well. In general, a two-dimensional region can be any of the following

- horizontally simple (but not vertically simple),
- vertically simple (but not horizontally simple),
- both horizontally simple and vertically simple, or
- neither horizontally simple or vertically simple.

As we will see later, switching the description of a region that is both horizontally simple and vertically simple can make the algebra of evaluating an iterated integral easier. ▲

Example 12.4.12 In this example, we will look at region that is neither horizontally or vertically simple. Let D be the parallelogram with vertices $(0,0)$, $(2,1)$, $(4,3)$, and $(2,2)$ as shown in Figure 12.4.13. You should take a moment to verify each of the equations for the lines that bound D.

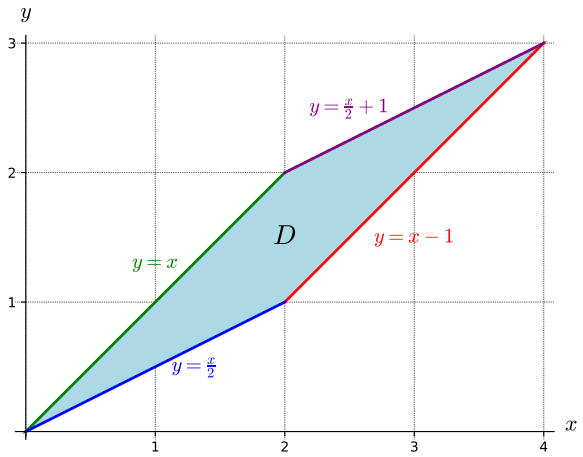
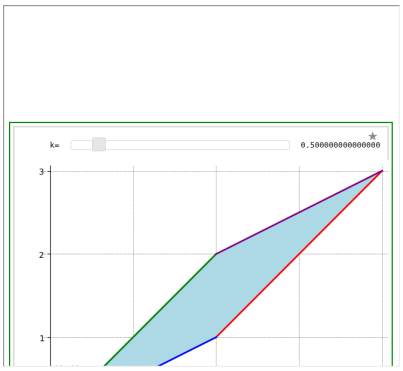


Figure 12.4.13 The region D with vertices $(0,0)$, $(2,1)$, $(4,3)$, and $(2,2)$

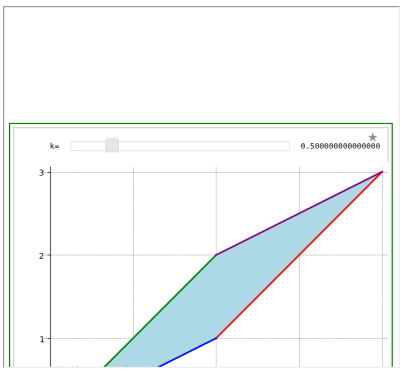
(a) If we try to look at D along vertical slices of the form $x = k$, different values of k will have different upper and lower bounds. For instance, the default k value of 0.5 shown in Figure 12.4.14 is a vertical segment with lower bound at the point $\left(k, \frac{k}{2}\right)$ and upper bound at the point (k, k). Use the slider at the top of Figure 12.4.14 to change the value of k to a number greater than 2. For such values, the vertical slice of D along $x = k$ will have lower bound $(k, k-1)$ and upper bound $\left(k, \frac{k}{2}+1\right)$. Because the upper or lower bound *functions* depend on the location of a vertical slice, D is not vertically simple.



Standalone
In Context
Embed

Figure 12.4.14 The region D with vertical slices that change upper and lower bounds depending on the value of $x = k$

(b) If we try to look at D along horizontal slices of the form $y = k$, different values of k will have different upper and lower bounds. For instance, the default k value of 0.5 shown in Figure 12.4.15 is a vertical segment with lower bound at the point (k, k) and upper bound at the point $(2k, k)$. Use the slider at the top of Figure 12.4.15 to change the value of k to a number greater than 1. You can see that for k values between 1 and 2, the upper bound on the horizontal segment becomes $(k-1, k)$ and for k values above 2, the lower bound also changes to $\left(\frac{k}{2}+1, k\right)$. Because the upper or lower bound *functions* depend on the location of the horizontal slice, D is not horizontally simple.



Standalone
In Context
Embed

Figure 12.4.15 The region D with horizontal slices that change upper and lower bounds depending on the value of $y = k$

▲

Activity 12.4.3 In this activity, we will break the region D from Example 12.4.12 into smaller regions which are horizontally or vertically simple.

(a) In Figure 12.4.16, the region D is broken into two regions, V_1 and V_2.

Give inequalities for each of V_1 and V_2 that shows each region is vertically simple.

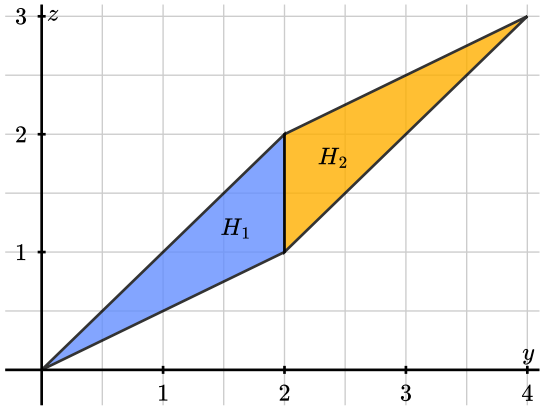


Figure 12.4.16 The region D split into two regions, V_1 and V_2, which are individually vertically simple

(b) In Figure 12.4.17, the region D is broken into three regions H_1, H_2, and H_3. Give inequalities for each of H_1, H_2, and H_3 that shows that each region is horizontally simple.

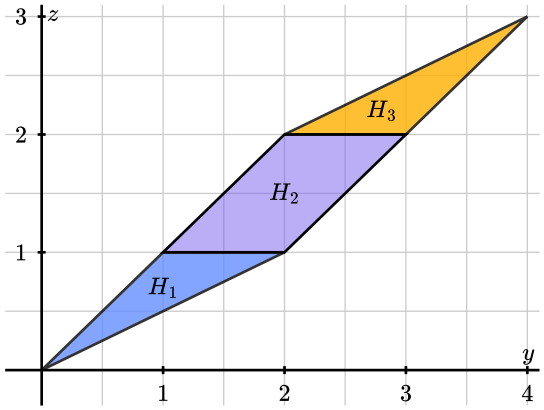


Figure 12.4.17 The region D split into three regions, H_1, H_2, and H_3, which are individually horizontally simple

In single variable calculus, we were able to split integrals into multiple parts. The integral of $f(x)$ over $[a, b]$ could be split to subintervals separately:

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

Similarly, we can split a region of integration for a double integral into subregions. If a region R can be split into two nonoverlapping regions R_1 and R_2, then

$$\iint_R f(x,y)\,dA = \iint_{R_1} f(x,y)\,dA + \iint_{R_2} f(x,y)\,dA.$$

This property is useful for a region similar to Figure 12.4.17 where D is split into parts H_1, H_2, and H_3 because

$$\iint_D f(x,y)\,dA = \iint_{H_1} f(x,y)\,dA + \iint_{H_2} f(x,y)\,dA + \iint_{H_3} f(x,y)\,dA,$$

where each of the double integrals on the right can be written as an iterated integral using a horizontally simple description.

The process for describing a two-dimensional region as horizontally or vertically simple is exactly the procedure needed to set up an iterated integral to evaluate a double integral.

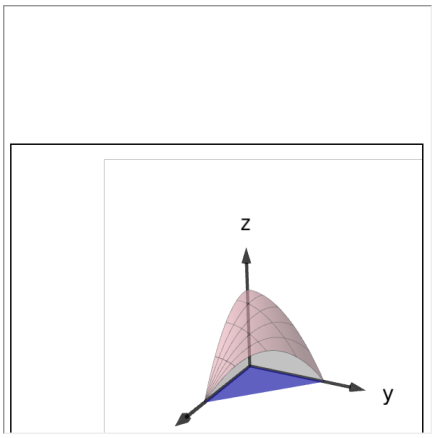
Key Idea 12.4.18 *If a vertically simple region D can be described by the inequalities $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where $g_1(x)$ and $g_2(x)$ are functions of only x, then the double integral can be written as an iterated integral in the following way*

$$\iint_D f(x,y)\,dA = \int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} f(x,y)\,dy\,dx.$$

Alternatively, if a horizontally simple region D is described by the inequalities $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where $h_1(y)$ and $h_2(y)$ are functions of only y, then the double integral can be written as an iterated integral in the following way

$$\iint_D f(x,y)\,dA = \int_{y=c}^{y=d} \int_{x=h_1(y)}^{x=h_2(y)} f(x,y)\,dx\,dy.$$

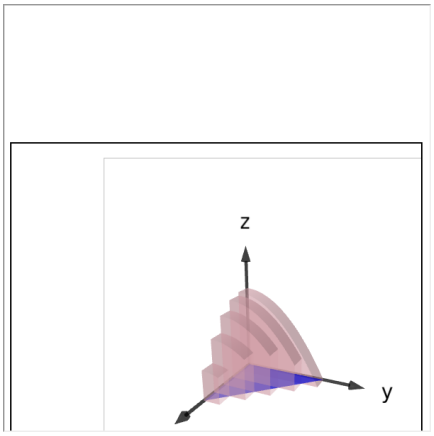
Proof. We will make a brief argument using the Classic Calculus Approach to give an outline for a proof of this result. We will visualize the double integral with a region similar to the one used in the Preview Activity, as shown in Figure 12.4.19. Remember that we will have a region of integration given by the inequalities $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where $g_1(x)$ and $g_2(x)$ are functions of only x



Standalone
In Context
Embed

Figure 12.4.19 The double integral $\iint_R f(x,y)dA$ visualized as the volume between the xy-plane and the surface given by $z = f(x,y)$, shown in pink, over the region R, shown in blue

For Step 1 of the Classic Calculus Approach, we approximate this double integral with n prisms of equal width along constant values of x as shown in Figure 12.4.20. The width of these n equal width prisms will be $\Delta x = \frac{b-a}{n}$. These prisms will involve the cross sectional area of our volume at the following values of x: $x_0 = a, x_1 = a + \Delta x, \ldots, x_i = a + i\Delta x, \ldots, x_{n-1} = a + (i-1)\Delta x$.



Standalone
In Context
Embed

Figure 12.4.20 The double integral $\iint_R f(x,y)dA$ approximated by prisms along constant values of x

The cross sectional area of the volume at $x = a$ is given by

$$CSA(a) = \int_{y=g_1(a)}^{y=g_2(a)} f(a,y)dy.$$

Note that this takes the height of the surface on this slice and integrates over only the relevant range of y-values. Our approximation for Step 1 of the Classic Calculus Approach is given by

$$\iint_R f(x,y)dA \approx \sum_{n=0}^{n-1} \text{CSA}(a)\Delta x$$

Note that this approximation already accomplishes Step 2 of the Classic Calculus Approach since a smaller scale means that we have an increase in n, the number of cross section prisms, which decreases Δx, the width of the prisms.

Applying a limit in Step 3 of our approximation yields

$$\iint_R f(x,y)dA = \lim_{n\to\infty} \sum_{n=0}^{n-1} \text{CSA}(x_i)\Delta x,$$

which you may recognize as a definite integral. In particular,

$$\iint_R f(x,y)dA = \int_{x=a}^{x=b} \text{CSA}(x)dx.$$

We can substitute this in our expression for CSA(x) to get

$$\iint_R f(x,y)dA = \int_{x=a}^{x=b} \left[\int_{y=g_1(x)}^{y=g_2(x)} f(x,y)dy \right] dx.$$

Note that when we substituted in our expression for the cross sectional area integral, we used the variable x since that corresponds to the value/variable used in the larger expression. The other result with prisms along constant values of y has a very analogous argument but does not offer any additional illumination, so we omit this discussion. ■

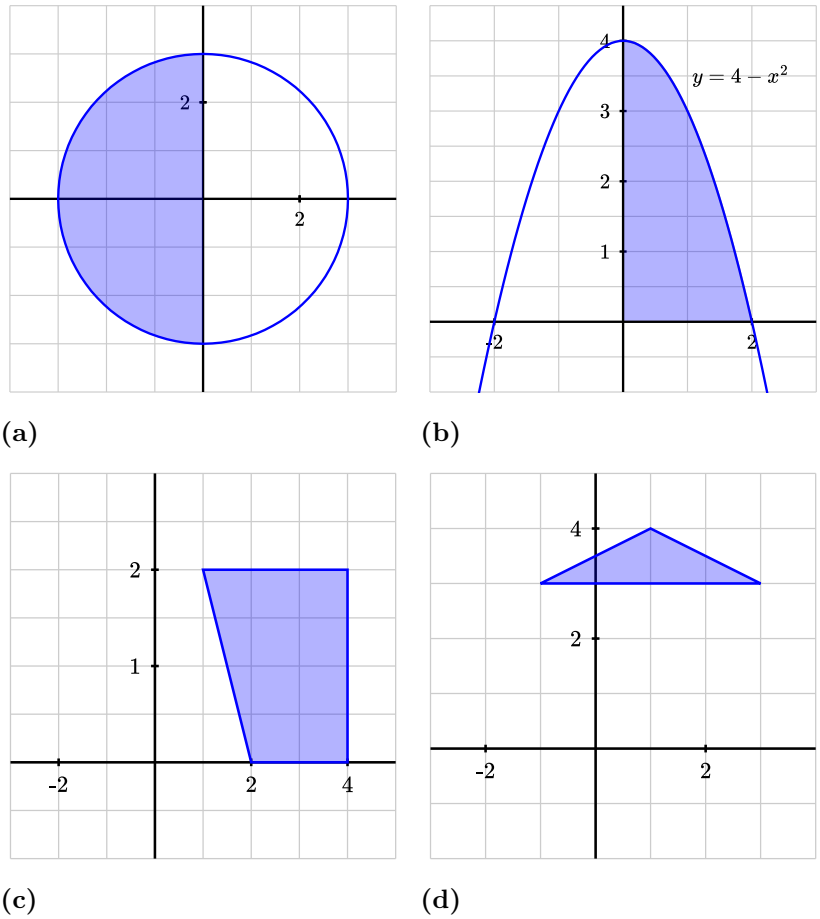
The structure of an iterated integral that will be used to evaluate a double integral must satisfy the following:

- The limits on the inner integral must be constants or expressed only in terms of the outer variable—that is, if the inner integral is with respect to y, then its limits may only involve x and constants.
- The limits on the outer integral must be constants.

The next activity gives you the opportunity to practice describing regions as horizontally simple, vertically simple, both, or neither. It then presents some iterated integrals and asks you to determine if they meet the requirements above to correspond to a double integral.

Activity 12.4.4

(a) For each region shown below, state whether the region is vertically simple, horizontally simple, both, or neither. State the appropriate inequalities to justify when a region is vertically simple or horizontally simple.



(a) **(b)**

(c) **(d)**

Figure 12.4.21 Four Regions of the xy-plane

(b) For each of the interated integrals below, write out the bounds of the integral as a set of inequalities. Use these inequalities to decide if the bounds of the iterated integrals make sense to compute a double integral. For any iterated integral with bounds that do not make sense as a double integral, explain your reasoning in a couple of sentences. For any bounds of an iterated integral that does make sense as a double integral, sketch a plot of the associated region of integration.

(i) $\displaystyle\int_0^1 \int_1^x f(x,y)\,dy\,dx$

(ii) $\displaystyle\int_0^1 \int_1^y f(x,y)\,dy\,dx$

(iii) $\displaystyle\int_0^1 \int_x^y f(x,y)\,dy\,dx$

(iv) $\displaystyle\int_0^1 \int_x^1 f(x,y)\,dy\,dx$

12.4.4 Double Integrals Evaluated as Iterated Integrals

We now turn our attention to using iterated integrals to evaluate double integrals. This is a powerful algebraic tool that we will use throughout the remainder of this chapter, but as the second part of Activity 12.4.4 showed, not every iterated integral that can be written down is meaningful in the sense of evaluating double integrals. As you read the rest of this section, pay careful attention to understanding the meaning behind the computations.

Example 12.4.22 Let $f(x,y) = x^2y$ be defined on the triangle D with vertices $(0,0)$, $(2,0)$, and $(2,3)$ as shown in Figure 12.4.23. This is the same region we described in Example 12.4.8. In this example, we will use the inequalities from that example that described the region as vertically simple to set up and evaluate an iterated integral that gives the value of the double integral.

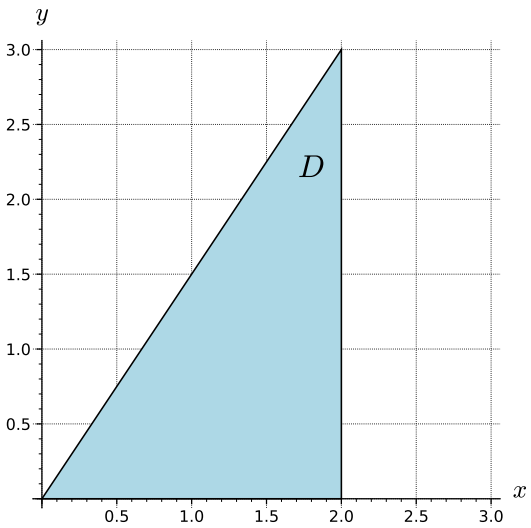


Figure 12.4.23 The triangular region D with vertices $(0,0)$, $(2,0)$, and $(2,3)$

(a) We use the inequalities from Example 12.4.8 that correspond to our vertically simple description:

$$\text{Along } x = k, \text{ the } y\text{-coordinates go from 0 to } \frac{3}{2}k,$$

which corresponds to the inequalities

$$0 \le x \le 2 \qquad 0 \le y \le \frac{3}{2}x.$$

Using these inequalities corresponds to integrating with respect to y first. That is, the variable in the inner integral is y.

Therefore,

$$\iint_D x^2 y \, dA = \int_{x=0}^{x=2} \int_{y=0}^{y=\frac{3}{2}x} x^2 y \, dy \, dx.$$

We evaluate the iterated integral by applying the Fundamental Theorem of Calculus first to the inner integral followed by the outer one. To do this, let's go carefully step-by-step and first evaluate the inner integral on its own. To do this, remember that, much like with computing partial derivatives, we treat x as a constant since we are integrating with respect to y. Therefore, we find an antiderivative *with respect to* y for the integrand proceed using the Fundamental Theorem of Calculus as follows:

$$\int_{y=0}^{y=\frac{3}{2}x} x^2 y \, dy = \left[x^2 \cdot \frac{y^2}{2} \right] \Big|_{y=0}^{y=\frac{3}{2}x} = \frac{9}{8} x^4 \tag{12.4.2}$$

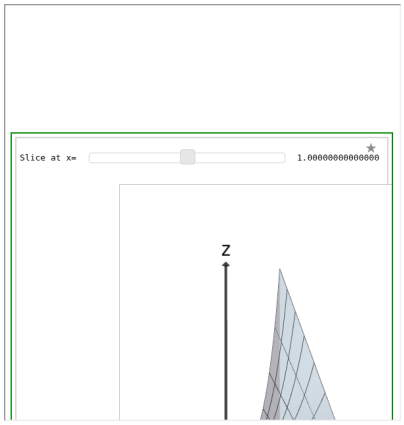
Using this to complete the evaluation of the interated integral, we now have

$$\begin{aligned} \int_{x=0}^{x=2} \int_{y=0}^{y=\frac{3}{2}x} x^2 y \, dy \, dx &= \int_{x=0}^{x=2} \frac{9}{8} x^4 \, dx \\ &= \frac{9}{8} \frac{x^5}{5} \Big|_{x=0}^{x=2} \\ &= \left(\frac{9}{8} \right) \left(\frac{32}{5} \right) \\ &= \frac{36}{5}. \end{aligned}$$

Figure 12.4.24 provides an opportunity to visualize both the double integral and the iterated integral.

i. The double integral of $f(x, y) = x^2 y$ over the region D can be seen as the volume of the solid below the surface $z = f(x, y)$ and above the xy-plane on the region D. This is shown as the pink volume in Figure 12.4.24. We can also use each of the iterated integrals to visualize the result.

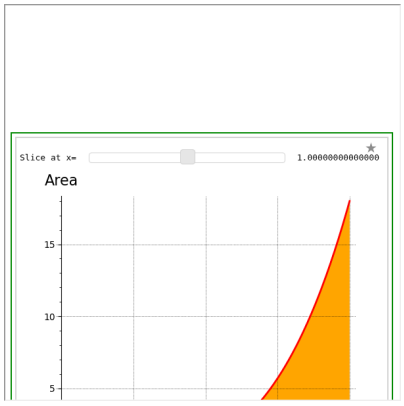
ii. The inner integral, which we evaluated in (12.4.2) to have value $\frac{9}{8}x^4$, corresponds to the area of a cross section along a slice with a fixed x-value, as shown by the blue cross section in Figure 12.4.24. The result of the inner integral changes depending on which x-value is used. Move the slider at the top of Figure 12.4.24 to change the x-value at which the slice of the solid is taken and observe that the area of the cross section increases as x increases.



Standalone
In Context
Embed

Figure 12.4.24 A plot of the volume below $f(x, y)$ over the region D with a highlighted cross section for $x = k$

(b) We conclude this example by discussing what the function obtained by evaluating the *inner* integral in the iterated integrals tells us geometrically. Evaluating the inner integral gave us the function $\frac{9}{8}x^4$. Figure 12.4.25 shows a plot of this function on the interval $[0, 2]$. The value of this function for a specific value of x gives the *area* of the cross section of the original solid at that location. Move the slider in the figure to see the value of this area function at different values of x. As you noticed when examining Figure 12.4.24, the cross sections with x close to zero are very small, and the areas of the cross sections increase as x increases.



Standalone
In Context
Embed

Figure 12.4.25 A plot of $\frac{9}{8}x^4$ on the interval $[0, 2]$

We know from single-variable calculus that evaluating the *outer integral* in the iterated integral must give the area under the graph in Figure 12.4.25. One of the interpretations we have been considering of a double integral tells us that this ultimately yields the *volume* of the three-dimensional solid. This makes sense because the values of the function being integrated in the outer integral are areas of cross sections.

▲

Activity 12.4.5 Let D be the triangular region with vertices $(0, 0)$, $(4, 0)$, and $(0, 2)$. Consider the double integral $\iint_D (4 - x - 2y)\, dA$.

(a) Draw and label a plot of D with relevant cross sections for a vertically simple description. Give the inequalities that show D is vertically simple.

(b) Write the double integral as an iterated integral of the form $\iint_D (4 - x - 2y)\,dy\,dx$, including correct bounds on the integral.

(c) Evaluate the iterated integrals from the previous part. Write a sentence describing at least one interpretation of the meaning of the value of the double integral.

The region we considered as vertically simple in Example 12.4.22 was shown in Example 12.4.8 to be *both* vertically simple and horizontally simple. We will now see that it is also possible to set up an iterated integral for the double integral using the horizontally simple description of the region.

Example 12.4.26 Consider again the function $f(x, y) = x^2 y$ on the triangle D with vertices $(0, 0)$, $(2, 0)$, and $(2, 3)$ as shown in Figure 12.4.23. This is the same region we described as both vertically simple and horizontally simple in Example 12.4.8. In Example 12.4.22, we set up an iterated integral for $\iint_D f(x, y)\,dA$ using the vertically simple description of D. In this example, we will use the horizontally simple description of D to write another iterated integral that computes thd double integral.

(a) We will use the inequalities from Example 12.4.8 that correspond to the horizontally simple description of D:

$$\text{Along } y = k, \text{ the } x\text{-coordinates go from } \frac{2}{3}k \text{ to } 2$$

This description corresponds to the inequalities

$$0 \le y \le 3 \qquad \frac{2}{3}y \le x \le 2.$$

Using these inequalities corresponds to integrating with respect to x in the inner integral of the iterated integral. That is, the variable in the inner integral is x and we treat y as a constant when antidifferentiating.

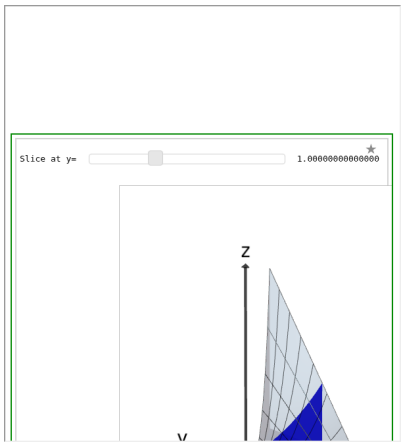
Therefore we can write the double integral as a $dx\,dy$ iterated integral in the form

$$\iint_D x^2 y\,dA = \int_{y=0}^{y=3} \int_{x=\frac{2}{3}y}^{x=2} x^2 y\,dx\,dy.$$

We can now evaluate the resulting iterated integral as before by twice applying the Fundamental Theorem of Calculus. In the inner integral, we treat y as a constant and integrate with respect to x. When we do so, we find that

$$\begin{aligned}
\int_{y=0}^{y=3} \int_{x=\frac{2}{3}y}^{2} x^2 y\,dx\,dy &= \int_{y=0}^{y=3} \left[\frac{x^3}{3} \cdot y\right]\Big|_{x=\frac{2}{3}y}^{x=2} dx \\
&= \int_{y=0}^{y=3} \left[\frac{8}{3}y - \frac{8}{81}y^4\right] dy \\
&= \left[\frac{8}{3}\frac{y^2}{2} - \frac{8}{81}\frac{y^5}{5}\right]\Big|_{y=0}^{y=3} \\
&= \left(\frac{8}{3}\right)\left(\frac{9}{2}\right) - \left(\frac{8}{81}\right)\left(\frac{243}{5}\right) \\
&= 12 - \frac{24}{5} \\
&= \frac{36}{5}.
\end{aligned}$$

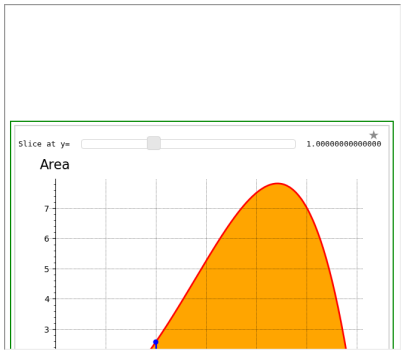
In Figure 12.4.27, you can see the same solid as before in pink. The volume of this solid is the value of the double integral. In this example, we used a horizontally simple description which means the slices correspond to constant values of y. A cross section along a constant y-value is shown in blue in the figure. Use the slider at the top of Figure 12.4.27 to see how the size of cross section changes with respect to y. Notice how as y increases, the area of the cross section first increases and then decreases.



Standalone
In Context
Embed

Figure 12.4.27 A plot of the solid below $f(x, y)$ over the region D

(b) As with our previous consideration of this double integral, we conclude this example by examining what the function obtained by evaluating the *inner* integral in the iterated integrals tells us geometrically. In part (a), the inner integral evaluated to $\frac{8}{3}y - \frac{8}{81}y^4$. The relevant y-values satisfy $0 \leq y \leq 3$. In Figure 12.4.28, we show a plot of this function on the interval $[0, 3]$ along with a slider. Moving the slider corresponds to changing the location of the cross section in Figure 12.4.27. In Figure 12.4.28, it helps illustrate the value of the inner integral at that value of y. The value of this function for a particular value of y is the *area* of the cross section of the solid at that y-position. Therefore, just as you visually observed by moving the slider in Figure 12.4.27, the cross section is small for y values near 0 or 3. The area of the cross section is largest when the y-value is slightly larger than 2. This figure demonstrates how the function obtained by evaluating the inner integral is the function giving the area of a cross section for a constant value of y.



Standalone
In Context
Embed

Figure 12.4.28 A plot of $\frac{8}{3}y - \frac{8}{81}y^4$ on the interval $[0, 3]$

Recall that evaluating the *outer* integral in part (a) finds the area of the shaded region in Figure 12.4.28. Remembering that the height of the

function's graph in that plot is the area of a cross section tells us that the outer integral corresponds to accumulating area. This leads to the value of the *double* integral measuring the *volume* of the solid.

▲

We have now seen that in a situation where D can be described as both horizontally simple and vertically simple, we can set up an iterated integral in either order. If we can reasonably evaluate the iterated integral in either order, then the iterated integrals yield the same value in each case. However, there are times where finding an antiderivative to evaluate the inner integral may be easier in one order than the other.

The process of switching between a vertically simple description and a horizontally simple description for the region of integration is often called **switching the order of integation** because the corresponding iterated integrals used to calculate a double integral will have the order in which you integrate these variables switched.

The possible meanings of a double integral $\iint_D f(x,y)\,dA$ over a non-rectangular region D are the same as those we stated for a double integral over a rectangular region:

- The volume of the solids the graph of f bounds above the xy-plane over the region D minus the volume of the solids the graph of f bounds below the xy-plane under the region D.
- If $f(x,y)$ is the density of a substance over a region D, then the double integral of f over D will measure the total amount of the substance in D.
- Dividing the double integral by the area of D gives the average value of f over the region D.

The next activity asks you to return to the function and triangular region considered in Activity 12.4.5 but to set up an iterated integral from the perspective of the region being horizontally simple.

Activity 12.4.6 Let D be the triangular region with vertices $(0,0)$, $(4,0)$, and $(0,2)$. Consider the double integral $\iint_D (4-x-2y)\,dA$.

(a) Draw and label a plot of D with relevant cross sections for a horizontally simple description. Give the inequalities that show D is horizontally simple.

(b) Write the double integral as an iterated integral of the form $\iint_D (4-x-2y)\,dx\,dy$, include correct bounds on the integral.

(c) Evaluate the iterated integral above, and verify that you get the same value as in Activity 12.4.5.

The next activity explores a region of integration with non-linear sides, given as bounds in an iterated integral. You will also switch the order of integration, evaluate whichever order is most convinient, and finally interpret the result as the average value.

Activity 12.4.7 Consider the iterated integral $\int_{x=0}^{x=1} \int_{y=x}^{y=\sqrt{x}} (4x+10y)\,dy\,dx$. Remember that not every iterated integral makes sense as a double integral, so as you proceed through this activity, pause to think about *why* this iterated integral does correspond to a double integral.

(a) Sketch the region of integration D for which

$$\iint_D (4x + 10y)\, dA = \int_{x=0}^{x=1} \int_{y=x}^{y=\sqrt{x}} (4x + 10y)\, dy\, dx.$$

(b) Determine the equivalent iterated integral that results from integrating in the opposite order ($dx\, dy$, instead of $dy\, dx$). That is, determine limits of integration for which

$$\iint_D (4x + 10y)\, dA = \int_{y=?}^{y=?} \int_{x=?}^{x=?} (4x + 10y)\, dx\, dy.$$

(c) Evaluate whichever one of the two iterated integrals above you find easier to compute. Explain what the value means in at least one of the interpretations of Key Idea 12.2.8.

(d) Set up and evaluate a *single* definite integral to determine the exact area $A(D)$ of D.

(e) Determine the exact average value of $f(x, y) = 4x + 10y$ over D.

We mentioned earlier that sometimes antidifferentiating in an iterated integral is difficult. To illustrate this, Exercise 12.4.5.23 asks you to consider

$$\int_{x=0}^{x=4} \int_{y=x/2}^{y=2} e^{y^2}\, dy\, dx$$

and explores how how switching the order of integration can sometimes yield a significantly easier iterated integral to evaluate algebraically.

Summary

- For a double integral $\iint_D f(x, y)\, dA$, we can write appropriate iterated integrals if the region D is either horizontally or vertically simple.
  - A horizontally simple region is described by the inequalities $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$.
  - A vertically simple region is described by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$.
- In an iterated double integral, the limits on the outer integral must be constants while the limits on the inner integral must be constants or in terms of only the remaining variable. In other words, an iterated integral corresponding to a double integral has one of the following forms (which result in the same value):

  $$\int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} f(x, y)\, dy\, dx,$$

  where $g_1 = g_1(x)$ and $g_2 = g_2(x)$ are functions of x only and the region D is described by the inequalities $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$ or

  $$\int_{y=c}^{y=d} \int_{x=h_1(y)}^{x=h_2(y)} f(x, y)\, dx\, dy,$$

  where $h_1 = h_1(y)$ and $h_2 = h_2(y)$ are functions of y only and the region D is described by the inequalities $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$.

12.4.5 Exercises

1. Decide, without calculation, if each of the integrals below are positive, negative, or zero. Let D be the region inside the unit circle centered at the origin. Let T, B, R, and L denote the regions enclosed by the top half, the bottom half, the right half, and the left half of unit circle, respectively.

(a) $\displaystyle\iint_B (y^3 + y^5)\, dA$

(b) $\displaystyle\iint_T (y^3 + y^5)\, dA$

(c) $\displaystyle\iint_D (y^3 + y^5)\, dA$

(d) $\displaystyle\iint_L (y^3 + y^5)\, dA$

(e) $\displaystyle\iint_R (y^3 + y^5)\, dA$

2. Match the following integrals with the verbal descriptions of the solids whose volumes they give. Put the letter of the verbal description to the left of the corresponding integral.

(a) $\displaystyle\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 - x^2 - y^2\, dydx$

(b) $\displaystyle\int_{-2}^{2}\int_{4}^{4+\sqrt{4-x^2}} 4x + 3y\, dydx$

(c) $\displaystyle\int_{0}^{\frac{1}{\sqrt{3}}}\int_{0}^{\frac{1}{2}\sqrt{1-3y^2}} \sqrt{1 - 4x^2 - 3y^2}\, dxdy$

(d) $\displaystyle\int_{0}^{1}\int_{y^2}^{\sqrt{y}} 4x^2 + 3y^2\, dxdy$

(e) $\displaystyle\int_{0}^{2}\int_{-2}^{2} \sqrt{4 - y^2}\, dydx$

A. One eighth of an ellipsoid.

B. Solid under an elliptic paraboloid and over a planar region bounded by two parabolas.

C. Solid under a plane and over one half of a circular disk.

D. One half of a cylindrical rod.

E. Solid bounded by a circular paraboloid and a plane.

3. In evaluating a double integral over a region D, a sum of iterated integrals was obtained as follows:

$$\iint_D f(x,y)\,dA = \int_0^2 \int_0^{(5/2)y} f(x,y)\,dxdy + \int_2^7 \int_0^{7-y} f(x,y)\,dxdy\,.$$

Sketch the region D and express the double integral as an iterated integral with reversed order of integration.

$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dydx$

$a =$ _____ $b =$ _____

$g_1(x) =$ _____ $g_2(x) =$ _____

4. Evaluate the integral

$$\int_0^1 \int_y^1 \sin(x^2)\ dx\,dy$$

by reversing the order of integration.

With order reversed,

$$\int_a^b \int_c^d \sin(x^2)\ dy\,dx,$$

where $a =$ _____, $b =$ _____, $c =$ _____, and $d =$ _____.

Evaluating the integral, $\int_0^1 \int_y^1 \sin(x^2)\ dx\,dy$ = _____

5. Evaluate the iterated integral $\int_0^3 \int_0^{e^v} \sqrt{1+e^v}\,dw\,dv$.

Answer: _____

6. Evaluate the double integral $I = \iint_{\mathbf{D}} xy\ dA$ where **D** is the triangular region with vertices $(0,0), (3,0), (0,3)$.

7. Evaluate the double integral $I = \iint_{\mathbf{D}} xy\ dA$ where **D** is the triangular region with vertices $(0,0), (4,0), (0,1)$.

8. Evaluate the integral by reversing the order of integration.

$\int_0^1 \int_{7y}^7 e^{x^2}\,dxdy =$ _____

9. Consider the integral $\int_0^7 \int_0^{\sqrt{49-y}} f(x,y)dxdy$. If we change the order of integration we obtain the sum of two integrals:

$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)dydx + \int_c^d \int_{g_3(x)}^{g_4(x)} f(x,y)dydx$

$a =$ _____ $b =$ _____

$g_1(x) =$ _____ $g_2(x) =$ _____

$c =$ _____ $d =$ _____

$g_3(x) =$ _____ $g_4(x) =$ _____

10. The region W lies below the surface $f(x,y) = 4e^{-(x-1)^2-y^2}$ and above the disk $x^2 + y^2 \le 4$ in the xy-plane.

(a) Think about what the contours of f look like. You may want to using $f(x,y) = 1$ as an example. Sketch a rough contour diagram on a separate sheet of paper.

(b) Write an integral giving the area of the cross-section of W in the plane $x = 1$.

Area = $\int_a^b$ _____ d_____,
where a = _____ and b =_____

(c) Use your work from (b) to write an iterated double integral giving the volume of W, using the work from (b) to inform the construction of the inside integral.

Volume = $\int_a^b \int_c^d$ _____ d_____ d_____,
where a = _____, b =_____ c = _____ and d =_____

11. Set up a double integral in rectangular coordinates for calculating the volume of the solid under the graph of the function $f(x, y) = 16 - x^2 - y^2$ and above the plane $z = 7$.

Instructions: Please enter the integrand in the first answer box. Depending on the order of integration you choose, enter *dx* and *dy* in either order into the second and third answer boxes with only one *dx* or *dy* in each box. Then, enter the limits of integration.

$\int_A^B \int_C^D$ _____ _____ _____
A = _____
B = _____
C = _____
D = _____

12. Find the volume of the solid bounded by the planes x = 0, y = 0, z = 0, and x + y + z = 2.

13. A pile of earth standing on flat ground has height 9 meters. The ground is the xy-plane. The origin is directly below the top of the pile and the z-axis is upward. The cross-section at height z is given by $x^2 + y^2 = 9 - z$ for $0 \le z \le 9$, with x, y, and z in meters.

(a) What equation gives the edge of the base of the pile?

◯ $x + y = 3$

◯ $x + y = 9$

◯ $x^2 + y^2 = 3$

◯ $x^2 + y^2 = 9$

◯ None of the above

(b) What is the area of the base of the pile?

(c) What equation gives the cross-section of the pile with the plane $z = 8$?

◯ $x^2 + y^2 = 8$

◯ $x^2 + y^2 = \sqrt{1}$

◯ $x^2 + y^2 = 64$

◯ $x^2 + y^2 = 1$

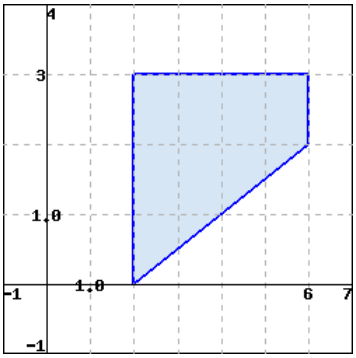
◯ None of the above

(d) What is the area of the cross-section $z = 8$ of the pile?

(e) What is $A(z)$, the area of a horizontal cross-section at height z?
$A(z)$ = _____ square meters

(f) Use your answer in part (e) to find the volume of the pile.
Volume = _____ cubic meters

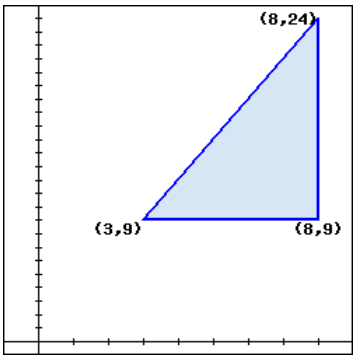
14. Consider the shaded region in the graph below.



Write $\int_R f\,dA$ on this region as an iterated integral:
$\int_R f\,dA = \int_a^b \int_c^d f(x,y)\,d_____\;d_____$ where
$a =$ _____,
$b =$ _____,
$c =$ _____, and
$d =$ _____.

15. Evaluate the integral $\int_R 4xy\,dA$, where R is the triangle $5x + y \le 5,\ x \ge 0,\ y \ge 0$.
$\int_R 4xy\,dA =$ _____

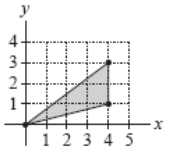
16. Find the volume under the graph of the function $f(x,y) = 8x^2y$ over the region shown in the figure below.



volume = _____

17. Evaluate the double integral $\iint_D \left(x^2 + 6y\right)\,dA$, where D is bounded by $y = x$, $y = x^3$, and $x \ge 0$.
Answer: _____

18. Calculate the double integral of $f(x,y)$ over the triangle indicated in the following figure:



$f(x,y) = 24ye^x$
Answer : _____

19. For each of the following iterated integrals,

- sketch the region of integration,
- write an equivalent iterated integral expression in the opposite order of integration,
- choose one of the two orders and evaluate the integral.

(a) $\int_{x=0}^{x=1} \int_{y=x^2}^{y=x} xy \, dy \, dx$

(b) $\int_{y=0}^{y=2} \int_{x=-\sqrt{4-y^2}}^{x=0} xy \, dx \, dy$

(c) $\int_{x=0}^{x=1} \int_{y=x^4}^{y=x^{1/4}} x + y \, dy \, dx$

(d) $\int_{y=0}^{y=2} \int_{x=y/2}^{x=2y} x + y \, dx \, dy$

20. The temperature at any point on a metal plate in the xy-plane is given by $T(x, y) = 100 - 4x^2 - y^2$, where x and y are measured in inches and T in degrees Celsius. Consider the portion of the plate that lies on the region D that is the finite region that lies between the parabolas $x = y^2$ and $x = 3 - 2y^2$.

(a) Construct a labeled sketch of the region D.

(b) Set up an iterated integral whose value is $\iint_D T(x, y) \, dA$, using $dA = dxdy$

Hint. It is possible that more than one integral is needed.

(c) Set up an integrated integral whose value is $\iint_D T(x, y) \, dA$, using $dA = dydx$.

Hint. It is possible that more than one integral is needed.

(d) Use the Fundamental Theorem of Calculus to evaluate the integrals you determined in (b) and (c).

(e) Determine the exact average temperature, $T_{\text{AVG}(D)}$, over the region D.

21. Consider the solid that is given by the following description: the base is the given region D, while the top is given by the surface $z = p(x, y)$. In each setting below, set up, but do not evaluate, an iterated integral whose value is the exact volume of the solid. Include a labeled sketch of D in each case.

(a) D is the interior of the quarter circle of radius 2, centered at the origin, that lies in the second quadrant of the plane; $p(x, y) = 16 - x^2 - y^2$.

(b) D is the finite region between the line $y = x + 1$ and the parabola $y = x^2$; $p(x, y) = 10 - x - 2y$.

(c) D is the triangular region with vertices $(1, 1)$, $(2, 2)$, and $(2, 3)$; $p(x, y) = e^{-xy}$.

(d) D is the region bounded by the y-axis, $y = 4$ and $x = \sqrt{y}$; $p(x, y) = \sqrt{1 + x^2 + y^2}$.

22. Consider the iterated integral $I = \int_{x=0}^{x=4} \int_{y=\sqrt{x}}^{y=2} \cos(y^3)\, dy\, dx$.

(a) Sketch the region of integration.

(b) Write an equivalent iterated integral with the order of integration reversed.

(c) Choose one of the two orders of integration and evaluate the iterated integral you chose by hand. Explain the reasoning behind your choice.

(d) Determine the exact average value of $\cos(y^3)$ over the region D that is determined by the iterated integral I.

23. Consider the iterated integral $\int_{x=0}^{x=4} \int_{y=x/2}^{y=2} e^{y^2}\, dy\, dx$.

(a) Explain why we cannot find a simple antiderivative for e^{y^2} with respect to y, and thus are unable to evaluate $\int_{x=0}^{x=4} \int_{y=x/2}^{y=2} e^{y^2}\, dy\, dx$ in the indicated order using the Fundamental Theorem of Calculus.

(b) Given that $\iint_D e^{y^2}\, dA = \int_{x=0}^{x=4} \int_{y=x/2}^{y=2} e^{y^2}\, dy\, dx$, sketch the region of integration, D.

(c) Rewrite the given iterated integral in the opposite order using $dA = dx\, dy$.

Hint. You may need more than one integral.

(d) Use the Fundamental Theorem of Calculus to evaluate the iterated integral you wrote in the previous part. Write one sentence to explain the meaning of the value you found in terms of an interpretation from Key Idea 12.2.8.

(e) Write a sentence or two to describe what you have learned in this activity regarding the order in which you set up an iterated integral.

12.5 Applications of Double Integrals

Motivating Questions

- How can a double integral be used to find the area of a region in the xy-plane?
- If we have a mass density function for a thin plate, how does a double integral determine the mass of the lamina?
- What is a joint probability density function? How do we determine the probability of an event if we know a probability density function?
- Given a mass density function on a lamina, how can we find the lamina's center of mass?

12.5.1 Introduction

So far, we have interpreted the double integral of a function f over a region D in three main ways, as described in Key Idea 12.2.8. However, our focus has been primarily on using double integrals to find the volume the surface defined by $z = f(x, y)$ bounds above the xy-plane over D minus the volume the surface bounds below the xy-plane over D. In this section, we investigate several other applications of double integrals, using the integration process as seen in Preview Activity 12.5.1: we partition into small regions, approximate the desired quantity on each small region, then use the integral to sum these values exactly in the limit.

The following preview activity explores the second interpretation of double integrals in Key Idea 12.2.8: a double integral can be used to determine the mass of a thin plate with a density distribution, $\delta(x, y)$. Recall that in single-variable calculus, we considered a similar problem and computed the mass of a one-dimensional rod with a mass-density distribution. There, as here, the key idea is that if density is constant, mass is the product of density and volume.

Preview Activity 12.5.1 Suppose that we have a flat, thin object (called a **lamina**) whose density varies across the object. We can think of the density on a lamina as a measure of mass per unit area. As an example, consider a circular plate D of radius 1 cm centered at the origin whose density δ varies depending on the distance from its center so that the density in grams per square centimeter at point (x, y) is

$$\delta(x, y) = 10 - 2(x^2 + y^2).$$

(a) Partition the plate into subrectangles R_{ij}, where $1 \le i \le m$ and $1 \le j \le n$, so that each subrectangle has area ΔA. Select a point (x^*_{ij}, y^*_{ij}) in R_{ij} for each i and j. Write a couple of sentences to explain the meaning of the quantity $\delta(x^*_{ij}, y^*_{ij})\Delta A$.

(b) Write a double Riemann sum that computes an approximation of the mass of the plate.

(c) Explain why the double integral $\iint_D \delta(x, y)\, dA$ tells us the exact mass of the plate.

(d) Write an iterated integral which, if evaluated, would give the exact mass of the plate. Do not actually evaluate the integral. (This integral is considerably easier to evaluate in polar coordinates, which we will learn more about in Section 12.6.)

12.5.2 Area

The simplest possible density function we could have considered in the Preview Activity would have been $\delta(x, y) = 1$, so we begin by considering what such a double integral represents. In fact, we can do this without needing to think of density at all. If the integrand in a double integral is 1, then our volume interpretation tells us that the double integral represents the volume of the solid underneath the plane $z = 1$ and above the region D of integration in the xy-plane. Since this solid is cylindrical and has height 1, its volume is also the area of the base region D, as we can see in the activity below.

Activity 12.5.2 Suppose we want to find the area of the bounded region D between the curves

$$y = 1 - x^2 \qquad \text{and} \qquad y = x - 1$$

A picture of this region is shown in Figure 12.5.1.

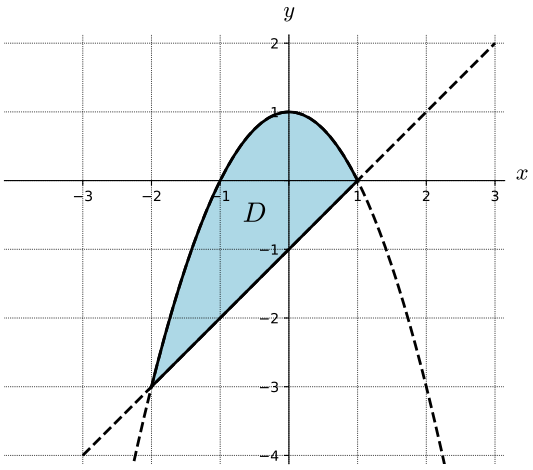


Figure 12.5.1 The graphs of $y = 1 - x^2$ and $y = x - 1$

(a) The volume of a solid with constant height is given by the area of the base times the height. Hence, we may interpret the area of the region D as the volume of a solid with base D and of uniform height 1. That is, the area of D is given by $\iint_D 1\,dA$. Write an iterated integral whose value is $\iint_D 1\,dA$.

Hint. Which order of integration might be more efficient? Why?

(b) Evaluate the iterated integral from (a). What does the result tell you?

We now formally state the conclusion from our earlier discussion and Activity 12.5.2.

The Double Integral and Area.

Given a closed, bounded region D in the plane, the area of D, denoted $A(D)$, is given by the double integral

$$A(D) = \iint_D 1\,dA.$$

12.5.3 Mass

Density is a measure of some quantity per unit area or volume. For example, we can measure the human population density of some region as the number of humans in that region divided by the area of that region. In physics, the mass density of an object is the mass of the object per unit area or volume. As suggested by Preview Activity 12.5.1, the following holds in general.

The Mass of a Lamina.

If $\delta(x, y)$ describes the density of a lamina defined by a planar region D, then the **mass** of D is given by the double integral $\iint_D \delta(x, y)\, dA$.

Activity 12.5.3 Let D be a half-disk lamina of radius 3 cm oriented in quadrants IV and I, centered at the origin as shown in Figure 12.5.2. Assume the density at point (x, y) is given by $\delta(x, y) = x$ with units of grams per square centimeter. Find the exact mass of the lamina.

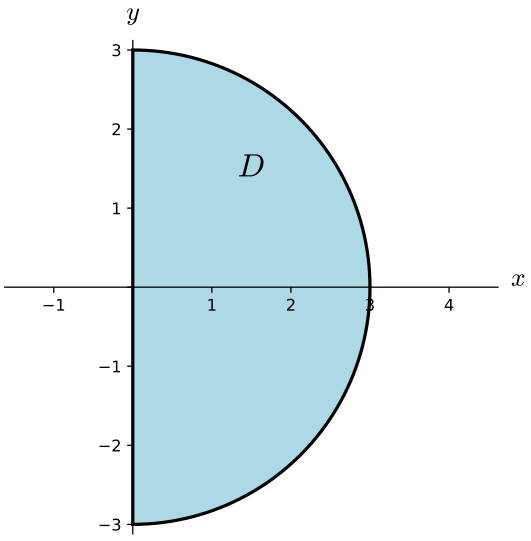


Figure 12.5.2 A half-disk lamina

12.5.4 Probability

Calculating probabilities is an important application of integration in the physical, social, and life sciences. To understand the basics, consider the game of darts in which a player throws a dart at a board and tries to hit a particular target. Suppose that a dart board is in the form of a disk D with radius 10 inches. If we assume that a player throws a dart at random without aiming at any particular point, then it is equally probable that the dart will strike any single point on the board. For instance, the probability that the dart will strike a particular 1 square inch region is $\frac{1}{100\pi}$, or the ratio of the area of the desired target to the total area of D. (This makes the perhaps unreasonable assumption that the dart thrower always hits the board itself at some point.) Similarly, the probability that the dart strikes a point in a disk D_3 of radius 3 inches inside the larger disk is given by the area of D_3 divided by the area of D. In other words, the probability that the dart strikes the disk D_3 is

$$\frac{9\pi}{100\pi} = \iint_{D_3} \frac{1}{100\pi}\, dA.$$

The integrand, $\frac{1}{100\pi}$, may be thought of as a **distribution function**, describing how the dart strikes are distributed across the board. In this case the distribution function is constant since we are assuming a uniform distribution. However, we can envision situations where the distribution function varies. For example, if the player is fairly good and is aiming for the bullseye (the center

of D), then the distribution function f could be skewed toward the center, perhaps

$$f(x,y) = Ke^{-(x^2+y^2)}$$

for some constant positive K. If we assume that the player is consistent enough so that the dart always strikes the board, then the total probability that the dart strikes the board somewhere is 1, and the distribution function f must satisfy[1]

$$\iint_D f(x,y)\,dA = 1.$$

For such a function f, the probability that the dart strikes in the disk D_1 of radius 1 would be

$$\iint_{D_1} f(x,y)\,dA.$$

Indeed, the probability that the dart strikes in any region R that lies within D is given by

$$\iint_R f(x,y)\,dA.$$

The preceding discussion highlights the general idea behind calculating probabilities. We assume we have a **joint probability density function** f, which is a function of two independent variables x and y defined on a domain D that satisfies the conditions

- $f(x,y) \geq 0$ for all x and y in D,
- the probability that x is between some values a and b while y is between some values c and d is given by

$$\int_a^b \int_c^d f(x,y)\,dy\,dx$$

and

- the probability that the point (x,y) is in D is 1, that is

$$\iint_D f(x,y)\,dA = 1. \tag{12.5.1}$$

Note that it is possible that D could be an infinite region and the limits on the integral in Equation (12.5.1) could be infinite. When we have such a probability density function $f = f(x,y)$, the probability that the point (x,y) is in some region R contained in the domain D is determined by

$$P((x,y) \text{ in } R) = \iint_R f(x,y)\,dA.$$

Activity 12.5.4 A firm manufactures smoke detectors. Two components for the detectors come from different suppliers—one in Michigan and one in Ohio. The company studies these components for their reliability and their data suggests that if x is the life span in years of a randomly chosen component from the Michigan supplier and y the life span in years of a randomly chosen component from the Ohio supplier, then the joint probability density function f could be given by

$$f(x,y) = e^{-x}e^{-y}.$$

[1] This makes $K = \frac{1}{\pi\left(1-e^{-100}\right)}$, which you can check.

(a) Theoretically, the components might last forever, so the domain D of the function f is the set D of all (x, y) such that $x \geq 0$ and $y \geq 0$. To show that f is a probability density function on D we need to demonstrate that

$$\iint_D f(x, y)\, dA = 1,$$

or that

$$\int_0^\infty \int_0^\infty f(x, y)\, dy\, dx = 1.$$

Use your knowledge of improper integrals to verify that f is indeed a probability density function.

(b) Assume that the smoke detector fails only if both of the supplied components fail. To determine the probability that a randomly selected detector will fail within one year, we will need to determine the probability that the life span of each component is between 0 and 1 years. Set up an appropriate iterated integral, and evaluate the integral to determine the probability.

(c) What is the probability that a randomly chosen smoke detector will fail between years 3 and 7?

(d) Suppose that the manufacturer determines that one of the components is more likely to fail than the other, and hence conjectures that the probability density function is instead $f(x, y) = Ke^{-x}e^{-2y}$. What is the value of K?

12.5.5 Center of Mass

The center of mass of an object is a point at which the object will balance perfectly. For example, the center of mass of a circular disk of uniform density is located at its center. For any object, if we throw it through the air, it will spin around its center of mass and behave as if all the mass is located at the center of mass.

In order to understand the role that integrals play in determining the center of mass of an object with a nonuniform mass distribution, we start by finding the center of mass of a collection of N distinct point-masses in the plane.

Let m_1, m_2, ..., m_N be N masses located in the plane. Think of these masses as connected by rigid rods of negligible weight from some central point (x, y). A picture with four masses is shown in Figure 12.5.3. Now imagine balancing this system by placing it on a thin pole at the point (x, y) perpendicular to the plane containing the masses. Unless the masses are perfectly balanced, the system will fall off the pole. The point $(\overline{x}, \overline{y})$ at which the system will balance perfectly is called the **center of mass** of the system. Our goal is to determine the center of mass of a system of discrete masses, then extend this to a continuous lamina.

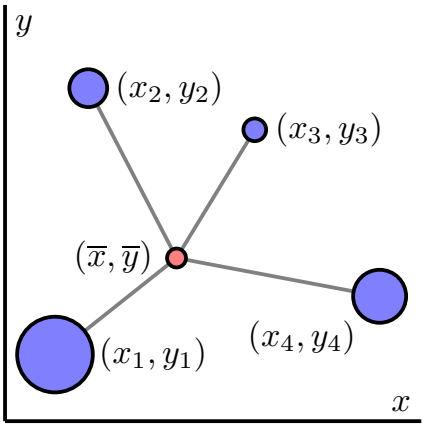


Figure 12.5.3 A center of mass $(\overline{x}, \overline{y})$ of four masses.

Each mass exerts a force (called a **moment**) around the lines $x = \overline{x}$ and $y = \overline{y}$ that causes the system to tilt in the direction of the mass. These moments are dependent on the mass and the distance from the given line. Let (x_1, y_1) be the location of mass m_1, (x_2, y_2) the location of mass m_2, etc. In order to balance perfectly, the moments in the x direction and in the y direction must be in equilibrium. We determine these moments and solve the resulting system to find the equilibrium point $(\overline{x}, \overline{y})$ at the center of mass.

The force that mass m_1 exerts to tilt the system from the line $y = \overline{y}$ is

$$m_1 g(\overline{y} - y_1),$$

where g is the gravitational constant. Similarly, the force mass m_2 exerts to tilt the system from the line $y = \overline{y}$ is

$$m_2 g(\overline{y} - y_2).$$

In general, the force that mass m_k exerts to tilt the system from the line $y = \overline{y}$ is

$$m_k g(\overline{y} - y_k).$$

For the system to balance, we need the forces to sum to 0, so that

$$\sum_{k=1}^{N} m_k g(\overline{y} - y_k) = 0.$$

Solving for $\overline{y}$, we find that

$$\overline{y} = \frac{\sum_{k=1}^{N} m_k y_k}{\sum_{k=1}^{N} m_k}.$$

A similar argument shows that

$$\overline{x} = \frac{\sum_{k=1}^{N} m_k x_k}{\sum_{k=1}^{N} m_k}.$$

The value $M_x = \sum_{k=1}^{N} m_k y_k$ is called the **total moment** with respect to the x-axis; $M_y = \sum_{k=1}^{N} m_k x_k$ is the **total moment** with respect to the y-axis. Hence, the respective quotients of the moments to the total mass M determines the center of mass of a point-mass system:

$$(\overline{x}, \overline{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M} \right)$$

Now, suppose that rather than a point-mass system, we have a continuous lamina with a variable mass-density $\delta(x,y)$. We may estimate its center of mass by partitioning the lamina into mn subrectangles of equal area ΔA, and treating the resulting partitioned lamina as a point-mass system. In particular, we select a point (x_{ij}^*, y_{ij}^*) in the ijth subrectangle and observe that the quanity

$$\delta(x_{ij}^*, y_{ij}^*)\Delta A$$

is density times area, so $\delta(x_{ij}^*, y_{ij}^*)\Delta A$ approximates the mass of the small portion of the lamina determined by the subrectangle R_{ij}.

We now treat $\delta(x_{ij}^*, y_{ij}^*)\Delta A$ as a point mass at the point (x_{ij}^*, y_{ij}^*). The coordinates $(\overline{x}, \overline{y})$ of the center of mass of these mn point masses are thus given by

$$\overline{x} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m} x_{ij}^*\delta(x_{ij}^*, y_{ij}^*)\Delta A}{\sum_{j=1}^{n}\sum_{i=1}^{m} \delta(x_{ij}^*, y_{ij}^*)\Delta A} \quad \text{and} \quad \overline{y} = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m} y_{ij}^*\delta(x_{ij}^*, y_{ij}^*)\Delta A}{\sum_{j=1}^{n}\sum_{i=1}^{m} \delta(x_{ij}^*, y_{ij}^*)\Delta A}.$$

If we take the limit as m and n go to infinity, we obtain the exact center of mass $(\overline{x}, \overline{y})$ of the continuous lamina.

The center of mass of a lamina.

The coordinates $(\overline{x}, \overline{y})$ of the **center of mass of a lamina** D with density $\delta = \delta(x,y)$ are given by

$$\overline{x} = \frac{\iint_D x\delta(x,y)\,dA}{\iint_D \delta(x,y)\,dA} \quad \text{and} \quad \overline{y} = \frac{\iint_D y\delta(x,y)\,dA}{\iint_D \delta(x,y)\,dA}.$$

The center of mass of a lamina can then be thought of as a *weighted average* of all of the points in the lamina with the weights given by the density at each point. The **centroid** of a lamina is the just the average of all of the points in the lamina, which is the same as the center of mass in the case where the density at each point is 1.

The numerators of $\overline{x}$ and $\overline{y}$ are called the respective **moments** of the lamina about the coordinate axes. Thus, the moment of a lamina D with density $\delta = \delta(x,y)$ about the y-axis is

$$M_y = \iint_D x\delta(x,y)\,dA$$

and the moment of D about the x-axis is

$$M_x = \iint_D y\delta(x,y)\,dA.$$

If M is the mass of the lamina, it follows that the center of mass is

$$(\overline{x}, \overline{y}) = \left(\frac{M_y}{M}, \frac{M_x}{M}\right).$$

Activity 12.5.5 In this activity we determine integrals that represent the center of mass of a lamina D described by the triangular region bounded by the x-axis and the lines $x = 1$ and $y = 2x$ in the first quadrant if the density at point (x,y) is $\delta(x,y) = 6x + 6y + 6$. A picture of the lamina is shown in Figure 12.5.4.

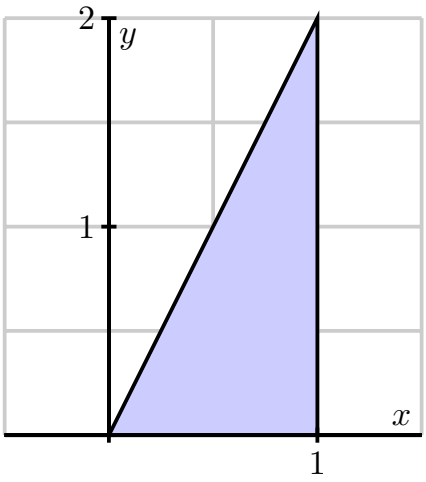


Figure 12.5.4 The lamina bounded by the x-axis and the lines $x = 1$ and $y = 2x$ in the first quadrant.

(a) Set up an iterated integral that computes the total mass of the lamina.

(b) Assume the total mass of the lamina is 14. Set up two iterated integrals that represent the coordinates of the center of mass of the lamina.

Summary

- The area of a region D in the plane has the same numerical value as the volume of a solid of uniform height 1 and base D, so the area of D is given by $\iint_D 1\,dA$.
- The mass of a lamina D with a mass density function $\delta = \delta(x, y)$ is $\iint_D \delta(x, y)\,dA$.
- Given a joint probability density function f is a function of two independent variables x and y defined on a domain D, if R is some subregion of D, then the probability that (x, y) is in R is given by

$$\iint_R f(x, y)\,dA.$$

- The center of mass, $(\overline{x}, \overline{y})$, of a continuous lamina with a variable density $\delta(x, y)$ is given by

$$\overline{x} = \frac{\iint_D x\delta(x, y)\,dA}{\iint_D \delta(x, y)\,dA} \qquad \text{and} \qquad \overline{y} = \frac{\iint_D y\delta(x, y)\,dA}{\iint_D \delta(x, y)\,dA}.$$

12.5.6 Exercises

1. The masses m_i are located at the points P_i. Find the center of mass of the system.

$m_1 = 7$, $m_2 = 8$, $m_3 = 3$.
$P_1 = (5, -3)$, $P_2 = (6, -5)$, $P_3 = (8, 8)$.
$\bar{x}$= _____
$\bar{y}$= _____

2. Find the centroid $(\bar{x}, \bar{y})$ of the triangle with vertices at $(0, 0)$, $(5, 0)$, and $(0, 1)$.

$\bar{x}=$ _____

$\bar{y}=$ _____

3. Find the mass of the rectangular region $0 \leq x \leq 1$, $0 \leq y \leq 1$ with density function $\rho(x, y) = 1 - y$.

4. Find the mass of the triangular region with vertices (0, 0), (2, 0), and (0, 6), with density function $\rho(x, y) = x^2 + y^2$.

5. A lamina occupies the region inside the circle $x^2 + y^2 = 6y$ but outside the circle $x^2 + y^2 = 9$. The density at each point is inversely proportional to its distance from the origin.

Where is the center of mass?

(_____, _____)

6. A lamina occupies the part of the disk $x^2 + y^2 \leq 9$ in the first quadrant and the density at each point is given by the function $\rho(x, y) = (x^2 + y^2)$.

A. What is the total mass? _____

D. Where is the center of mass? (_____, _____)

7. A lamina occupies the part of the disk $x^2 + y^2 \leq 9$ in the first quadrant and the density at each point is given by the function $\rho(x, y) = 5(x^2 + y^2)$.

A. What is the total mass? _____

B. What is the moment about the x-axis? _____

C. What is the moment about the y-axis? _____

D. Where is the center of mass? (_____, _____)

E. What is the moment of inertia about the origin? _____

8. A sprinkler distributes water in a circular pattern, supplying water to a depth of e^{-r} feet per hour at a distance of r feet from the sprinkler.

A. What is the total amount of water supplied per hour inside of a circle of radius 7?

_____ ft^3 per hour

B. What is the total amount of water that goes through the sprinkler per hour?

_____ ft^3 per hour

9. Let p be the joint density function such that $p(x, y) = \frac{1}{9}xy$ in R, the rectangle $0 \leq x \leq 6, 0 \leq y \leq 1$, and $p(x, y) = 0$ outside R. Find the fraction of the population satisfying the constraint $x + y \leq 7$

fraction = _____

10. A lamp has two bulbs, each of a type with an average lifetime of 11 hours. The probability density function for the lifetime of a bulb is $f(t) = \frac{1}{11}e^{-t/11}, t \geq 0$.

What is the probability that both of the bulbs will fail within 5 hours?

11. For the following two functions $p(x, y)$, check whether p is a joint density function. Assume $p(x, y) = 0$ outside the region R.

(a) $p(x, y) = 1$, where R is $0 \leq x \leq 0.5, -1 \leq y \leq 0$.

$p(x, y)$ (☐ is a joint density function ☐ is not a joint density function)

(b) $p(x, y) = 2$, where R is $-1 \leq x \leq 1, 2 \leq y \leq 3$.

$p(x, y)$ (☐ is a joint density function ☐ is not a joint density function)

Then, for the region R given by $2 \le x \le 5, 1 \le y \le 4$, what constant function $p(x, y)$ is a joint density function?

$p(x, y) =$ _____

12. Let x and y have joint density function

$$p(x,y) = \begin{cases} \frac{2}{3}(x+2y) & \text{for } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that

(a) $x > 1/2$:

probability = _____

(b) $x < \frac{1}{2} + y$:

probability = _____

13. A lamp has two bulbs, each of a type with an average lifetime of 4 hours. The probability density function for the lifetime of a bulb is $f(t) = \frac{1}{4}e^{-t/4}, t \ge 0$.

What is the probability that both of the bulbs will fail within 5 hours?

14. Let p be the joint density function such that $p(x, y) = \frac{1}{16}xy$ in R, the rectangle $0 \le x \le 4, 0 \le y \le 2$, and $p(x, y) = 0$ outside R. Find the fraction of the population satisfying the constraint $x \ge y$

fraction = _____

15. A point is chosen at random from the region S in the xy-plane containing all points (x, y) such that $-2 \le x \le 2, -1 \le y \le 1$ and $x - y \le 0$ (at random means that the density function is constant on S).

(a) Determine the joint density function for x and y in S

$p(x, y) =$ _____

(with $p(x, y) = 0$ for all (x, y) not in S.)

(b) If T is a subset of S with area a, find the probability that a point (x, y) is in T.

probability = _____

16. Electric charge is distributed over the disk

$x^2+y^2 \le 12$ so that the charge density at (x,y) is $\sigma(x, y) = 14+x^2+y^2$ coulombs per square meter.

Find the total charge on the disk.

17. Electric charge is distributed over the disk

$x^2+y^2 \le 14$ so that the charge density at (x,y) is $\sigma(x, y) = 7+x^2+y^2$ coulombs per square meter.

Find the total charge on the disk.

18. A triangular plate is bounded by the graphs of the equations $y = 2x$, $y = 4x$, and $y = 4$. The plate's density at (x, y) is given by $\delta(x, y) = 4xy^2 + 1$, measured in grams per square centimeter (and x and y are measured in centimeters).

(a) Set up an iterated integral whose value is the mass of the plate. Include a labeled sketch of the region of integration. Why did you choose the order of integration you did?

(b) Determine the mass of the plate.

(c) Determine the exact center of mass of the plate. Draw and label the point you find on your sketch from (a).

(d) What is the average density of the plate? Include units on your answer.

19. Let D be a half-disk lamina of radius 3 in quadrants IV and I, centered at the origin as in Activity 12.5.3. Assume the density at point (x, y) is equal to x.

(a) Before doing any calculations, what do you expect the y-coordinate of the center of mass to be? Why?

(b) Set up iterated integral expressions which, if evaluated, will determine the exact center of mass of the lamina.

(c) Use appropriate technology to evaluate the integrals to find the center of mass numerically.

20. Let x denote the time (in minutes) that a person spends waiting in a checkout line at a grocery store and y the time (in minutes) that it takes to check out. Suppose the joint probability density for x and y is

$$f(x, y) = \frac{1}{8} e^{-x/4 - y/2}.$$

(a) What is the exact probability that a person spends between 0 to 5 minutes waiting in line, and then 0 to 5 minutes waiting to check out?

(b) Set up, but do not evaluate, an iterated integral whose value determines the exact probability that a person spends at most 10 minutes total both waiting in line and checking out at this grocery store.

(c) Set up, but do not evaluate, an iterated integral expression whose value determines the exact probability that a person spends at least 10 minutes total both waiting in line and checking out, but not more than 20 minutes.

12.6 Double Integrals in Polar Coordinates

Motivating Questions

- What is the area element dA in polar coordinates?
- How do we convert a double integral in rectangular coordinates to a double integral in polar coordinates?

12.6.1 Introduction

Suppose that we would like to integrate $f(x, y) = e^{x^2+y^2}$ over the the unit disc D centered at the origin. Using the methods we have learned thus far tin the chapter, we could evaluate this double integral using the iterated integral

$$\iint_D f(x, y)\, dA = \int_{x=-1}^{x=1} \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} e^{x^2+y^2}\, dy\, dx.$$

This relies on treating D as vertically simple, as suggested in Figure 12.6.1.

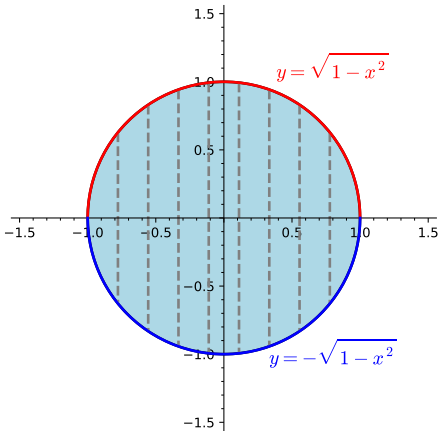


Figure 12.6.1 A unit disc centered at the origin with slices showing a vertically simple description

For this particular integral, we are unable to find an antiderivative of the integrand. Changing the order of integration does not help, as the region and the integrand are symmetric in x and y. Furthermore, even if we could find an antiderivative, the inner limits of integration involve relatively complicated functions. However, the region of integration has rotational symmetry about the origin. This should make you think about polar coordinates, which we introduced in Subsection 9.9.2. In this section, we will see how we can use polar coordinates to write double integrals as iterated integrals, allowing us to both streamline the writing of the limits of integration and make evaluating the iterated integral easier.

Recall that a point P in rectangular coordinates that is described by an ordered pair (x, y), where x is the displacement from P to the y-axis and y is the displacement from P to the x-axis can also be described with polar coordinates (r, θ), where r is the distance from P to the origin and θ is the angle formed by the line segment $\overline{OP}$ and the positive x-axis, as shown in Figure 12.6.2.

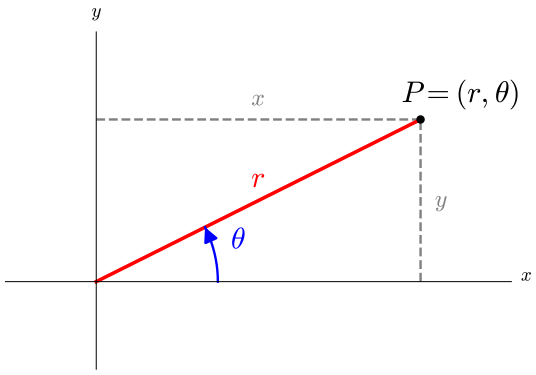


Figure 12.6.2 A plot with polar coordinate measurements labeled

Trigonometry and the Pythagorean Theorem allow for straightforward conversion from rectangular to polar, and vice versa. A reminder of these formulas, which were developed in Subsection 9.9.2, is included below.

Converting between rectangular and polar coordinates.

From polar to rectangular Given polar coordinates (r, θ) of a point P, the rectangular coordinates (x, y) of P satisfy

$$x = r\cos(\theta) \qquad y = r\sin(\theta).$$

From rectangular to polar Given rectangular coordinates (x, y) of a point P, the polar coordinates (r, θ) of P satisfy

$$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x},$$

assuming $x \neq 0$.

The labels in the box above are useful when thinking about converting the *coordinates* of a point from one coordinate system to the other. As we think about iterated integrals, however, we will often need to take a function $z = f(x, y)$ in which the inputs are described in rectangular coordinates and rewrite it using polar coordinates. This task simply requires using substitutions for x and y given in the first part of the conversions.

Recall that when considering polar coordinates, it is important to think about how the sign of r influences the position of a point. There are also multiple angle measurements in polar coordinates that can be used for a single point, so it will be helpful to think flexibly about ranges of θ when describing regions in polar coordinates. Refer to Example 9.9.12 and later examples for different ways to interpret and make conversions between rectangular and polar coordinates.

We can draw graphs of curves in polar coordinates in a similar way to how we do in rectangular coordinates. However, when plotting in polar coordinates, we use a grid that considers changes in angles and changes in distance from the origin. In particular, the angles θ and distances r partition the plane into small wedges as shown in Figure 12.6.3.

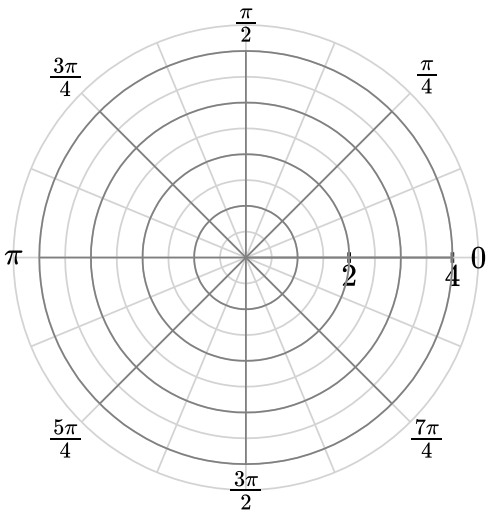


Figure 12.6.3 The polar coordinate grid

Preview Activity 12.6.1 This Preview Activity asks you to practice converting some equations and regions between rectangular and polar coordinates.

(a) For each of the following polar equations, draw a plot in the plane of the curve. Write a sentence for each graph to explain why the equation is satisfied.

i. $r = 1$ ii. $r = 3$ iii. $r = 0$ iv. $\theta = 1$ v. $\theta = \frac{3\pi}{4}$

(b) Draw a plot of the region described by the following inequalities:

$$1 \leq r \leq 3 \quad \text{and} \quad \frac{\pi}{2} \leq \theta \leq \frac{3\pi}{4}$$

(c) Sketch the region in the xy-plane with $x \leq 0$, $y \leq 0$, and inside the circle $x^2 + y^2 = 25$. State inequalities for r and θ that describe the region.

(d) Sketch the region in the xy-plane that lies between the circles $x^2 + y^2 = 4$ and $x^2 + y^2 = 9$. State inequalities for r and θ that describe the region.

12.6.2 Integration in Polar Coordinates

To find a way to evaluate the double integral

$$\iint_D e^{x^2+y^2}\, dA$$

where D is the unit disk using polar coordinates, we will find a way to write this double integral as an iterated integral in the variables r and θ. We will need to complete three steps to transform this double integral into polar coordinates:

1. Convert the area element to polar coordinates.
2. Describe the region of integration with inequalities that correspond to fixing one coordinate and bounding the other in a consistent way.
3. Convert the integrand $f(x, y)$ to polar coordinates.

The second and third of these steps are already things we have discussed in the context of understanding polar coordinates, and we will return to them for this specific integral shortly. However, understanding how to write the area element dA in polar coordinates is a new topic. To do this, we focus on how the area element dA can be written in terms of dr and $d\theta$ in the context of polar coordinates.

To write dA in polar coordinates, we start by remembering what dA represents. It comes from ΔA, which is the *area* of the base of a rectangular prism used to approximate a volume. In rectangular coordinates, $\Delta A = \Delta x\, \Delta y$, which leads to $dA = dx\, dy$. To think in polar coordinates, we need to consider the area of a *polar* rectangle with $R_0 \leq r \leq R_1$ and $\theta_0 \leq \theta \leq \theta_1$. Figure 12.6.4 shows two polar rectangles with the same values of θ_0 and θ_1. However, the figure illustrates that a polar rectangle located farther from the origin has greater area than one located closer to the origin, even when $\Delta\theta = \theta_1 - \theta_0$ and $\Delta r = R_1 - R_0$ are the same for the two polar rectangles. This suggests that dA depends on not only dr and $d\theta$ but also r.

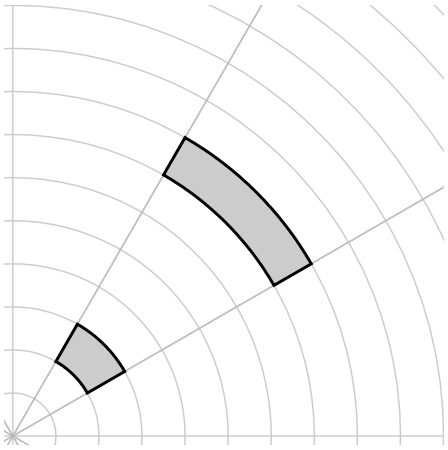


Figure 12.6.4 Polar rectangles with the same $\Delta\theta$ and Δr at different distances from the origin have different areas

Before stating and justifying the formula for dA in polar coordinates, it's worth also thinking about what the units on dA are. They need to be units of area, so suppose that x and y are both measured in centimeters. Then $dA = dx\,dy$ in rectangular coordinates has units square centimeters, which is an area unit. However, in polar coordinates, θ is measured in radians. Conventionally, we treat quantities measured in radians as lacking units, so $d\theta\,dr$ would have units of centimeters, which is a unit of length, not area.

Key Idea 12.6.5 *In polar coordinates, the area element is*

$$dA = r\,dr\,d\theta.$$

Proof. Remember that the Classic Calculus Approach to defining the double integral involved breaking each coordinate into smaller pieces which created a collection of smaller (rectangular) regions that we used to create our Riemann sum. We can do the same procedure in terms of polar coordinates, and we could have our smaller pieces be in terms of some small changes in r and θ, such as the regions on the polar grid in Figure 12.6.3. Notice however that, as we showed in Figure 12.6.4, even though the changes in r and θ used to define the grid are all the same, a region farther from the origin has larger area than one closer to the origin. This suggests that understanding how to measure the area of a piece described by $R_0 \le r \le R_1$ and $\theta_0 \le \theta \le \theta_1$, as shown in Figure 12.6.7, will be more complicated than in rectangular coordinates.

We first consider the area of the annulus between $r = R_0$ and $r = R_1$, shown in Figure 12.6.6. The area of this region is the difference between the area of a disk of radius R_1 and the area of a disk of radius R_0:

$$A(r = R_1) - A(r = R_0) = \pi R_1^2 - \pi R_0^2 = \pi\left(R_1^2 - R_0^2\right)$$

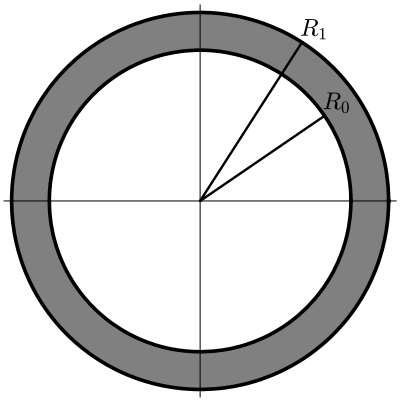


Figure 12.6.6 The annular region in the plane bounded by the circles of radius R_0 and R_1.

We need to measure the area of the portion of the annulus satisfying $\theta_0 \leq \theta \leq \theta_1$. Let $\Delta\theta = \theta_1 - \theta_0$ be the change in θ for the small piece for which we want to find the area. Because we are using radian measures for angles, the fraction of the annulus we are considering is $\frac{\Delta\theta}{2\pi}$. This means that the area of this small piece is

$$\Delta A = \frac{\Delta\theta}{2\pi}\left(\pi\left(R_1^2 - R_0^2\right)\right).$$

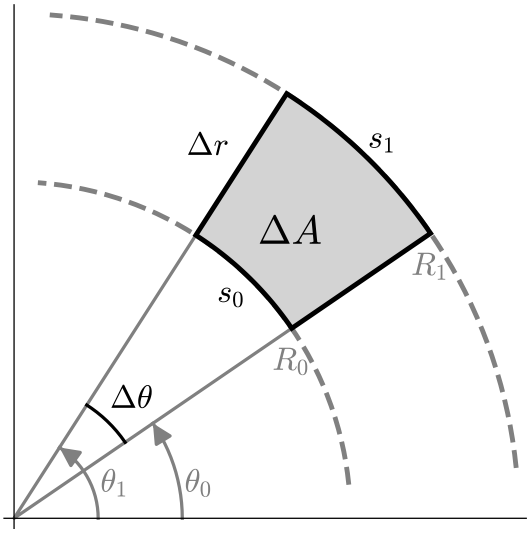


Figure 12.6.7 An area described by $R_0 \leq r \leq R_1$ and $\theta_0 \leq \theta \leq \theta_1$

Now we can reorganize our expression for ΔA by writing $R_1^2 - R_0^2$ as the difference of two squares. This gives

$$\begin{aligned}\Delta A &= \frac{\Delta\theta}{2\pi}\left(\pi(R_1 + R_0)(R_1 - R_0)\right) \\ &= \Delta\theta\left(\frac{R_1 + R_0}{2}\right)\Delta r,\end{aligned}$$

where $\Delta r = R_1 - R_0$. Note that as Δr gets smaller, R_1 and R_0 get closer to the same value, which means that

$$\Delta A \approx R_1 \Delta r\, \Delta\theta.$$

Since we can easily convert the function $f(x,y)$ to polar coordinates with the conversion equations to be $f(r\cos(\theta), r\sin(\theta))$, we can then approximate a double integral $\iint_D f(x,y)\,dA$ with a Riemann sum of the form

$$\begin{aligned}\iint_D f(x,y)\,dA &= \sum_{j=1}^{n}\sum_{i=1}^{m} f(r_i^*\cos(\theta_j^*), r_i^*\sin(\theta_j^*))\ \Delta A \\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} f(r_i^*\cos(\theta_j^*), r_i^*\sin(\theta_j^*))\ r_i^*\,\Delta r\,\Delta\theta.\end{aligned}$$

Notice that here, instead of the $\Delta x\,\Delta y$ that we had in a Riemann sum using rectangular coordinates, our integrand is multiplied by $r_i^*\,\Delta r\,\Delta\theta$. This leads us to the realization that we need to use $dA = r\,dr\,d\theta$ when writing a double integral as an iterated integral in polar coordinates. ■

Returning briefly to our discussion of units, notice that $dA = r\,dr\,d\theta$ has units of square centimeters if x and y (and thus r) are measured in centimeters. This is an area unit, as desired. Before getting into describing more complicated regions in polar coordinates, we return to the example of evaluating $\iint_D e^{x^2+y^2}\,dA$ over the unit disc D.

Example 12.6.8 Let $f(x,y) = e^{x^2+y^2}$ on the disk $D = \{(x,y) : x^2+y^2 \le 1\}$. Our goal in this example is to evaluate $\displaystyle\iint_D f(x,y)\,dA$.

In rectangular coordinates the double integral $\iint_D f(x,y)\,dA$ can be written as the iterated integral

$$\iint_D f(x,y)\,dA = \int_{x=-1}^{x=1}\int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} e^{x^2+y^2}\,dy\,dx,$$

where the vertically simple slicing is shown in Figure 12.6.1

We cannot evaluate this iterated integral because $e^{x^2+y^2}$ does not have an elementary antiderivative with respect to either x or y. However, since $r^2 = x^2+y^2$ and the region D is circular, it is natural to wonder whether converting to polar coordinates will allow us to evaluate the new integral. To do so, we replace x with $r\cos(\theta)$, y with $r\sin(\theta)$, and $dy\,dx$ with $r\,dr\,d\theta$ to obtain

$$\iint_D f(x,y)\,dA = \iint_D e^{r^2}\,r\,dr\,d\theta.$$

The disc D is described in polar coordinates by the inequalities $0 \le r \le 1$ and $0 \le \theta \le 2\pi$. Therefore, it follows that

$$\iint_D e^{r^2}\,r\,dr\,d\theta = \int_{\theta=0}^{\theta=2\pi}\int_{r=0}^{r=1} e^{r^2}\,r\,dr\,d\theta.$$

We can evaluate the resulting iterated polar integral as follows:

$$\begin{aligned}\int_{\theta=0}^{\theta=2\pi}\int_{r=0}^{r=1} e^{r^2}\,r\,dr\,d\theta &= \int_{\theta=0}^{2\pi}\left(\frac{1}{2}e^{r^2}\Big|_{r=0}^{r=1}\right)d\theta \\ &= \frac{1}{2}\int_{\theta=0}^{\theta=2\pi}(e-1)\,d\theta \\ &= \frac{1}{2}(e-1)\int_{\theta=0}^{\theta=2\pi} d\theta \\ &= \frac{1}{2}(e-1)\,[\theta]\Big|_{\theta=0}^{\theta=2\pi}\end{aligned}$$

$$= \pi(e - 1).$$

▲

We now focus on the second conversion step: convert the region of integration described by inequalities that correspond to fixing one coordinate and bounding the other in a consistent way. Similar to how we used vertically simple and horizontally simple descriptions for rectangular regions, we want to set up inequalities by holding one of our coordinates constant and setting boundary functions for the other coordinate.

Key Idea 12.6.9 *A region is called* ***radially simple*** *if it can be described by a set of inequalities of the form*

$$a \leq \theta \leq b$$
$$g_1(\theta) \leq r \leq g_2(\theta),$$

corresponding geometrically to the region in Figure 12.6.10

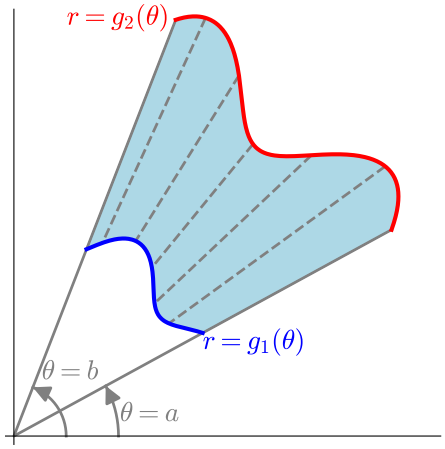


Figure 12.6.10 *A radially simple region with slices along constant θ values shown in gray*

A region is called ***angularly simple*** *if it can be described by a set of inequalities of the form*

$$c \leq r \leq d$$
$$h_1(r) \leq \theta \leq h_2(r),$$

corresponding geometrically to the region in Figure 12.6.11

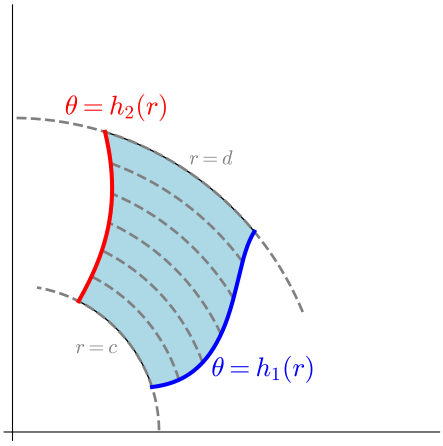


Figure 12.6.11 *An angularly simple region with slices along constant r values shown in gray*

Note that the unit disc D we have considered throughout this section can be considered both radially simple and angularly simple since all of the bounds of integration are constant. Figure 12.6.12 shows how D can be considered radially simple.

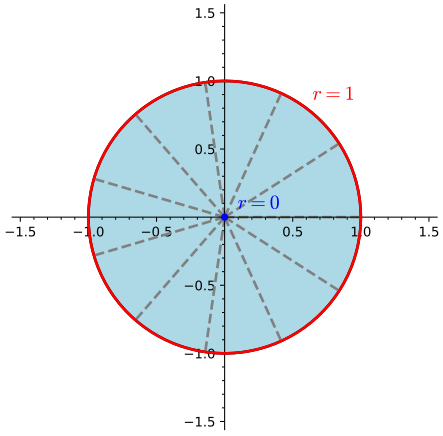


Figure 12.6.12 The unit disk centered at the origin with slices showing a radially simple description

Regions that are angularly simple but *not* radially simple do not arise with much frequency. Thus, while we have defined both types of simple regions for completeness, we will not spend time on an example of a region that is angularly simple but not radially simple.

Each of the types of regions above corresponds to exactly the necessary description to use an iterated integral. With a radially simple region, for any fixed value of θ in the interval $[a, b]$, a slice along this constant θ value corresponds to a ray from the origin in the sector between the grey lines in Figure 12.6.10. Along this ray with $\theta = k$, the region of interest has upper and lower bounds on the r-coordinate. In fact, $r = g_2(\theta)$ describes the upper boundary for every value of θ in $[a, b]$. Similarly, $r = g_1(\theta)$ describes the lower boundary for every value of θ in $[a, b]$. Geometrically, this means that along every gray dashed slice where θ is constant, the region has the same upper and lower boundary curves, shown in red and blue, respectively.

While angularly simple regions are used less frequently than radially simple regions, a parallel structure is used. Along every slice with a constant value of

r in $[c, d]$, represented by the gray, dashed curves in Figure 12.6.11, the upper bound on θ is given by the function $\theta = h_2(r)$ and the lower bound is given by $\theta = h_1(r)$, shown by the red and blue bounding curves in Figure 12.6.11.

Just as in the rectangular coordinate case, especially complicated regions may need to be split into several different regions that are radially or angularly simple.

Double integrals in polar coordinates.

The double integral $\iint_D f(x, y)\, dA$ in rectangular coordinates can be converted to a double integral in polar coordinates as $\iint_D f(r\cos(\theta), r\sin(\theta))\, r\, dr\, d\theta$.

If the region D is radially simple with $a \le \theta \le b$ and $g_1(\theta) \le r \le g_2(\theta)$, then

$$\iint_D f(x, y)\, dA = \int_a^b \int_{g_1(\theta)}^{g_2(\theta)} f(r\cos(\theta), r\sin(\theta))\, r\, dr\, d\theta.$$

If the region D is angularly simple with $c \le r \le d$ and $h_1(r) \le \theta \le h_2(r)$, then

$$\iint_D f(x, y)\, dA = \int_c^d \int_{h_1(r)}^{h_2(r)} f(r\cos(\theta), r\sin(\theta))\, r\, d\theta\, dr.$$

Note again that the order of integration corresponds to the manner in which we have sliced the region of integration. The variable being held constant along a slice is the last variable to be integrated, just as was the case with iterated integrals in rectangular coordinates.

Example 12.6.13 Suppose that we have a piece of metal in the shape of a quarter annulus R, as shown in Figure 12.6.14. The circular arcs bounding the region have radius 2 and 3, and the region lies in the first quadrant. The metal is heated so that the temperature in degrees Celsius at the point (x, y) is given by $T(x, y) = \sqrt{x^2 + y^2}$. That is, the temperature is 12 times the distance between the origin and the point. In this example, we will find the *average value* of T over the region R.

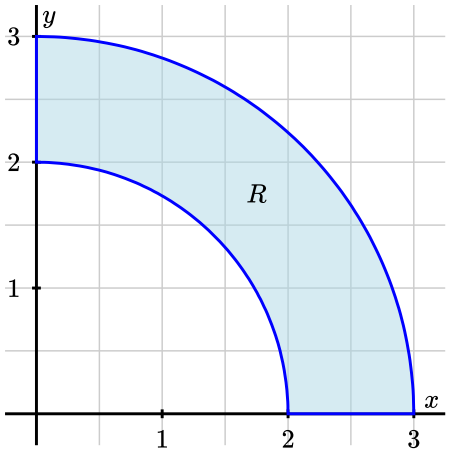


Figure 12.6.14 A planar region representing a metal plate in the shape of one quarter of the annulus between circles of radius 2 and 3

To find the average value of T on R, we must compute

$$\frac{1}{\text{Area}(R)} \iint_R T\, dA.$$

We can compute the area of R using geometry, as its area is one quarter the area of the annulus between the circles of radius 2 and 3 centered at the origin. That annulus has area $\pi \cdot 3^2 - \pi \cdot 2^2 = 5\pi$, so R has area $5\pi/4$. To compute the double integral, we will use an iterated integral. This will require finding inequalities that allow us to describe R. Notice first that R is neither horizontally simple nor vertically simple, so evaluating $\iint_R T\, dA$ rectangular coordinates would require that we set up multiple iterated integrals. However, in polar coordinates, we see that R is both radially simple and angularly simple. That is, R is a polar rectangle, which we can describe as

$$2 \le r \le 3 \qquad 0 \le \theta \le \frac{\pi}{2}.$$

Using this description of R in polar coordinates along with the facts that $r^2 = x^2 + y^2$ and $dA = r\, dr, d\theta$, we can evaluate the double integral:

$$\begin{aligned}
\iint_R T(x,y)\, dA &= \iint_R 12\sqrt{x^2+y^2}\, dA \\
&= \int_0^{\pi/2} \int_2^3 12r \cdot r\, dr\, d\theta \\
&= \int_0^{\pi/2} 4r^3 \Big|_2^3 d\theta \\
&= \int_0^{pi/2} 76\, d\theta = 76\theta \Big|_0^{\pi/2} = 38\pi.
\end{aligned}$$

Therefore, the average temperature of the metal plate is

$$\frac{1}{\frac{5\pi}{4}} 38\pi = \frac{4}{5\pi} \cdot 38\pi = \frac{152}{5} = 30.4^\circ \text{ C}.$$

▲

While there is no firm rule for when polar coordinates can or should be used, they are a natural alternative anytime the domain of integration may be expressed simply in polar form and/or when the integrand involves expressions such as $\sqrt{x^2+y^2}$.

Activity 12.6.2 In this activity, we will find the volume of the solid region V above the xy-plane, below the paraboloid $z = f(x,y) = 25 - x^2 - y^2$, and outside the cylinder $x^2 + y^2 = 1$.

(a) We will find the volume by computing a double integral $\iint_R f(x,y)\, dA$ where R is an appropriate region in the xy-plane. Sketch the region R.

(b) Write inequalities to describe R in polar coordinates. Write a couple of sentences explaining why it would be very difficult to describe R using rectangular coordinates.

(c) Use an interated integral in polar coordinates to find the volume of V.

The next activity asks you to find the volume between two surfaces in three-dimensional space. To do this, you will find it useful to think of the functions involved in polar coordinates from the very beginning.

Activity 12.6.3 Use polar coordinates to find the volume of the solid between the cone $z = f(x,y) = \sqrt{x^2+y^2}$ and the paraboloid $z = g(x,y) = \frac{1}{8}\left(x^2+y^2\right)+2$.

Perhaps the surfaces in Activity 12.6.3 brought thoughts of cylindrical coordinates to your mind. The next section will generalize the idea of a double integral to a triple integral, and after that, we will consider integration using cylindrical and spherical coordinates as well.

Summary

- The area element dA in polar coordinates is determined by the area of a slice of an annulus and is given by
$$dA = r\,dr\,d\theta$$
- To convert the double integral $\iint_D f(x,y)\,dA$ to an iterated integral in polar coordinates, we substitute $r\cos(\theta)$ for x, $r\sin(\theta)$ for y, and $r\,dr\,d\theta$ for dA to obtain the iterated integral (where the order of integration is determined by whether the region is described as radially or angularly simple)
$$\iint_D f(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta$$

12.6.3 Exercises

You can find many more problems involving the conversion of points, graphs, and region between rectangular and polar coordinates in Section 9.9. The exercises of this section will be focused on applications of these conversion to double integrals.

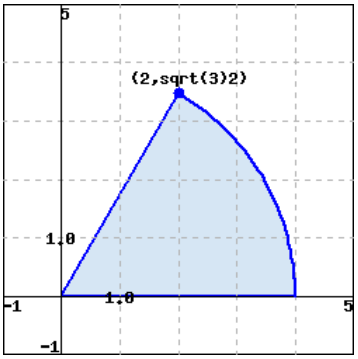
1. Sketch the region of integration for the following integral.
$$\int_0^{\pi/4}\int_0^{6/\cos(\theta)} f(r,\theta)\,r\,dr\,d\theta$$
The region of integration is bounded by

- ◯ $y = 0, y = x,$ and $y = 6$
- ◯ $y = 0, y = x,$ and $x = 6$
- ◯ $y = 0, y = \sqrt{36-x^2},$ and $x = 6$
- ◯ $y = 0, x = \sqrt{36-y^2},$ and $y = 6$
- ◯ None of the above

2. Convert the integral
$$\int_0^{\sqrt{10}}\int_{-x}^{x} dy\,dx$$
to polar coordinates and evaluate it (use t for θ):
With $a =$ _____, $b =$ _____, $c =$ _____ and $d =$ _____,
$\int_0^{\sqrt{10}}\int_{-x}^{x} dy\,dx = \int_a^b\int_c^d$ _____$dr\,dt$
$= \int_a^b$ _____ dt
$=$ _____$\Big|_a^b$
$=$ _____.

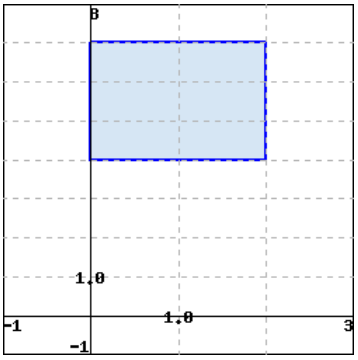
3. For each of the following, set up the integral of an arbitrary function $f(x, y)$ over the region in whichever of rectangular or polar coordinates is most appropriate. (Use t for θ in your expressions.)

(a) The region



With $a =$ _____, $b =$ _____,
$c =$ _____, and $d =$ _____,
integral $= \int_a^b \int_c^d$ _____ d_____ d_____

(b) The region



With $a =$ _____, $b =$ _____,
$c =$ _____, and $d =$ _____,
integral $= \int_a^b \int_c^d$ _____ d_____ d_____

4. By changing to polar coordinates, evaluate the integral

$$\iint_D (x^2 + y^2)^{7/2}\, dxdy$$

where D is the disk $x^2 + y^2 \leq 4$.

Answer = _____

5. Using polar coordinates, evaluate the integral which gives the area which lies in the first quadrant between the circles $x^2+y^2 = 36$ and $x^2-6x+y^2 = 0$.

6. ***(a)*** Graph $r = 1/(4\cos\theta)$ for $-\pi/2 \leq \theta \leq \pi/2$ and $r = 1$. Then write an iterated integral in polar coordinates representing the area inside the curve $r = 1$ and to the right of $r = 1/(4\cos\theta)$. (Use t for θ in your work.)

With $a =$_____, $b =$_____,
$c =$_____, and $d =$_____,
area $= \int_a^b \int_c^d$ _____d_____ d_____

(b) Evaluate your integral to find the area.

area = _____

7. Using polar coordinates, evaluate the integral $\iint_R \sin(x^2+y^2)dA$ where R is the region $4 \le x^2+y^2 \le 81$.

8. Use the polar coordinates to find the volume of a sphere of radius 6.

9. Consider the solid under the graph of $z = e^{-x^2-y^2}$ above the disk $x^2+y^2 \le a^2$, where $a > 0$.

(a) Set up the integral to find the volume of the solid.

Instructions: Please enter the integrand in the first answer box, typing *theta* for θ. Depending on the order of integration you choose, enter *dr* and *dtheta* in either order into the second and third answer boxes with only one *dr* or *dtheta* in each box. Then, enter the limits of integration.

$\int_A^B \int_C^D$ _____ _____ _____

A = _____

B = _____

C = _____

D = _____

(b) Evaluate the integral and find the volume. Your answer will be in terms of a.

Volume V = _____

(c) What does the volume approach as $a \to \infty$?

$\lim_{a\to\infty} V$ = _____

10. Sketch the region whose area is given by the integral and evaluate it.

$$\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} \int_3^8 r\,dr\,d\theta$$

Answer: _____

11. Sketch the region whose area is given by the integral and evaluate it.

$$\int_0^{\pi/2} \int_0^{8\sin\theta} r\,dr\,d\theta$$

Answer: _____

12. Using polar coordinates, evaluate the integral $\iint_R \sin(x^2+y^2)dA$ where R is the region $16 \le x^2+y^2 \le 25$.

13. A cylindrical drill with radius 1 is used to bore a hole through the center of a sphere of radius 6. Find the volume of the ring shaped solid that remains.

14. Evaluate the double integral $\iint_D x^2 y\,dA$, where D is the top half of the disc with center the origin and radius 4, by changing to polar coordinates.

Answer: _____

15. Evaluate the double integral $\iint_R (2x-y)\,dA$, where R is the region in the first quadrant enclosed by the circle $x^2+y^2=4$ and the lines $x=0$ and $y=x$, by changing to polar coordinates.

Answer: _____

16. Use polar coordinates to find the volume of the solid under the cone $z = \sqrt{x^2 + y^2}$ and above the disc $x^2 + y^2 \leq 16$.
Answer: _____

17. Use polar coordinates to find the volume of the solid below the paraboloid $z = 48 - 3x^2 - 3y^2$ and above the xy-plane.
Answer: _____

18. A volcano fills the volume between the graphs $z = 0$ and $z = \dfrac{1}{(x^2 + y^2)^{21}}$, and outside the cylinder $x^2 + y^2 = 1$. Find the volume of this volcano.

19. Consider the iterated integral $I = \int_{-3}^{0} \int_{-\sqrt{9-y^2}}^{0} \frac{y}{x^2+y^2+1}\, dx\, dy$.

a. Sketch (and label) the region of integration.

b. Convert the given iterated integral to one in polar coordinates.

c. Evaluate the iterated integral in (b).

d. State one possible interpretation of the value you found in (c).

20. Let D be the region that lies inside the unit circle in the plane.

a. Set up and evaluate an iterated integral in polar coordinates whose value is the area of D.

b. Determine the exact average value of $f(x, y) = y$ over the upper half of D.

c. Find the exact center of mass of the lamina over the portion of D that lies in the first quadrant and has its mass density distribution given by $\delta(x, y) = 1$. (Before making any calculations, where do you expect the center of mass to lie? Why?)

d. Find the exact volume of the solid that lies under the surface $z = 8 - x^2 - y^2$ and over the unit disk, D.

21. For each of the following iterated integrals,

- sketch and label the region of integration,
- convert the integral to the other coordinate system (if given in polar, to rectangular; if given in rectangular, to polar), and
- choose one of the two iterated integrals to evaluate exactly.

a. $\displaystyle\int_{\pi}^{3\pi/2} \int_{0}^{3} r^3\, dr\, d\theta$

b. $\displaystyle\int_{0}^{2} \int_{-\sqrt{1-(x-1)^2}}^{\sqrt{1-(x-1)^2}} \sqrt{x^2 + y^2}\, dy\, dx$

c. $\displaystyle\int_{0}^{\pi/2} \int_{0}^{\sin(\theta)} r\sqrt{1 - r^2}\, dr\, d\theta.$

d. $\displaystyle\int_{0}^{\sqrt{2}/2} \int_{y}^{\sqrt{1-y^2}} \cos(x^2 + y^2)\, dx\, dy.$

22. Consider the circle given by $x^2 + (y - 1)^2 = 1$ as shown in Figure 12.6.15.

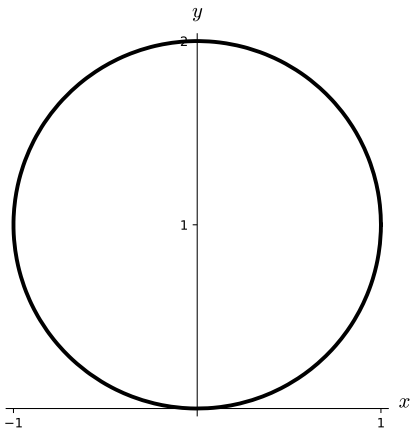


Figure 12.6.15 The circle $x^2 + (y-1)^2 = 1$

(a) Determine a polar curve in the form $r = f(\theta)$ that traces out the circle $x^2+(y-1)^2 = 1$. To do this, you should substitute $x = r\cos(\theta)$ and $y = r\sin(\theta)$ into the rectangular coordinate equation and solve for r.

Hint. After you substitute, you should expand and combine like terms. You will also want to factor your expression and disregard the $r = 0$ solution.

(b) Find the exact average value of $g(x, y) = \sqrt{x^2 + y^2}$ over the interior of the circle $x^2 + (y-1)^2 = 1$.

Hint. Use a radially simple description of your region with your result from part a as the upper bound.

(c) Find the volume under the surface $h(x, y) = x$ over the region D, where D is the region bounded above by the line $y = x$ and below by the circle. We have provided a sketch of this as the shaded region in Figure 12.6.16.

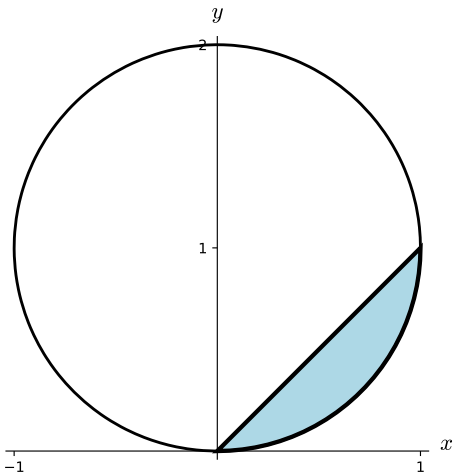


Figure 12.6.16 The region between the circle $x^2+(y-1)^2 = 1$ and the line $y = x$.

(d) Explain why in both of the previous parts it is advantageous to use polar coordinates.

12.7 Triple Integrals

Motivating Questions

- How do the ideas of Riemann sums, integrals, and interpretations of integrals generalize to functions of three variables?
- How can iterated integrals be used to efficiently calculate integrals involving three variables?

12.7.1 Introduction

In this chapter, we defined the double integral of a continuous function $f = f(x, y)$ over a rectangle $R = [a, b] \times [c, d]$ as a limit of a double Riemann sum, which paralleled the definition of a single-variable integral for a function $g = g(x)$ on an interval $[a, b]$. We have also repeatedly emphasized the interpretations and applications of the double integral as stated in Key Idea 12.2.8. These ideas will naturally extend to functions with more than two variables, but the geometric elements of these extensions become harder to visualize. In the Preview Activity below, we will setup a Riemann sum as an approximation to integrating a function of three variables (step 1 of a classic calculus approach).

Preview Activity 12.7.1 In this activity, we want to try to estimate the mass of large piece of granite. Granite is composed of different minerals such as feldspar and quartz that give distinctive patterns of color and texture. The large piece of granite we are looking at is 4 feet wide, six feet deep, and 8 feet tall which we will describe by

$$B = \{(x, y, z) : 0 \leq x \leq 4, 0 \leq y \leq 6, 0 \leq z \leq 8\}.$$

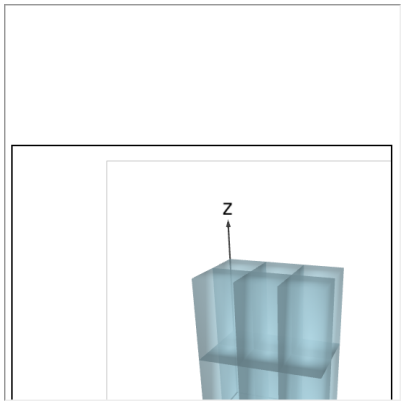
This very special piece of granite formed in a region with many geological folds and has its density given by $\delta(x, y, z) = 163 + 3\sin(xy) + \frac{2z}{3}$. The units for x, y, z are measured in feet and the density is given in pounds per cubic foot.

For a solid of constant density, we can find the mass by multiplying the density and volume. For our block of granite, the density varies from point to point based on the function $\delta(x, y, z)$. In this activity, we will approximate the mass of the block (step 1 of the classic calculus approach) by slicing the solid into smaller pieces on which there are smaller density changes and thus the density is closer to constant.

(a) For a first approximation using smaller pieces, we partition the block as follows

- the width, given by the x-interval $[0, 4]$, into two subintervals of equal length
- the depth, given by the y-interval $[0, 6]$ into three subintervals of equal length
- the height, given by the z-interval $[0, 8]$ into two subintervals of equal length

This partitions the box B into sub-boxes as shown in Figure 12.7.1.



Standalone
In Context
Embed

Figure 12.7.1 A partitioned three-dimensional domain

Let $0 = x_0 < x_1 < x_2 = 4$ be the endpoints of the x-subintervals of $[0, 4]$ after partitioning. Label these endpoints on Figure 12.7.2. Repeat this process with $0 = y_0 < y_1 < y_2 < y_3 = 6$ and $0 = z_0 < z_1 < z_2 = 8$.

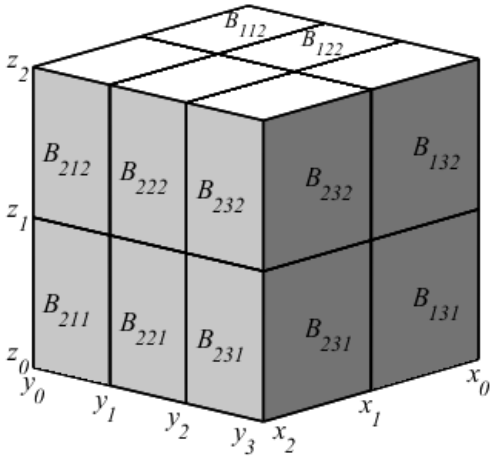


Figure 12.7.2 A partitioning of our granite block

(b) Our partitions of the x, y, z-intervals created sub-boxes for our granite block. How many sub-boxes were created and what is the volume of the sub-boxes?

(c) For each sub-box, pick a point in that region and evaluate the density function at that point. Make sure you note what point you chose and the density at that point.

(d) For each sub-box, multiply the density at your chosen point by the volume of the sub-box to get an estimate of the mass for the sub-box. Sum over all of your sub-boxes to give an estimate of the mass for the entire block, which will complete step 1 of the classic calculus approach to measuring the total mass of our granite block.

(e) We partitioned our block, the space of inputs for $\delta(x, y, z)$, into a relatively small number of pieces. Write a couple of sentences about what

you would change in the process above to improve the approximation of the mass of the block. Be sure to explain why your changes will give a better approximation.

In the Preview Activity, we completed step 1 and justified step 2 of the, classic calculus approach toward integrating a density function in three dimensions. The generalization of our work in the previous part of this chapter on double integrals works much as you would expect, so we be brief in our statements of these generalizations and will spend the rest of this section getting used to the mechanics and geometric arguments that are typical in these types of problems.

12.7.2 Triple Riemann Sums and Triple Integrals

Through the application of a mass density distribution over a three-dimensional solid block of granite, Preview Activity 12.7.1 suggests a natural generalization from double Riemann sums of functions of two variables to triple Riemann sums of functions of three variables. In the same way, we can generalize from double integrals to triple integrals. By simply adding a z-coordinate to our earlier work, we can define both a triple Riemann sum and the corresponding triple integral.

Definition 12.7.3 Let $f = f(x, y, z)$ be a continuous function on a box $B = [a, b] \times [c, d] \times [r, s]$. The **triple Riemann sum of f over B** is created as follows.

- Partition the x-interval $[a, b]$ into m subintervals of equal length, the y-interval $[c, d]$ into n subintervals of equal length, and the z-interval $[r, s]$ into ℓ subintervals of equal length.
- For each sub-box created by this partition, let P_{ijk} be a point in the i-th segment of x, the j-th segment of y, and k-th segment of z. The volume of each sub-box is $\frac{1}{mn\ell}$ of the total volume.
- The resulting triple Riemann sum for f on B is

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{\ell} f(P_{ijk})\,\Delta V.$$

◆

If $f(x, y, z)$ represents the density of a material in the box B, then as we saw in Preview Activity 12.7.1, the triple Riemann sum approximates the total mass of material in the box B as step 1 of our classic calculus approach. In order to find the exact mass of the box, we will need to let the number of sub-boxes increase without bound (in other words, let m, n, and ℓ go to infinity). When we do this, the finite sum of the mass approximations becomes the exact mass of the solid B. More generally, we have the following definition of the triple integral.

Definition 12.7.4 Using the notation introduced for a triple Riemann sum, the **triple integral of f over B** is

$$\iiint_B f(x, y, z)\,dV = \lim_{m,n,\ell\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{\ell} f(P_{ijk})\,\Delta V.$$

◆

In the above description, we used equally spaced division of each of the coordinates to separate the region of integration into pieces. As the number of pieces increases (m, n, and ℓ go to infinity), the volume of each piece (ΔV) will go to zero. A more general treatment of this process will show that the pieces do not need to be the same size, but the size of all of the pieces must approach zero in the limit used for the triple Riemann sum.

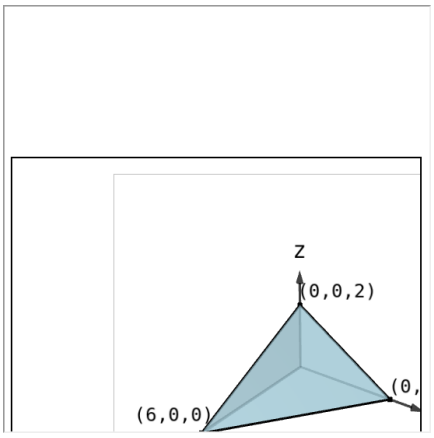
Just as we did with double integrals, we will use a series of iterated integrals to efficiently evaluate our triple integrals. As you might expect, the volume element dV is $dz\,dy\,dx$ in rectangular coordinates, since we measure the volume of a box as the height times the width times the depth. Thus, a triple integral of a function f over a box $B = [a, b] \times [c, d] \times [r, s]$ in rectangular coordinates can be evaluated as an iterated integral of the form

$$\iiint_B f(x, y, z)\, dV = \int_a^b \int_c^d \int_r^s f(x, y, z)\, dz\, dy\, dx.$$

Activity 12.7.2 Set up and evaluate an iterated integral that will evaluate the triple integral of $f(x, y, z) = x - y + 2z$ over the box $B = [-2, 3] \times [1, 4] \times [0, 2]$.

Remember that it was our descriptions of two dimensional regions of integration using inequalities corresponding to horizontally or vertically simple regions that allowed us to convert double integrals into iterated integrals. We will generalize this process to three dimensional regions of integration in order to convert triple integrals into three iterated integrals where the order in which state the bounds on each coordinate corresponds to the order in which variables are integrated. In the next example, we will describe the process of setting up these inequalities, then using three iterated integrals to calculate the triple integral of a three variable function over a region that does not have constant coordinate bounds.

Example 12.7.5 In this example, we will find the mass of the tetrahedron S in the first octant bounded by the coordinate planes and the plane $x+2y+3z = 6$ if the density at point (x, y, z) is given by $\delta(x, y, z) = x + y + z$. A picture of the solid tetrahedron is shown in Figure 12.7.6.



Standalone
In Context
Embed

Figure 12.7.6 The tetrahedron S in the first octant bounded by the coordinate planes and the plane $x + 2y + 3z = 6$

We find the mass M of the tetrahedron using the triple integral

$$M = \iiint_S \delta(x, y, z)\, dV.$$

To do this, we will need to generalize our ideas from Section 12.4 and describe S using three sets of inequalities, one for each variable. In this example, we

choose to integrate with respect to z first for the innermost integral. We will first need to consider how to give bounds on the "top" and "bottom" functions for z as a function of x and y. In other words, we need to give functions $z = g(x, y)$ and $z = h(x, y)$ such that for any (x, y) point in our region, $g(x, y)$ will give the largest z-value we need to consider and $h(x, y)$ will need to give the smallest z-value. This description will give us an iterated integral of the form

$$\iint_D \left[\int_{z=h(x,y)}^{z=g(x,y)} \delta(x, y, z)\, dz \right] dA$$

Note that the inner integral will be considered with x and y held constant and D being the set of (x, y) points over which our three dimensional region of integration sits.

You can see from Figure 12.7.6 that the plane containing the points $(6, 0, 0), (0, 3, 0), (0, 0, 2)$ will give us $z = g(x, y)$ and the xy-plane will give us $z = h(x, y)$. So we have

$$\iint_D \left[\int_{z=0}^{z=2-\frac{2}{3}y-\frac{x}{6}} \delta(x, y, z)\, dz \right] dA$$

as our first iterated integral. We now need to consider D, the region of the xy-plane over which our region S sits. In this example, D coincides with the triangle in the xy plane with vertices $(0, 0), (6, 0), (0, 3)$, as drawn in Figure 12.7.7.

In order to complete our transformation to iterated integrals, we need to describe D as either horizontally simple or vertically simple. We can see that D is both vertically simple and horizontally simple, so we could use either description. In this example, we chose to describe the region as vertically simple, as suggested by the dashed lines in Figure 12.7.7. We will cut our region into vertical slices for $0 \leq x \leq 6$. Furthermore, the lower bound on each slice is $y = 0$, so we just need to find the equation of the top boundary of the region. We can find the equation of the line that determines this top boundary as $x + 2y = 6$. Solving for y gives $y = 3 - \frac{1}{2}x$. Therefore, we can describe the base of the tetrahedron as a vertically simple region using the inequalities

$$0 \leq x \leq 6 \qquad 0 \leq y \leq 3 - \frac{1}{2}x.$$

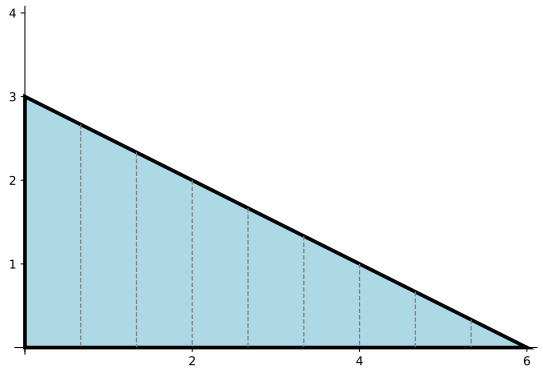


Figure 12.7.7 The projection of our tetrahedron to the xy-plane, shown with vertical slices

We can now combine our vertically simple description of D with our iterated integral above to get

$$M = \iint_D \left[\int_{z=h(x,y)}^{z=g(x,y)} \delta(x, y, z)\, dz \right] dA$$

$$= \int_{x=0}^{x=6} \int_{y=0}^{y=3-\frac{1}{2}x} \left[\int_{z=h(x,y)}^{z=g(x,y)} \delta(x,y,z)\, dz \right] dy\, dx$$

We can restate our work above as a description of the solid S using the inequalities

$$0 \le x \le 6 \qquad 0 \le y \le 3 - \frac{1}{2}x \qquad 0 \le z \le \frac{1}{3}(6 - x - 2y).$$

With our description of S in terms of inequalities in hand, we can write an iterated triple integral to find the mass of the tetrahedron by integrating the density function $\delta(x,y,z) = x + y + z$:

$$M = \int_0^6 \int_0^{3-(1/2)x} \int_0^{(1/3)(6-x-2y)} (x + y + z)\, dz\, dy\, dx. \tag{12.7.1}$$

Evaluating the three iterated integrals yields

$$\begin{aligned}
M &= \int_0^6 \int_0^{3-(1/2)x} \int_0^{(1/3)(6-x-2y)} (x + y + z)\, dz\, dy\, dx \\
&= \int_0^6 \int_0^{3-(1/2)x} \left[xz + yz + \frac{z^2}{2} \right] \Big|_0^{(1/3)(6-x-2y)} dy\, dx \\
&= \int_0^6 \int_0^{3-(1/2)x} \left(\frac{4}{3}x - \frac{5}{18}x^2 - \frac{7}{9}xy + \frac{2}{3}y - \frac{4}{9}y^2 + 2 \right) dy\, dx \\
&= \int_0^6 \left[\frac{4}{3}xy - \frac{5}{18}x^2 y - \frac{7}{18}xy^2 + \frac{1}{3}y^2 - \frac{4}{27}y^3 + 2y \right] \Big|_0^{3-(1/2)x} dx \\
&= \int_0^6 \left(5 + \frac{1}{2}x - \frac{7}{12}x^2 + \frac{13}{216}x^3 \right) dx \\
&= \left[5x + \frac{1}{4}x^2 - \frac{7}{36}x^3 + \frac{13}{864}x^4 \right] \Big|_0^6 \\
&= \frac{33}{2}.
\end{aligned}$$

▲

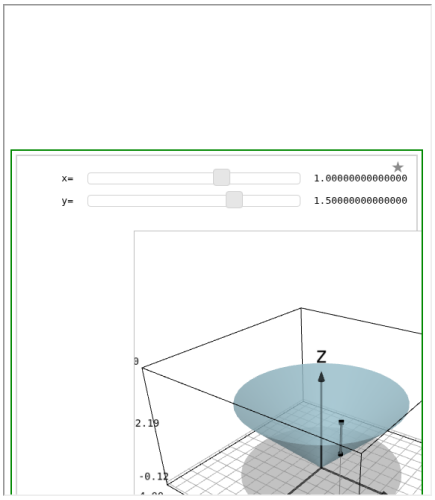
The next activity asks you to find the mass of a solid region that does not have a base lying in the xy-plane.

Activity 12.7.3 Let S be the solid cone bounded by $z = \sqrt{x^2 + y^2}$ and $z = 3$. The goal of this activity is to set up an iterated integral of the form

$$\int_{x=?}^{x=?} \int_{y=?}^{y=?} \int_{z=?}^{z=?} \delta(x,y,z)\, dz\, dy\, dx \tag{12.7.2}$$

to represent the mass of S in the setting where $\delta(x,y,z)$ tells us the density of S at the point (x,y,z). In particular, we must find the limits on each of the three integrals.

A picture of S is shown in Figure 12.7.8. Adjust the sliders, which will move the vertical line segment shown inside the solid. Notice that some combinations of slider values show a red point and no line segment, as the point (x,y) does not correspond to any points inside the solid.



Standalone
In Context
Embed

Figure 12.7.8 The region S with a vertical segment highlighted for fixed x and y values

(a) For the innermost integral of equation (12.7.2), we need bounds on the z-coordinate for fixed values of x and y. In Figure 12.7.8, you can use the sliders to change the values of x and y. When your choices of x and y correspond to points inside the solid, you see a vertical line segment in the plot from the bottom of the solid to the top of the solid over the point (x, y) in the xy-plane.

Try several values of x and y. Look at how the length of the segment changes in the z-direction. In particular, *for every x and y pair*, the bottom boundary of the solid is the same. Similarly, for every pair of values the top boundary of the solid is the same surface. This allows you to use functions, in terms of x and y, that describe the top and bottom boundaries of the solid. These are the z-coordinates of the points at the top and bottom of the vertical line segments through the solid shown in the figure.

$$z_{\text{top}}(x, y) = _____$$
$$z_{\text{bottom}}(x, y) = _____$$

(b) Complete the compound inequality below to provide lower and upper bounds for the z-coordinates of points in S (in terms of x and y).

$$_____ \leq z \leq _____$$

(c) Having established upper and lower bounds for z as a function of a fixed choice of x and y, we need to describe the set of points (x, y) in the xy-plane such that a vertical line through (x, y) will pass through the solid. Notice that if you choose values of x and y in Figure 12.7.8 that does not intersect the solid (e.g., $(x, y) = (2.6, -2.3)$), then the point is shown in red.

You can see from Figure 12.7.8 that there will be x and y values from -3 to 3 that will correspond to points in our solid. It is tempting to give the region of the xy-plane we need to consider using the inequalities $-3 \leq x \leq 3$ and $-3 \leq y \leq 3$. Write a couple of sentences to explain why the set of (x, y) points we need to consider is *not* the square $[-3, 3] \times [-3, 3]$.

(d) On Figure 12.7.9, draw a plot of D, the region of (x, y) points that correspond to points of S. We refer to D as the **projection** of S onto the xy-plane.

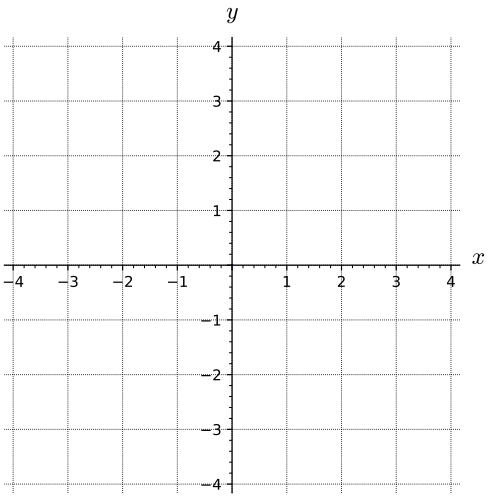


Figure 12.7.9 A grid of $[-4, 4] \times [-4, 4]$

(e) To complete the task of writing

$$\iiint_S \delta(x, y, z) dV$$

as an iterated integral, you need to describe the region D from the previous part using inequalities, as with double integrals. Do this using a *vertically simple* description in order to have your iterated integral fit the form of equation (12.7.2).

Give the inequalities that describe the region D from the previous part as vertically simple.

$$_____ \leq x \leq _____$$
$$_____ \leq y \leq _____$$

(f) Write an iterated integral of the form (12.7.2) that represents the mass of the cone S.

We did not evaluate the three iterated integrals in the previous activity because the bounds of integration stated would give iterated integrals that turn out to be difficult to do by hand. You may have noticed that the region in Activity 12.7.3 would be easily described in cylindrical coordinates and might yield more convenient bounds of integration. We will look at this type of transformation for triple integrals in Section 12.8.

Algebraic Note: When setting up iterated integrals, the limits on a given variable can be *only* in terms of constants and the variable(s) whose integral(s) are *outside* that integral in the iterated integral. In addition, there are multiple different ways we can choose to set up such an integral. For example, two possibilities for iterated integrals that represent a triple integral $\iiint_S f(x, y, z)\, dV$ over a solid S are

$$\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)\, dz\, dy\, dx$$

$$\int_r^s \int_{p_1(z)}^{p_2(z)} \int_{q_1(x,z)}^{q_2(x,z)} f(x,y,z)\,dy\,dx\,dz$$

where g_1, g_2, h_1, h_2, p_1, p_2, q_1, and q_2 are functions of the indicated variables. There are four other options beyond the two stated here, since the variables x, y, and z can, in theory, be arranged in any order. In many circumstances, an insightful choice of variable order will make it easier to set up an iterated integral, just as was the case when we worked with double integrals.

Setting up limits on iterated integrals often requires considerable geometric intuition. It is important not only to create carefully labeled figures for the region of integration, but also to think about how we wish to slice the solid region of inputs for a triple integral. Further, note that when we say "we will integrate first with respect to x," by "first" we are referring to the *innermost* integral in the iterated integral. The next activity explores several different ways we might set up the integral in Example 12.7.5.

Activity 12.7.4 There are several other ways to set up the integral to give the mass of the tetrahedron in Example 12.7.5.

(a) How many different orders of integration could be used for iterated integrals that are equal to the integral in equation (12.7.1)?

(b) Set up an iterated integral, integrating first with respect to z, then x, then y that is equivalent to the integral in equation (12.7.1). Before you write down the integral, think about Figure 12.7.6 and draw a plot of the appropriate projection.

Hint. After considering the upper and lower bound surfaces in terms of z, you will need to project onto the xy-plane and think of this projection as a horizontally simple region.

(c) Set up an iterated integral, integrating first with respect to y, then z, then x, that is equivalent to the integral in equation (12.7.1). As above, think carefully about the geometry first and draw a plot of the appropriate projection.

(d) Set up an iterated integral, integrating first with respect to x, then y, then z that is equivalent to the integral in equation (12.7.1).

As described earlier, if $f(x,y,z)$ represents the density of the solid B at each point (x,y,z), then

$$M = \iiint_B f(x,y,z)\,dV$$

is the mass of B. Even more importantly, for any continuous function f over the solid B, we can use a triple integral to determine the average value of f over B, $f_{\text{AVG}(B)}$. We note this generalization of our work with functions of two variables along with several others in Key Idea 12.7.10. Note that each of these quantities may actually be considered over a general domain S in $\mathbb{R}^3$, not simply a box, B.

Key Idea 12.7.10 Interpretations of Triple Integrals. *For a solid region S in $\mathbb{R}^3$, the following are common and important interpretations of triple integrals over S:*

1. *The triple integral*

$$V(S) = \iiint_S 1\,dV$$

represents the ***volume*** *of the solid S.*

2. *If S has density given by $\delta = \delta(x,y,z)$, then $\iiint_S \delta(x,y,z)\,dV$ is the* ***mass*** *of S.*

3. *If S has density given by $\delta = \delta(x,y,z)$, then the* ***center of mass*** *$(\overline{x},\overline{y},\overline{z})$ can be computed as*

$$\overline{x} = \frac{\iiint_S x\ \delta(x,y,z)\,dV}{M},$$
$$\overline{y} = \frac{\iiint_S y\ \delta(x,y,z)\,dV}{M},$$
$$\overline{z} = \frac{\iiint_S z\ \delta(x,y,z)\,dV}{M},$$

where M is the mass of S.

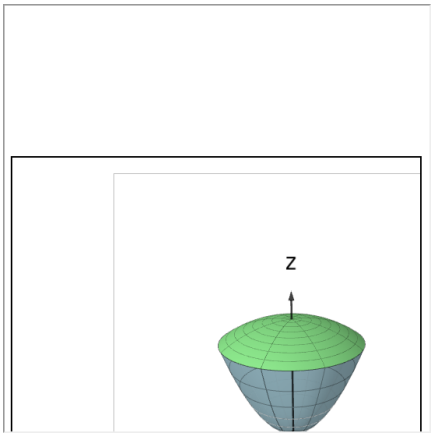
4. *The* ***average value*** *of a function f over S is given by*

$$f_{\text{AVG}(S)} = \left(\frac{1}{V(S)}\right)\iiint_S f(x,y,z)\,dV,$$

where $V(S)$ is the volume of the solid S.

The next activity asks you to practice setting up triple integrals as iterated integrals, including thinking about the interpretations of Key Idea 12.7.10.

Activity 12.7.5 A solid S is bounded below by the paraboloid $z = x^2+y^2$ and above by the sphere $x^2+y^2+z^2 = 6$. A picture of S is shown in Figure 12.7.11.



Standalone
In Context
Embed

Figure 12.7.11 The solid bounded below by the paraboloid $z = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 6$

(a) Set up an iterated triple integral to find the volume of S. You do not need to evaluate the integral at this time.

Hint. You need to find three compound inequalities that bound x, y, and z to get your limits of integration. You may find it helpful to recall the approach of Activity 12.7.3 and imagine a line segment through S (parallel to the axis of the inner most variable). When finding the bounds on x and y, you will also likely find it helpful to set up your inequalities in terms of $x^2 + y^2$.

(b) Suppose the density at point (x,y,z) is $\delta(x,y,z) = x^2+1$. Set up, but do not evaluate, the necessary iterated integrals to find the center of mass of S.

(c) Let $f(x, y, z) = xy + z^2$. Set up but do not evaluate an iterated triple integral to find the average value of f on S.

(d) Use technology to evaluate the iterated triple integrals you wrote in the other three parts of this activity. Write a couple of sentences to explain why the location of the center of mass makes sense.

Summary

- There is a natural generalization of Riemann sums and integration to functions of three variables that yields the definition for a triple integral as

$$\iiint_B f(x, y, z)\, dV = \lim_{\Delta V \to 0} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} f(P_{ijk})\, \Delta V,$$

  where the triple Riemann sum is defined in the usual way. The definition of the triple integral naturally extends to non-rectangular solid regions.

- Iterated integrals can be used to efficiently evaluate triple integrals by bounding surfaces in one coordinate, then projecting the solid region onto the remaining coordinate plane. There are six different orders of integration possible in terms of rectangular coordinates.

- For a solid region S in $\mathbb{R}^3$, the following are common and important interpretations of triple integrals over S:

  1. The triple integral
  $$V(S) = \iiint_S 1\, dV$$
  represents the **volume** of the solid S.

  2. If S has density given by $\delta = \delta(x, y, z)$, then $\iiint_S \delta(x, y, z)\, dV$ is the **mass** of S.

  3. If S has density given by $\delta = \delta(x, y, z)$, then the **center of mass** $(\overline{x}, \overline{y}, \overline{z})$ can be computed as
  $$\overline{x} = \frac{\iiint_S x\ \delta(x, y, z)\, dV}{M},$$
  $$\overline{y} = \frac{\iiint_S y\ \delta(x, y, z)\, dV}{M},$$
  $$\overline{z} = \frac{\iiint_S z\ \delta(x, y, z)\, dV}{M},$$
  where M is the mass of S.

  4. The **average value** of a function f over S is given by
  $$f_{\text{AVG}(S)} = \frac{1}{V(S)} \iiint_S f(x, y, z)\, dV,$$
  where $V(S)$ is the volume of the solid S.

12.7.3 Exercises

1. Let W_1 be the solid half-cone bounded by $z = \sqrt{x^2 + y^2}$, $z = 4$ and the yz-plane with $x \leq 0$, and let Let W_2 be the solid half-cone bounded by $z = \sqrt{x^2 + y^2}$, $z = 4$ and the xz-plane with $y \geq 0$.

For each of the following, decide (without calculating its value) whether the integral is positive, negative, or zero.

(a) $\int_{W_1} x\,dV$ is (☐ positive ☐ negative ☐ zero)

(b) $\int_{W_1} x^2 y\,dV$ is (☐ positive ☐ negative ☐ zero)

(c) $\int_{W_2} yz\,dV$ is (☐ positive ☐ negative ☐ zero)

2. Without calculation, decide if each of the integrals below are positive, negative, or zero. Let W be the solid bounded by $z = \sqrt{x^2 + y^2}$ and $z = 2$.

(a) $\iiint_W \left(z - \sqrt{x^2 + y^2}\right)\,dV$

(b) $\iiint_W (z - 2)\,dV$

(c) $\iiint_W e^{-xyz}\,dV$

3. Find the triple integral of the function $f(x, y, z) = x^3 \cos(y + z)$ over the cube $6 \leq x \leq 7$, $0 \leq y \leq \pi$, $0 \leq z \leq \pi$.

4. Find the average value of the function $f\,(x, y, z) = x^2 + y^2 + z^2$ over the rectangular prism $0 \leq x \leq 1$, $0 \leq y \leq 4$, $0 \leq z \leq 5$.

5. Evaluate the triple integral

$$\iiint_{\mathbf{E}} xyz\,dV$$

where $\boldsymbol{E}$ is the solid: $0 \leq z \leq 3$, $0 \leq y \leq z$, $0 \leq x \leq y$.

6. Express the integral $\iiint_E f(x, y, z)dV$ as an iterated integral in six different ways, where E is the solid bounded by $z = 0, x = 0, z = y - 3x$ and $y = 12$.

1. $\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)dzdydx$

$a =$ _____ $b =$ _____

$g_1(x) =$ _____ $g_2(x) =$ _____

$h_1(x, y) =$ _____ $h_2(x, y) =$ _____

2. $\int_a^b \int_{g_1(y)}^{g_2(y)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z)dzdxdy$

$a =$ _____ $b =$ _____

$g_1(y) =$ _____ $g_2(y) =$ _____

$h_1(x, y) =$ _____ $h_2(x, y) =$ _____

3. $\int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(y,z)}^{h_2(y,z)} f(x, y, z)dxdydz$

$a =$ _____ $b =$ _____

$g_1(z) =$ _____ $g_2(z) =$ _____

$h_1(y, z) =$ _____ $h_2(y, z) =$ _____

4. $\int_a^b \int_{g_1(y)}^{g_2(y)} \int_{h_1(y,z)}^{h_2(y,z)} f(x,y,z)dxdzdy$

$a =$ _____ $b =$ _____

$g_1(y) =$ _____ $g_2(y) =$ _____

$h_1(y,z) =$ _____ $h_2(y,z) =$ _____

5. $\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,z)}^{h_2(x,z)} f(x,y,z)dydzdx$

$a =$ _____ $b =$ _____

$g_1(x) =$ _____ $g_2(x) =$ _____

$h_1(x,z) =$ _____ $h_2(x,z) =$ _____

6. $\int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(x,z)}^{h_2(x,z)} f(x,y,z)dydxdz$

$a =$ _____ $b =$ _____

$g_1(z) =$ _____ $g_2(z) =$ _____

$h_1(x,z) =$ _____ $h_2(x,z) =$ _____

7. Find the volume of the solid enclosed by the paraboloids $z = 4\left(x^2 + y^2\right)$ and $z = 18 - 4\left(x^2 + y^2\right)$.

8. Evaluate the triple integral

$$\iiint_{\mathbf{E}} xyz\, dV$$

where $\boldsymbol{E}$ is the solid: $0 \le z \le 6, \quad 0 \le y \le z, \quad 0 \le x \le y$.

9. Rewrite the triple integral $\int_0^1 \int_0^x \int_0^y f(x,y,z)dzdydx$ as $\int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(y,z)}^{h_2(y,z)} f(x,y,z)dxdydz$

$a =$ _____ $b =$ _____

$g_1(z) =$ _____ $g_2(z) =$ _____

$h_1(y,z) =$ _____ $h_2(y,z) =$ _____

10. Evaluate $\iiint_{\mathcal{W}} f(x,y,z)\, dV$ for the function f and region $\mathcal{W}$ specified:

$$f(x,y,z) = 42(x+y) \quad \mathcal{W} : y \le z \le x, 0 \le y \le x, 0 \le x \le 1$$

$\iiint_{\mathcal{W}} (42(x+y))\, dV =$ _____

11. Integrate $f(x,y,z) = 18xz$ over the region in the first octant $(x,y,z \ge 0)$ above the parabolic cylinder $z = y^2$ and below the paraboloid $z = 8 - 2x^2 - y^2$.

Answer : _____

12. Find the volume of the solid in $\mathbf{R}^3$ bounded by $y = x^2$, $x = y^2$, $z = x + y + 36$, and $z = 0$.

V = _____

13. Consider the triple integral $\iiint_{\mathcal{W}} xyz^2\, dV$, where $\mathcal{W}$ is the region bounded by

$$z = 4 - y^2, \quad z = 0, \quad y = 6x, \quad x = 0, \quad y \ge 0.$$

Write the triple integral as an iterated integral in the order $dz\, dx\, dy$, and describe the region of integration:

_____ $\le x \le$ _____

_____ $\le y \le$ _____

_____ $\le z \le$ _____

14. Evaluate the iterated integral $\int_0^1 \int_x^{3x} \int_0^y 3xyz \, dz \, dy \, dx$.

Answer: _____

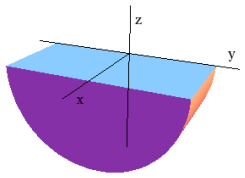
15. Evaluate the iterated integral $\int_0^{\pi/3} \int_0^y \int_0^x \cos(x+y+z) \, dz \, dx \, dy$.

Answer: _____

16. Evaluate the triple integral $\iiint_T x^2 \, dV$, where T is the solid tetrahedron with vertices $(0,0,0)$, $(3,0,0)$, $(0,3,0)$, and $(0,0,3)$.

Answer: _____

17. Write limits of integration for the integral $\int_W g(x,y,z) \, dV$, where W is the half cylinder shown, if the length of the cylinder is 1 and its radius is 1.



$\int_W g(x,y,z) \, dV = \int_a^b \int_c^d \int_e^f g(x,y,z) \, d___ \, d___ \, d___$
where $a =$ _____, $b =$ _____,
$c =$ _____, $d =$ _____,
$e =$ _____, and $f =$ _____.

(Note: values for all answer blanks must be supplied for this problem to be able to check the answers provided.)

18. Find the mass of the rectangular prism $0 \le x \le 3$, $0 \le y \le 4$, $0 \le z \le 3$, with density function $\rho(x,y,z) = x$.

19. Find the average value of the function $f(x,y,z) = ye^{-xy}$ over the rectangular prism $0 \le x \le 4$, $0 \le y \le 3$, $0 \le z \le 1$

20. Find the volume of the solid bounded by the planes x = 0, y = 0, z = 0, and x + y + z = 8.

21. Find the mass of the solid bounded by the xy-plane, yz-plane, xz-plane, and the plane $(x/4) + (y/4) + (z/16) = 1$, if the density of the solid is given by $\delta(x,y,z) = x + 4y$.

mass = _____

22. The *moment of inertia* of a solid body about an axis in 3-space relates the angular acceleration about this axis to torque (force twisting the body). The moments of inertia about the coordinate axes of a body of constant density and mass m occupying a region W of volume V are defined to be

$$I_x = \frac{m}{V}\int_W (y^2+z^2)\, dV \qquad I_y = \frac{m}{V}\int_W (x^2+z^2)\, dV \qquad I_z = \frac{m}{V}\int_W (x^2+y^2)\, dV$$

Use these definitions to find the moment of inertia about the z-axis of the rectangular solid of mass 9 given by $0 \le x \le 1$, $0 \le y \le 3$, $0 \le z \le 3$.

$I_x =$ _____
$I_y =$ _____
$I_z =$ _____

23. Calculate the volume under the elliptic paraboloid $z = x^2 + 2y^2$ and over the rectangle $R = [-4, 4] \times [-4, 4]$.

24. The motion of a solid object can be analyzed by thinking of the mass as concentrated at a single point, the *center of mass*. If the object has density $\rho(x, y, z)$ at the point (x, y, z) and occupies a region W, then the coordinates $(\overline{x}, \overline{y}, \overline{z})$ of the center of mass are given by

$$\overline{x} = \frac{1}{m}\int_W x\rho\, dV \quad \overline{y} = \frac{1}{m}\int_W y\rho\, dV \quad \overline{z} = \frac{1}{m}\int_W z\rho\, dV,$$

Assume x, y, z are in cm. Let C be a solid cone with both height and radius 2 and contained between the surfaces $z = \sqrt{x^2 + y^2}$ and $z = 2$. If C has constant mass density of 4 g/cm^3, find the z-coordinate of C's center of mass.

$\overline{z} =$ _____

(Include .)

25. Set up a triple integral to find the mass of the solid tetrahedron bounded by the xy-plane, the yz-plane, the xz-plane, and the plane $x/4 + y/3 + z/12 = 1$, if the density function is given by $\delta(x, y, z) = x + y$. Write an iterated integral in the form below to find the mass of the solid.

$$\iiint_R f(x, y, z)\, dV = \int_A^B \int_C^D \int_E^F _____\ dz\, dy\, dx$$

with limits of integration

A = _____

B = _____

C = _____

D = _____

E = _____

F = _____

26. Consider the solid S that is bounded by the parabolic cylinder $y = x^2$ and the planes $z = 0$ and $z = 1 - y$ as shown in Figure 12.7.12.

Assume the density of S is given by $\delta(x, y, z) = z$

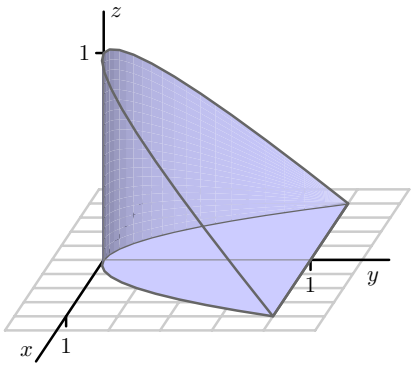


Figure 12.7.12 The solid bounded by $y = x^2$ and the planes $z = 0$ and $z = 1 - y$.

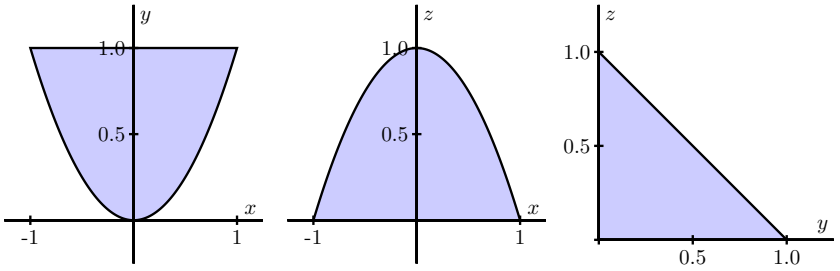


Figure 12.7.13 Projections of S onto the xy, xz, and yz-planes.

(a) Set up (but do not evaluate) an iterated integral that represents the mass of S. Integrate first with respect to z, then y, then x. A picture of the projection of S onto the xy-plane is shown at left in Figure 12.7.13.

(b) Set up (but do not evaluate) an iterated integral that represents the mass of S. In this case, integrate first with respect to y, then z, then x. A picture of the projection of S onto the xz-plane is shown at center in Figure 12.7.13.

(c) Set up (but do not evaluate) an iterated integral that represents the mass of S. For this integral, integrate first with respect to x, then y, then z. A picture of the projection of S onto the yz-plane is shown at right in Figure 12.7.13.

(d) Which of these three orders of integration is the most natural to you? Why?

27. This problem asks you to investigate the average value of some different quantities.

(a) Set up, but do not evaluate, an iterated integral expression whose value is the average sum of all real numbers x, y, and z that have the following property: y is between 0 and 2, x is greater than or equal to 0 but cannot exceed $2y$, and z is greater than or equal to 0 but cannot exceed $x + y$.

(b) Set up, but do not evaluate, an integral expression whose value represents the average value of $f(x, y, z) = x + y + z$ over the solid region in the first octant bounded by the surface $z = 4 - x - y^2$ and the coordinate planes $x = 0$, $y = 0$, $z = 0$.

(c) How are the quantities in (a) and (b) similar? How are they different?

28. Consider the solid that lies between the paraboloids $z = g(x, y) = x^2 + y^2$ and $z = f(x, y) = 8 - 3x^2 - 3y^2$.

(a) By eliminating the variable z, determine the curve of intersection between the two paraboloids, and sketch this curve in the xy-plane.

(b) Set up, but do not evaluate, an iterated integral expression whose value determines the mass of the solid, integrating first with respect to z, then y, then x. Assume the the solid's density is given by $\delta(x, y, z) = \frac{1}{x^2+y^2+z^2+1}$.

(c) Set up, but do not evaluate, iterated integral expressions whose values determine the mass of the solid using all possible remaining

orders of integration. Use $\delta(x, y, z) = \frac{1}{x^2+y^2+z^2+1}$ as the density of the solid.

(d) Set up, but do not evaluate, iterated integral expressions whose values determine the center of mass of the solid. Again, assume the the solid's density is given by $\delta(x, y, z) = \frac{1}{x^2+y^2+z^2+1}$.

(e) Which coordinates of the center of mass can you determine *without* evaluating any integral expression? Why?

29. In each of the following problems, your task is to

i. Sketch, by hand, the region over which you integrate

ii. set up iterated integral expressions which, when evaluated, will determine the value sought

iii. use appropriate technology to evaluate each iterated integral expression you develop

In some problems you may be able to use a double rather than a triple integral, and polar coordinates may be helpful in some cases.

(a) Consider the solid created by the region enclosed by the circular paraboloid $z = 4-x^2-y^2$ over the region R in the xy-plane enclosed by $y = -x$ and the circle $x^2 + y^2 = 4$ in the first, second, and fourth quadrants. Determine the solid's volume.

(b) Consider the solid region that lies beneath the circular paraboloid $z = 9-x^2-y^2$ over the triangular region between $y = x$, $y = 2x$, and $y = 1$. Assuming that the solid has its density at point (x, y, z) given by $\delta(x, y, z) = xyz + 1$, measured in grams per cubic cm, determine the center of mass of the solid.

(c) In a certain room in a house, the walls can be thought of as being formed by the lines $y = 0$, $y = 12 + x/4$, $x = 0$, and $x = 12$, where length is measured in feet. In addition, the ceiling of the room is vaulted and is determined by the plane $z = 16 - x/6 - y/3$. A heater is stationed in the corner of the room at $(0, 0, 0)$ and causes the temperature in the room at a particular time to be given by

$$T(x, y, z) = \frac{80}{1 + \frac{x^2}{1000} + \frac{y^2}{1000} + \frac{z^2}{1000}}$$

What is the average temperature in the room?

(d) Consider the solid enclosed by the cylinder $x^2 + y^2 = 9$ and the planes $y + z = 5$ and $z = 1$. Assuming that the solid's density is given by $\delta(x, y, z) = \sqrt{x^2 + y^2}$, find the mass and center of mass of the solid.

12.8 Triple Integrals in Cylindrical and Spherical Coordinates

Motivating Questions

- What is the volume element in cylindrical coordinates? How does this inform us about evaluating a triple integral as an iterated integral in cylindrical coordinates?
- What is the volume element in spherical coordinates? How does this inform us about evaluating a triple integral as an iterated integral in spherical coordinates?

12.8.1 Introduction

In Section 12.6, we saw how advantageous it could be to use polar coordinates to write iterated integrals when trying to evaluate a double integral. In particular, we saw that double integrals with integrands and regions of integration that are simple to write in polar coordinates are the most convenient to convert, as well as how the area element would be equal to $dA = rdr\,d\theta$ when using polar coordinates.

In Section 9.9, we introduced both cylindrical and spherical coordinate systems and explored converting points, equations, and regions between these and rectangular coordinate systems in three dimensions. Cylindrical coordinates are most commonly used in situations with rotational symmetry around the z-axis, whereas spherical coordinates are primarily used when there is rotational symmetry around the origin. We encourage you to review these elements, including the conversion equations between cylindrical and rectangular coordinates and conversion equations between spherical and rectangular coordinates. The following preview activity will include some descriptions of common surfaces given in either cylindrical or spherical coordinates.

Preview Activity 12.8.1 In this Preview Activity, we will be reviewing the meaning of each of the cylindrical and spherical coordinates by looking at a description of a common surface in either cylindrical or spherical coordinates. For each task, you should draw a plot of the surface described by hand and write a few sentences describing how your plot relates to the cylindrical/spherical coordinates.

(a) What familiar surface is described by the points in cylindrical coordinates with $r = 2$, $0 \leq \theta \leq 2\pi$, and $0 \leq z \leq 2$? How does this example suggest that we call these coordinates *cylindrical coordinates*? How does the surface change if we restrict θ to $0 \leq \theta \leq \pi$?

(b) What familiar surface is described by the points in cylindrical coordinates with $\theta = 2$, $0 \leq r \leq 2$, and $0 \leq z \leq 2$?

(c) Plot the graph of the cylindrical equation $z = r$, where $0 \leq \theta \leq 2\pi$ and $0 \leq r \leq 2$. What familiar surface is this a plot of?

(d) What familiar surface is described by the points in spherical coordinates with $\rho = 1$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$? How does this particular example demonstrate the reason for the name of this coordinate system? What if we restrict ϕ to $0 \leq \phi \leq \frac{\pi}{2}$?

(e) What familiar surface is described by the points in spherical coordinates with $\phi = \frac{\pi}{3}$, $0 \leq \rho \leq 1$, and $0 \leq \theta \leq 2\pi$?

In this section, we will look at converting triple integrals to use cylindrical and spherical coordinates. Similar to our approach in Section 12.6, we will focus conceptually on converting

1. the volume element dV
2. the integrand (the function being integrated)
3. the region of integration

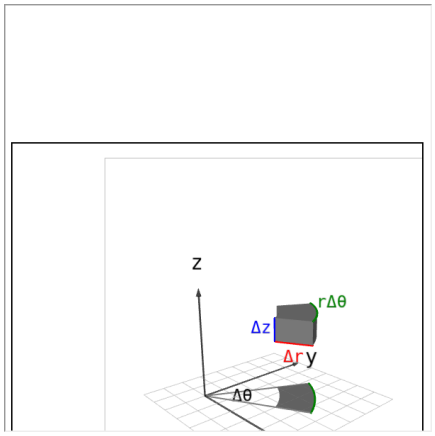
to cylindrical and spherical coordinates.

12.8.2 Triple Integrals in Cylindrical Coordinates

To evaluate a triple integral $\iiint_S f(x, y, z)\, dV$ as an iterated integral in Cartesian coordinates, we use the fact that the volume element dV is equal to $dz\, dy\, dx$ (which corresponds to the volume of a small box). To evaluate a triple integral in cylindrical coordinates, we similarly must understand the volume element dV in cylindrical coordinates.

In particular, we need to look at the volume of a region created by a small step in each of the coordinates $(\Delta r, \Delta\theta, \Delta z)$. You can see a plot of an example region in Figure 12.8.1 and you should recognize the shaded area in the xy-plane as the area element from our work on double integrals in polar coordinates. Recall that ΔA was $\frac{1}{2}(r_{i+1} + r_i)\Delta r\ \Delta\theta$ and since z is the perpendicular height of our region, we have

$$\Delta V = \Delta A\ \Delta z = \frac{(r_{i+1} + r_i)}{2}\Delta r\, \Delta\theta\, \Delta z.$$



Standalone
In Context
Embed

Figure 12.8.1 A cylindrical box

Remember that as Δr shrinks, $\frac{(r_{i+1}+r_i)}{2}$ approaches the r-coordinate of this block. Therefore the volume element dV in cylindrical coordinates is given by $dV = r\, dz\, dr\, d\theta$.

We can also convert our integrand to polar coordinates by using $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Slicing our region into a radially simple or angularly simple description will provide the bounds of our iterated integrals used to evaluate the triple integral. Remember that cylindrical coordinates are really polar coordinates plus z.

Triple integrals in cylindrical coordinates.

Given a continuous function $f = f(x, y, z)$ over a region S in $\mathbb{R}^3$,

$$\iiint_S f(x, y, z)\, dV = \iiint_S f(r\cos(\theta), r\sin(\theta), z)\, r\, dz\, dr\, d\theta.$$

The latter expression is an *iterated integral in cylindrical coordinates.*

While the volume element and the conversion of the integrand will work the same for every case, the task of writing an iterated integral in cylindrical coordinates will require us to slice the particular region in terms of simple bounds on θ, r, and z. In the following activity, we explore how to do this in several situations where cylindrical coordinates are natural and advantageous.

Activity 12.8.2 In this activity, we will work with a few triple integrals and look at the advantages of utilizing cylindrical coordinates for the iterated integral calculation.

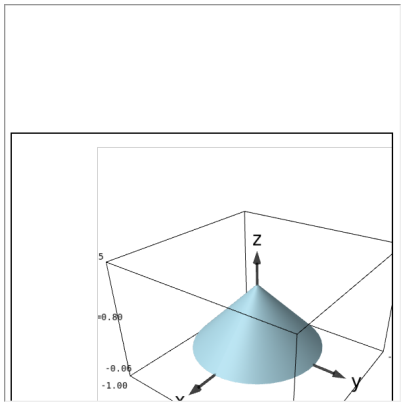
(a) Let S be the solid bounded above by the graph of $z = x^2 + y^2$ and below by $z = 0$ on the unit disk in the xy-plane.

i. The projection of the solid S onto the xy-plane is a disk. Describe this disk using inequalities in polar coordinates.

ii. Use your inequalities from above to write inequalities in cylindrical coordinates that describe the solid S.

iii. Write out a set of iterated integrals that will evaluate to the volume of S. You do not need to evaluate this integral.

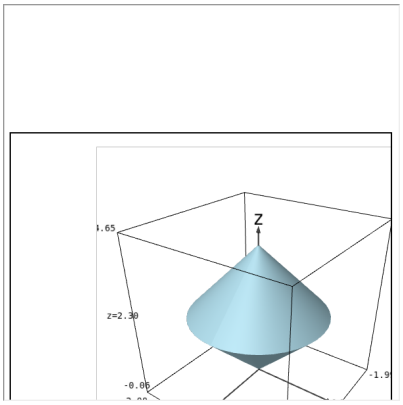
(b) Suppose the density of the cone defined by $r = 1 - z$, with $z \geq 0$, is given by $\delta(r, \theta, z) = z$. A picture of the cone is shown in Figure 12.8.2. Set up an iterated integral in cylindrical coordinates that gives the mass of the cone. You do not need to evaluate this integral.



Standalone
In Context
Embed

Figure 12.8.2 The cylindrical cone $r = 1 - z$

(c) Write an iterated integral expression in cylindrical coordinates whose value is the volume of the solid bounded below by the cone $z = \sqrt{x^2 + y^2}$ and above by the cone $z = 4 - \sqrt{x^2 + y^2}$. A picture is shown in Figure 12.8.3. You do not need to evaluate this integral.



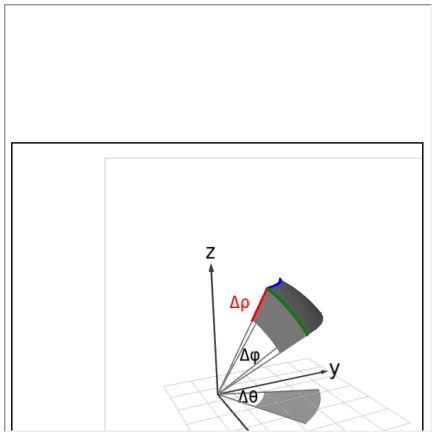
Standalone
In Context
Embed

Figure 12.8.3

12.8.3 Triple Integrals in Spherical Coordinates

Similar to our work with cylindrical and Cartesian coordinates, when it comes to thinking about particular surfaces in spherical coordinates, we usually write ρ as a function of θ and ϕ; this is a natural analog to polar coordinates, where we often think of our distance from the origin in the plane as being a function of θ. In spherical coordinates, we likewise often view ρ as a function of θ and ϕ, thus viewing distance from the origin as a function of two key angles.

As with rectangular and cylindrical coordinates, a triple integral $\iiint_S f(x,y,z)\,dV$ in spherical coordinates can be evaluated as an iterated integral once we understand the volume element dV. In Figure 12.8.4, we see a volume given by a change in each of the ρ, θ, ϕ coordinates. A conceptual description for where the volume element comes from is given in the proof to Key Idea 12.8.5.



Standalone
In Context
Embed

Figure 12.8.4

Key Idea 12.8.5 Triple integrals in spherical coordinates. *Given a continuous function $f = f(x,y,z)$ over a region S in $\mathbb{R}^3$, the triple integral $\iiint_S f(x,y,z)\,dV$ is converted to the integral*

$$\iiint_S f(\rho\sin(\phi)\cos(\theta), \rho\sin(\phi)\sin(\theta), \rho\cos(\phi))\,\rho^2\sin(\phi)\,d\rho\,d\theta\,d\phi$$

in spherical coordinates.

The latter expression is an iterated integral in spherical coordinates. *Note that the volume element for spherical coordinates is $dV = \rho^2\sin(\phi)\,d\rho\,d\theta\,d\phi$*

Proof. The conversion of the function being integrated is straightforward because $x = \rho\sin(\phi)\cos(\theta), y = \rho\sin(\phi)\sin(\theta), z = \rho\cos(\phi)$ does not require interpretation based on the coordinates used.

To find the volume element dV in spherical coordinates, we need to understand how to determine the volume of a spherical box of the form $\rho_1 \leq \rho \leq \rho_2$ (with $\Delta\rho = \rho_2 - \rho_1$), $\phi_1 \leq \phi \leq \phi_2$ (with $\Delta\phi = \phi_2 - \phi_1$), and $\theta_1 \leq \theta \leq \theta_2$ (with $\Delta\theta = \theta_2 - \theta_1$). An illustration of such a box is given in Figure 12.8.4. This spherical box is a bit more complicated than the cylindrical box we encountered earlier. In this situation, it is easier to approximate the volume ΔV than to compute it directly. Here we can approximate the volume ΔV of this spherical box with the volume of a Cartesian box whose sides have the lengths of the sides of this spherical box. We have highlighted the three sides we will use in red, blue, and green.

Along the red side, both θ and ϕ are constant and ρ is changing. In fact, the length of the red side is simply the change in ρ.

Along the green side, both θ and ρ are constant and ϕ is changing. We can simplify our plot by looking along the plane $\theta = \theta_0$ which will give us the following plot

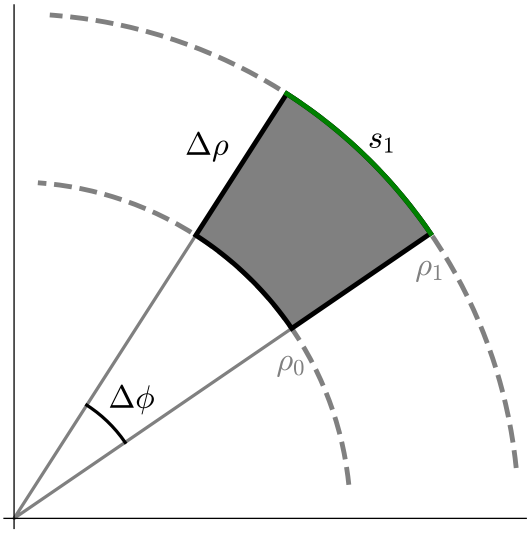


Figure 12.8.6 A plot of the face of our volume element with $\theta = \theta_0$

Hopefully you recognize this as the same kind of picture as we used when looking at the area element in polar coordinates. The length of the green curve will be $\rho_1\Delta\phi$ by the same kind of argument.

$$\text{(fraction of circle)(circumference of circle)} = \left(\frac{\Delta\phi}{2\pi}\right)(2\pi\rho_1) = \rho_1\Delta\phi$$

The blue curve occurs at a constant value of ρ and ϕ, with θ changing. In other words, the blue curve consists of $x = \rho_1\sin(\phi_0)\cos(\theta), y = \rho_1\sin(\phi_0), z = \rho_1\cos(\phi_0)$ which corresponds to a circle of radius $\rho_1\sin(\phi_0)$ at a height $z = \rho_1\cos(\phi_0)$. We can use the same type of argument as above that the arc length of a curve corresponding to an angular change of $\Delta\theta$ along a circle of radius $\rho_1\sin(\phi_0)$ will be $\rho_1\sin(\phi_0)\Delta\theta$.

Remember that as we look at smaller and smaller changes in the coordinates, the difference between the values of the coordinates used will also shrink and $\rho_0 \approx \rho_1 \approx \rho$, for example. So we can get an idea of how large this volume element (dV) will be by multiplying the lengths of the red, blue, and green

segments to get

$$\Delta V \approx (\Delta\rho)\,(\rho\,\Delta\phi)\,(\rho\sin(\phi)\,\Delta\theta) = \rho^2\sin(\phi)\Delta\phi\,\Delta\theta\,\Delta\rho$$

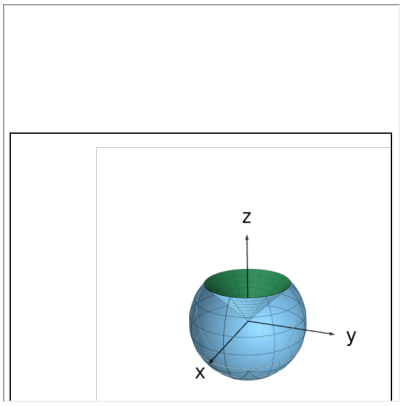
■

Finally, in order to actually evaluate an iterated integral in spherical coordinates, we must of course determine the limits of integration in terms of ϕ, θ, and ρ. The process is similar to our earlier work in other coordinate systems and we will look a few simple cases in the next activities; The central idea in our process is to fix all but one coordinate and give the same bounding functions for the remaining coordinate, then project onto the remaining coordinates and repeat this process.

Activity 12.8.3 In this activity, we will use spherical coordinates to help us more easily understand some natural geometric objects.

(a) Recall that the sphere of radius a has spherical equation $\rho = a$. Set up and evaluate an iterated integral in spherical coordinates to determine the volume inside a sphere of radius a.

(b) Set up, but do not evaluate, an iterated integral expression in spherical coordinates whose value is the mass of the solid obtained by removing the region inside the cone $\phi = \frac{\pi}{4}$ from the sphere $\rho = 2$. Let δ at the point (x, y, z) be the density given by $\delta(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. An illustration of this solid is shown in Figure 12.8.7.



Standalone
In Context
Embed

Figure 12.8.7 The solid cut from the sphere $\rho = 2$ by the cone $\phi = \frac{\pi}{4}$

Summary

- The volume element dV in cylindrical coordinates is $dV = r\,dz\,dr\,d\theta$. Hence, a triple integral $\iiint_S f(x, y, z)\,dA$ can be evaluated as the iterated integral of the form

$$\iiint_S f(r\cos(\theta), r\sin(\theta), z)\,r\,dz\,dr\,d\theta.$$

- The volume element dV in spherical coordinates is $dV = \rho^2\sin(\phi)\,d\rho\,d\theta\,d\phi$. Thus, a triple integral $\iiint_S f(x, y, z)\,dA$ can be evaluated as the iterated integral of the form

$$\iiint_S f(\rho\sin(\phi)\cos(\theta), \rho\sin(\phi)\sin(\theta), \rho\cos(\phi))\,\rho^2\sin(\phi)\,d\rho\,d\theta\,d\phi.$$

12.8.4 Exercises

You can find many more problems involving the conversion of points, graphs, and region between rectangular, cylindrical, and spherical coordinates in Section 9.9. The exercises of this section will be focused on applications of these conversion to triple integrals.

1. Match the given equation with the verbal description of the surface:

A. Cone

B. Half plane

C. Circular Cylinder

D. Plane

E. Elliptic or Circular Paraboloid

F. Sphere

(a) $r^2 + z^2 = 16$

(b) $r = 4$

(c) $\rho = 4$

(d) $\phi = \dfrac{\pi}{3}$

(e) $z = r^2$

(f) $\rho\cos(\phi) = 4$

(g) $\theta = \dfrac{\pi}{3}$

(h) $\rho = 2\cos(\phi)$

(i) $r = 2\cos(\theta)$

2. Match the integrals with the type of coordinates which make them the easiest to do. Put the letter of the coordinate system to the left of the number of the integral.

(a) $\iiint_E z^2\, dV$ where E is: $-2 \le z \le 2, \quad 1 \le x^2 + y^2 \le 2$

(b) $\iiint_E z\, dV$ where E is: $1 \le x \le 2, \quad 3 \le y \le 4, \quad 5 \le z \le 6$

(c) $\displaystyle\int_0^1 \int_0^{y^2} \frac{1}{x}\, dx\, dy$

(d) $\iint_D \dfrac{1}{x^2 + y^2}\, dA$ where D is: $x^2 + y^2 \le 4$

(e) $\iiint_E dV$ where E is: $x^2 + y^2 + z^2 \le 4, \quad x \ge 0, \quad y \ge 0, \quad z \ge 0$

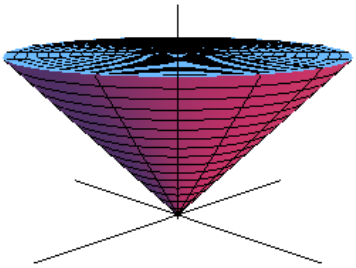
A. spherical coordinates

B. polar coordinates

C. cylindrical coordinates

D. cartesian coordinates

3. The region W is the cone shown below.



The angle at the vertex is $2\pi/3$, and the top is flat and at a height of $3/\sqrt{3}$.

Write the limits of integration for $\int_W dV$ in the following coordinates (do ***not*** reduce the domain of integration by taking advantage of symmetry):

(a) Cartesian:
With $a =$_____, $b =$_____,
$c =$_____, $d =$_____,
$e =$_____, and $f =$_____,
Volume $= \int_a^b \int_c^d \int_e^f$ _____ d_____ d_____ d_____

(b) Cylindrical:
With $a =$_____, $b =$_____,
$c =$_____, $d =$_____,
$e =$_____, and $f =$_____,
Volume $= \int_a^b \int_c^d \int_e^f$ _____ d_____ d_____ d_____

(c) Spherical:
With $a =$_____, $b =$_____,
$c =$_____, $d =$_____,
$e =$_____, and $f =$_____,
Volume $= \int_a^b \int_c^d \int_e^f$ _____ d_____ d_____ d_____

4. Evaluate the integral.

$$\int_0^7 \int_{-\sqrt{49-x^2}}^{\sqrt{49-x^2}} \int_{-\sqrt{49-x^2-z^2}}^{\sqrt{49-x^2-z^2}} \frac{1}{(x^2+y^2+z^2)^{1/2}}\, dy\, dz\, dx \quad =$$

_____

5. Use cylindrical coordinates to evaluate the triple integral $\iiint_{\mathbf{E}} \sqrt{x^2+y^2}\, dV$, where $\boldsymbol{E}$ is the solid bounded by the circular paraboloid $z = 9 - 16\left(x^2+y^2\right)$ and the xy -plane.

6. Use spherical coordinates to evaluate the triple integral $\iiint_{\mathbf{E}} x^2+y^2+z^2\, dV$, where $\boldsymbol{E}$ is the ball: $x^2+y^2+z^2 \leq 1$.

7. Find the volume of the solid enclosed by the paraboloids $z = 1\left(x^2+y^2\right)$ and $z = 32 - 1\left(x^2+y^2\right)$.

8. FInd the volume of the ellipsoid $x^2+y^2+6z^2 = 64$.

9. The density, δ, of the cylinder $x^2 + y^2 \leq 9$, $0 \leq z \leq 3$ varies with the distance, r, from the z-axis:

$$\delta = 3 + r \text{ g/cm}^3.$$

Find the mass of the cylinder, assuming x, y, z are in cm.

mass = _____

(Include .)

10. Suppose $f(x, y, z) = \dfrac{1}{\sqrt{x^2 + y^2 + z^2}}$ and W is the bottom half of a sphere of radius 4. Enter ρ as *rho,* ϕ as *phi,* and θ as *theta.*

(a) As an iterated integral,

$$\iiint_W f\, dV = \int_A^B \int_C^D \int_E^F _____\; d\rho\, d\phi\, d\theta$$

with limits of integration

A = _____

B = _____

C = _____

D = _____

E = _____

F = _____

(b) Evaluate the integral. _____

11. Use cylindrical coordinates to evaluate the triple integral $\iiint_{\mathbf{E}} \sqrt{x^2 + y^2}\, dV$, where $\boldsymbol{E}$ is the solid bounded by the circular paraboloid $z = 9 - 9\left(x^2 + y^2\right)$ and the xy -plane.

12. Evaluate the triple integral of $f(x, y, z) = \cos(x^2 + y^2)$ over the solid cylinder with height 7 and with base of radius 1 centered on the z axis at $z = -2$.

Integral = _____

13. Evaluate, in spherical coordinates, the triple integral of $f(\rho, \theta, \phi) = \sin\phi$, over the region $0 \leq \theta \leq 2\pi$, $0 \leq \phi \leq \pi/3$, $3 \leq \rho \leq 4$.

integral = _____

14. Write a triple integral including limits of integration that gives the volume of the cap of the solid sphere $x^2 + y^2 + z^2 \leq 25$ cut off by the plane $z = 3$ and restricted to the first octant. (In your integral, use theta, rho, and phi for θ, ρ and ϕ, as needed.)

What coordinates are you using? _____

*(Enter **cartesian**, **cylindrical**, or **spherical**.)*

With a =_____, b =_____,

c =_____, d =_____,

e =_____, and f =_____,

Volume = $\int_a^b \int_c^d \int_e^f$ _____ d_____ d_____ d_____

15. Find the volume between the cone $y = \sqrt{x^2 + z^2}$ and the sphere $x^2 + y^2 + z^2 = 9$.

volume = _____

16. A spherical shell centered at the origin has an inner radius of 3 cm and an outer radius of 6 cm. The density, δ, of the material increases linearly with the distance from the center. At the inner surface, $\delta = 9$ g/cm^3; at the outer surface, $\delta = 18$ g/cm^3.

(a) Using spherical coordinates, write the density, δ, as a function of radius, ρ. (Type rho for ρ.)

$\delta =$ _____

(b) Write an integral in spherical coordinates giving the mass of the shell (for this representation, do ***not*** reduce the domain of the integral by using symmetry; type phi and theta for ϕ and θ).

With $a =$_____, $b =$_____,

$c =$_____, $d =$_____,

$e =$_____, and $f =$_____,

Mass $= \int_a^b \int_c^d \int_e^f$ _____ d_____ d_____ d_____

(c) Find the mass of the shell.

Mass = _____

(Include .)

17. Use a triple integral to find the volume of the solid enclosed by the paraboloids $y = x^2 + z^2$ and $y = 18 - x^2 - z^2$.

Answer: _____

18. Use cylindrical coordinates to calculate :

$$\iiint_{\mathcal{W}} x^2 + y^2 \, dV \qquad \mathcal{W} : x^2 + y^2 \leq 25, \quad 0 \leq z \leq 4$$

$\iiint_{\mathcal{W}} (x^2 + y^2) \, dV =$ _____

19. Use cylindrical coordinates to calculate $\iiint_{\mathcal{W}} f(x, y, z) \, dV$ for the given function and region:

$$f(x, y, z) = z, \quad x^2 + y^2 \leq z \leq 36$$

$\iiint_{\mathcal{W}} f(x, y, z) \, dV =$ _____

20. Evaluate the triple integral of $f(x, y, z) = z(x^2 + y^2 + z^2)^{-3/2}$ over the part of the ball $x^2 + y^2 + z^2 \leq 36$ defined by $z \geq 3$.

$\iiint_{\mathcal{W}} f(x, y, z) \, dV =$ _____

21. In each of the following questions, set up an iterated integral expression whose value determines the desired result. Then, evaluate the integral first by hand, and then using appropriate technology.

(a) Find the volume of the "cap" cut from the solid sphere $x^2+y^2+z^2 = 4$ by the plane $z = 1$, as well as the z-coordinate of its centroid.

(b) Find the x-coordinate of the center of mass of the portion of the unit sphere that lies in the first octant (i.e., where x, y, and z are all nonnegative). Assume that the density of the solid given by $\delta(x, y, z) = \frac{1}{1+x^2+y^2+z^2}$.

(c) Find the volume of the solid bounded below by the xy-plane, on the sides by the sphere $\rho = 2$, and above by the cone $\phi = \pi/3$.

(d) Find the z coordinate of the center of mass of the region that is bounded above by the surface $z = \sqrt{\sqrt{x^2 + y^2}}$, on the sides by the cylinder $x^2 + y^2 = 4$, and below by the xy-plane. Assume that the density of the solid is uniform and constant.

(e) Find the volume of the solid that lies outside the sphere $x^2+y^2+z^2 = 1$ and inside the sphere $x^2 + y^2 + z^2 = 2z$.

22. For each of the following questions,

- sketch the region of integration,
- change the coordinate system in which the iterated integral is written to one of the remaining two,
- evaluate the iterated integral you deem easiest to evaluate by hand.

(a) $\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{2-x^2-y^2}} xy \, dz \, dy \, dx$

(b) $\int_0^{\pi/2} \int_0^{\pi} \int_0^1 \rho^2 \sin(\phi) \, d\rho \, d\phi \, d\theta$

(c) $\int_0^{2\pi} \int_0^1 \int_r^1 r^2 \cos(\theta) \, dz \, dr \, d\theta$

23. Consider the solid region S bounded above by the paraboloid $z = 16 - x^2 - y^2$ and below by the paraboloid $z = 3x^2 + 3y^2$.

(a) Describe parametrically the curve in $\mathbb{R}^3$ in which these two surfaces intersect.

(b) In terms of x and y, write an equation to describe the projection of the curve onto the xy-plane.

(c) What coordinate system do you think is most natural for an iterated integral that gives the volume of the solid?

(d) Set up, but do not evaluate, an iterated integral expression whose value is average z-value of points in the solid region S.

(e) Use technology to plot the two surfaces and evaluate the integral in part d. Write at least one sentence to discuss how your computations align with your intuition about where the average z-value of the solid should fall.

12.9 Change of Variables

Motivating Questions

- What is a change of variables in a multivariable coordinate system?
- How can we measure how area and volume are distorted for a change of variables in two or three dimensions?

12.9.1 Introduction

In single variable calculus, we encountered the idea of a change of variable in a definite integral through the method of substitution. For example, given the definite integral

$$\int_0^2 2x(x^2+1)^3 \, dx,$$

we naturally consider the change of variable $u = x^2+1$. From this substitution, it follows that $du = 2x \, dx$, and since $x = 0$ implies $u = 1$ and $x = 2$ implies $u = 5$, we have transformed the original integral in x into a new integral in u. In particular,

$$\int_0^2 2x(x^2+1)^3 \, dx = \int_1^5 u^3 \, du.$$

The latter integral, of course, is far easier to evaluate.

Through our work with polar, cylindrical, and spherical coordinates, we have already seen some of the issues and advantages that arise from using a change of variables with two or three variables. In the following Preview Activity, we will look geometrically at a couple of examples of changes of coordinates involving polar coordinates. In particular, we will explore exactly what properties an advantageous change of coordinates will have. In what follows, we seek to understand the general ideas behind any change of variables in a multiple integral.

Preview Activity 12.9.1

(a) Consider the double integral

$$I = \iint_D x^2 + y^2 \, dA \tag{12.9.1}$$

where D is the upper half of the unit disk ($y \geq 0$). Remember the polar conversion equations:

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta) \tag{12.9.2}$$

(i) Write the double integral I given in Equation (12.9.1) as an iterated integral in rectangular coordinates.

(ii) Write the double integral I given in Equation (12.9.1) as an iterated integral in polar coordinates.

(iii) Draw a plot of the region of integration D in the xy-plane and label the four "corners" of the polar coordinates (if D is described by $a \leq r \leq b, c \leq \theta \leq d$, then the corners will be $(r, \theta) = (a, c), (a, d), (b, c), (b, d)$).

Hint. Two of your "corners" will be at the same point

(b) Suppose you are calibrating a new radar system for your self driving car. This radar system is mounted at the center of your front bumper and can detect objects up to 45 degrees to either side of the forward direction at distances from 1 meter to 3 meters.

(i) On Figure 12.9.1, draw a plot of S, the region that your radar system can detect objects. Assume your car's front bumper is centered at the origin of your coordinate system and the car is facing the positive y-direction. You should label the four corners of S in terms of *both* rectangular and polar coordinates.

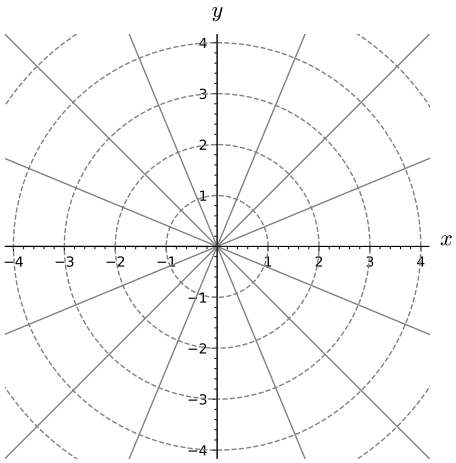


Figure 12.9.1 A polar grid

(ii) On Figure 12.9.2, plot the rectangle on the $r\theta$-plane that corresponds to the bounds $a \leq r \leq b \quad c \leq \theta \leq d$ for your radar system. Label the four corners of your figure with their r and θ coordinates.

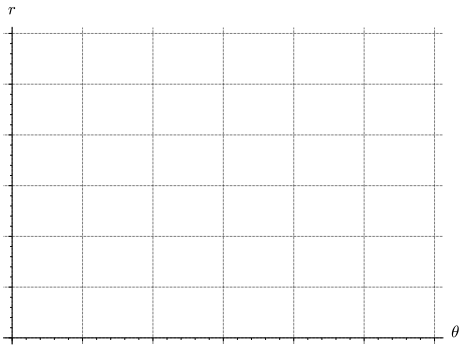


Figure 12.9.2 A grid on the θr-plane

12.9.2 Change of Variables in Polar Coordinates

The general idea behind a change of variables is suggested by Preview Activity 12.9.1. There, we saw that in a change of variables from rectangular coordinates to polar coordinates, a polar rectangle $[r_1, r_2] \times [\theta_1, \theta_2]$ gets mapped to a Cartesian rectangle under the transformation

$$x = r\cos(\theta) \quad \text{and} \quad y = r\sin(\theta).$$

Example 12.9.3 The vertices of the polar rectangle P given by $[r_1, r_2] \times [\theta_1, \theta_2]$ are transformed into the vertices of a closed and bounded region P' in rectangular coordinates. If we view the standard coordinate system as having the horizontal axis represent r and the vertical axis represent θ, then the polar rectangle P appears to us at left in Figure 12.9.4. The image P' of the polar rectangle P under the transformation given by (12.9.2) is shown at right in Figure 12.9.4. We thus see that there is a correspondence between a simple region (a traditional, right-angled rectangle) and a more complicated region (a fraction of an annulus) under the function T given by $T(r, \theta) = (r\cos(\theta), r\sin(\theta))$.

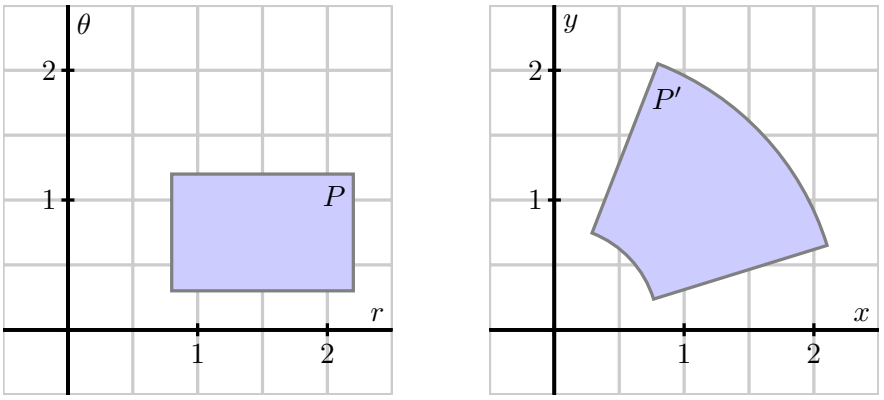


Figure 12.9.4 A rectangle P and its image P'.

▲

Furthermore, it follows generally that for an original polar rectangle $P = [r_1, r_2] \times [\theta_1, \theta_2]$, the area of the transformed rectangle P' is given by $\frac{r_2+r_1}{2}\Delta r \Delta\theta$. Therefore, as Δr and $\Delta\theta$ go to 0, this area becomes the familiar area element $dA = r\,dr\,d\theta$ in polar coordinates. When we work with other transformations for different changes in coordinates, we will have to understand how the transformation affects area so that we may use the correct area element in the new system of variables.

12.9.3 General Change of Coordinates

We first focus on double integrals. As with single integrals, we may be able to simplify a double integral of the form

$$\iint_D f(x, y)\, dA$$

by making a change of variables (that is, a substitution) of the form

$$x = x(s, t) \quad \text{and} \quad y = y(s, t)$$

where x and y are functions of new variables s and t. This transformation introduces a correspondence between a problem in the xy-plane and one in the the st-plane. The equations $x = x(s, t)$ and $y = y(s, t)$ convert s and t to x and y; we call these formulas the *change of variable* formulas.

To complete the change of our integral to the new s, t variables, we need to

1. understand the area element, dA, in terms of s and t
2. transform the integrand (with the change of variable formulas)
3. write the new region (in the st-plane) with a vertically or horizontally simple description

The following activity helps to illustrate the idea behind describing the new region in the st-plane.

Activity 12.9.2 In this activity, we will be looking at what happens to the region D, as shown in Figure 12.9.5, under the change of coordinates given by

$$x = \sqrt{st} \quad \text{and} \quad y = \sqrt{\frac{s}{t}}$$

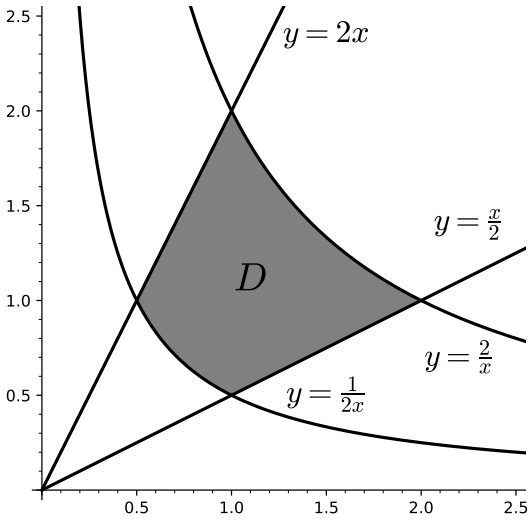


Figure 12.9.5

(a) Find the coordinates of the four corner points of the region D.

(b) Given the form of the change of variable equations above, it is not obvious how to change (x, y) points into (s, t) points. Show that the following equations are satisfied by our change of variable equations:

$$yx = s \quad \text{and} \quad \frac{y}{x} = \frac{1}{t}$$

Use these new relationships between our coordinates to convert each of these four corner points from xy-coordinates to st-coordinates and draw these points on Figure 12.9.6.

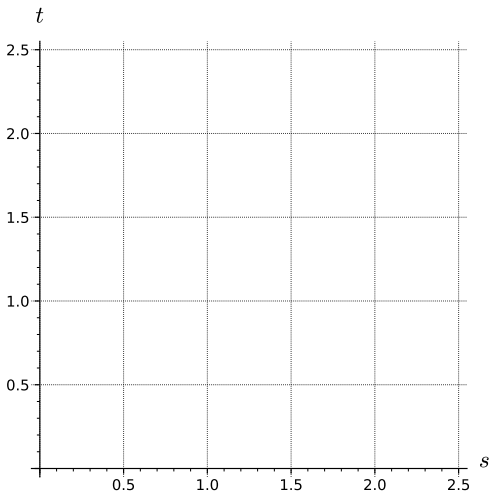


Figure 12.9.6

(c) We hope that the curved boundaries of the region D get transformed to horizontal and vertical lines in the st-plane, as is suggested by the result of the previous task. Let's look at how each of the boundary curves gets transformed by our change of variables formula.

Use the change of variables equations to write $y = 2x$ in terms of s and t coordinates and solve for either s or t (as would make sense). Plot the resulting line on Figure 12.9.6.

(d) Write each of the other boundary curves of D ($y = \frac{2}{x}, y = \frac{1}{2x}, y = \frac{x}{2}$) in terms of s and t coordinates, solve for either s or t (as would make sense), and plot the resulting lines on Figure 12.9.6.

(e) We have now confirmed that our irregular region D gets mapped to a rectangle in the st-plane by the change of variables formula given. Some areas of the region D needed to be stretched more than others in order to straighten out our sides into the nice rectangle you have drawn on Figure 12.9.6. Write a few sentences to describe what part of D you think is stretched the most in the transformation from xy-coordinates to st-coordinates.

In Activity 12.9.2, we used a clever relationship to convert our corner points from xy-coordinates to st-coordinates, but in general you would need to solve a system of equations to find the corresponding (s, t) points. For example, if we have change of coordinate equations given by $x = 3s - 2t$ and $y = s + 4t$, then the transformation of the point $(x, y) = (1, 2)$, would come from solving the system of equations $1 = 3s - 2t$ and $2 = s + 4t$.

If we shift our attention to the area element associated with a change of variables, we will need to look at how the area is stretched by the transformation between xy-coordinates to st-coordinates. Conceptually, derivatives will allow us to measure how quickly coordinates are changing in the context of a change of variables. Let's look back at the single variable example stated in the introduction to this section.

Example 12.9.7 If we look at the integral

$$\int_0^2 2x(x^2+1)^3\,dx,$$

and consider the change of variable $u = x^2 + 1$ with $du = 2x\,dx$, then $x = 0$ implies $u = 1$ and $x = 2$ implies $u = 5$. So we have transformed the original integral in x into a new integral in u. In particular,

$$\int_0^2 (x^2+1)^3(2x\,dx) = \int_1^5 u^3\,du$$

which we can visualize with the following plots of area under curves.

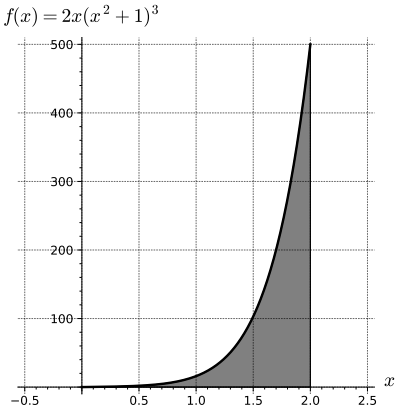


Figure 12.9.8

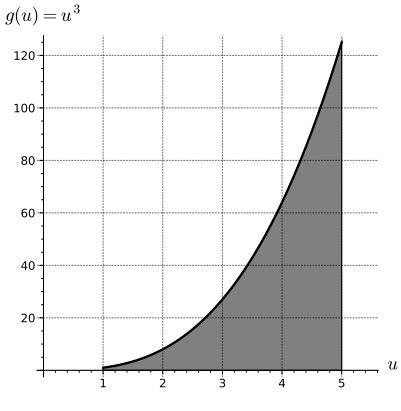


Figure 12.9.9

Both of these shaded regions have the same area but you can see that the scales on the vertical and horizontal axes changed. The vertical axis on each graph corresponds to the output of the function being integrated and the transformation comes from transforming the integrand. The bounds on the

integral come from the conversion of the interval $[0, 2]$ in x-coordinates to the interval $[1, 5]$ in u-coordinates. The final part of this substitution/change of variable problem is converting the differential. ▲

We have seen how in a two dimensional setting, we can convert the region of integration (Activity 12.9.2) and the integrand (the function being integrated) using our change of variable equations. As was suggested by the previous example, the partial derivatives of x and y with respect to s or t will measure how each variable (x or y) is stretched with respect to small a small change in the inputs (s or t). We will need to use vectors to quantify how these individual coordinate changes will amount to changes in the way we measure area.

The following result summarizes how to find the area element corresponding to a change of variables in general. We have included the details of the vector argument as a proof to this result.

Key Idea 12.9.10 *Given a change of coordinates given by* $x = x(s,t)$ *and* $y = y(s,t)$*, the associated area element is given by* $dA = dy\,dx = \left|\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s}\right| ds\,dt$.

Proof. We partition a rectangular domain in the st system into subrectangles. Let $T = [a, b] \times [a + \Delta s, b + \Delta t]$ be one of these subrectangles. Then we transform this into a region T' in the standard xy Cartesian coordinate system. The region T' is called the *image* of T; the region T is the *pre-image* of T'. Although the sides of this xy region T' aren't necessarily straight (linear), we will approximate the element of area dA for this region with the area of the parallelogram whose sides are given by the vectors $\vec{v}$ and $\vec{w}$, where $\vec{v}$ is the vector from $(x(a, b), y(a, b))$ to $(x(a + \Delta s, b), y(a + \Delta s, b))$, and $\vec{w}$ is the vector from $(x(a, b), y(a, b))$ to $(x(a, b + \Delta t), y(a, b + \Delta t))$.

An example of an image T' in the xy-plane that results from a transformation of a rectangle T in the st-plane is shown in Figure 12.9.11.

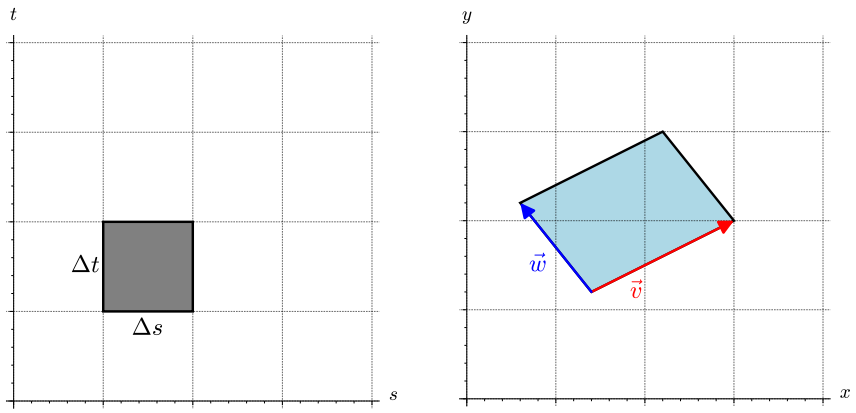


Figure 12.9.11 Approximating an area of an image resulting from a transformation

The components of the vector $\vec{v}$ are

$$\vec{v} = \langle x(a + \Delta s, b) - x(a, b), y(a + \Delta s, b) - y(a, b), 0\rangle$$

and for $\vec{w}$ the components are

$$\vec{w} = \langle x(a, b + \Delta t) - x(a, b), y(a, b + \Delta s) - y(a, b), 0\rangle\,.$$

Slightly rewriting $\vec{v}$ and $\vec{w}$, we have

$$\vec{v} = \left\langle \frac{x(a + \Delta s, b) - x(a, b)}{\Delta s}, \frac{y(a + \Delta s, b) - y(a, b)}{\Delta s}, 0 \right\rangle \Delta s, \text{ and}$$

$$\vec{w} = \left\langle \frac{x(a,b+\Delta t)-x(a,b)}{\Delta t}, \frac{y(a,b+\Delta s)-y(a,b)}{\Delta t}, 0 \right\rangle \Delta t.$$

For small Δs and Δt, the definition of the partial derivative tells us that

$$\vec{v} \approx \left\langle \frac{\partial x}{\partial s}(a,b), \frac{\partial y}{\partial s}(a,b), 0 \right\rangle \Delta s \qquad \text{and} \qquad \vec{w} \approx \left\langle \frac{\partial x}{\partial t}(a,b), \frac{\partial y}{\partial t}(a,b), 0 \right\rangle \Delta t.$$

Recall that the area of the parallelogram with sides $\vec{v}$ and $\vec{w}$ is the length of the cross product of the two vectors, $|\vec{v} \times \vec{w}|$. From this, we observe that

$$\begin{aligned} \vec{v} \times \vec{w} &\approx \left\langle \frac{\partial x}{\partial s}(a,b), \frac{\partial y}{\partial s}(a,b), 0 \right\rangle \Delta s \times \left\langle \frac{\partial x}{\partial t}(a,b), \frac{\partial y}{\partial t}(a,b), 0 \right\rangle \Delta t \\ &= \left\langle 0,\ 0,\ \frac{\partial x}{\partial s}(a,b)\frac{\partial y}{\partial t}(a,b) - \frac{\partial x}{\partial t}(a,b)\frac{\partial y}{\partial s}(a,b) \right\rangle \Delta s\, \Delta t. \end{aligned}$$

Finally, by computing the magnitude of the cross product, we see that

$$\begin{aligned} |\vec{v} \times \vec{w}| &\approx \left| \left\langle 0, 0, \frac{\partial x}{\partial s}(a,b)\frac{\partial y}{\partial t}(a,b) - \frac{\partial x}{\partial t}(a,b)\frac{\partial y}{\partial s}(a,b) \right\rangle \Delta s\, \Delta t \right| \\ &= \left| \frac{\partial x}{\partial s}(a,b)\frac{\partial y}{\partial t}(a,b) - \frac{\partial x}{\partial t}(a,b)\frac{\partial y}{\partial s}(a,b) \right| \Delta s\, \Delta t. \end{aligned}$$

Therefore, as the number of subdivisions increases without bound in each direction, Δs and Δt both go to zero, and we have

$$dA = \left| \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s} \right| ds\, dt. \tag{12.9.3}$$

■

Key Idea 12.9.10 determines the general change of variable formula in a double integral, and we can now convert all the parts of the double integral to s and t coordinates.

$$\iint_T f(x,y)\, dy\, dx = \iint_{T'} f(x(s,t), y(s,t)) \left| \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s} \right| ds\, dt$$

Definition 12.9.12 The quantity

$$\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s}$$

is called the *Jacobian*, and we denote the Jacobian using the shorthand notation

$$\frac{\partial(x,y)}{\partial(s,t)} = \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s}.$$

◆

Recall from Section 9.5 that we can also write this Jacobian as the determinant of the 2×2 matrix $\begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{bmatrix}$. Note that, as discussed in Section 9.5, the absolute value of the determinant of $\begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{bmatrix}$ is the area of the parallelogram determined by the vectors $\vec{v}$ and $\vec{w}$, and so the area element dA in xy-coordinates is also represented by the area element $\left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\, dt$ in st-coordinates, and $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$ is the factor by which the transformation magnifies area.

Example 12.9.13 We will look at the Jacobian for the change of variables in Activity 12.9.2. Given

$$x = \sqrt{st} \quad \text{and} \quad y = \sqrt{\frac{s}{t}}$$

we will need to find the partial derivatives of these with respect to both s and t.

$$\begin{aligned}
\frac{\partial x}{\partial s} &= \sqrt{t}\left(\frac{1}{2}\frac{1}{\sqrt{s}}\right) &&= \frac{1}{2}\sqrt{\frac{t}{s}} \\
\frac{\partial x}{\partial t} &= \sqrt{s}\left(\frac{1}{2}\frac{1}{\sqrt{t}}\right) &&= \frac{1}{2}\sqrt{\frac{s}{t}} \\
\frac{\partial y}{\partial s} &= \sqrt{\frac{1}{t}}\left(\frac{1}{2}\frac{1}{\sqrt{s}}\right) &&= \frac{1}{2}\sqrt{\frac{1}{st}} \\
\frac{\partial y}{\partial t} &= \sqrt{s}\left(-\frac{1}{2}\frac{1}{\sqrt{t^3}}\right) &&= -\frac{1}{2}\sqrt{\frac{s}{t^3}}
\end{aligned}$$

The Jacobian will then be

$$\det\left(\begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{bmatrix}\right) = \det\left(\begin{bmatrix} \frac{1}{2}\sqrt{\frac{t}{s}} & \frac{1}{2}\sqrt{\frac{s}{t}} \\ \frac{1}{2}\sqrt{\frac{1}{st}} & -\frac{1}{2}\sqrt{\frac{s}{t^3}} \end{bmatrix}\right)$$

which will evaluate to

$$\left(\frac{1}{2}\sqrt{\frac{t}{s}}\right)\left(-\frac{1}{2}\sqrt{\frac{s}{t^3}}\right) - \left(\frac{1}{2}\sqrt{\frac{s}{t}}\right)\left(\frac{1}{2}\sqrt{\frac{1}{st}}\right) = -\frac{1}{2t}$$

It is important to note a couple of interesting features of our Jacobian. First, the Jacobian is negative because in this example the region is flipped; Specifically the "lower" and "upper" bounds in Figure 12.9.5 given by $y = \frac{x}{2}$ and $y = 2x$ are respectively transformed to $t = 2$ and $t = \frac{1}{2}$. In other words, the top and bottom edges get flipped by the transformation which results in this area transformation being negative. Notice that we take the absolute value of the Jacobian when considering the area element because we only care about how the area is deformed in the transformation given by the change of coordinates (and do not care about any flips that occur as part of the transformation).

The second feature is that the Jacobian is not constant, which means that the deformation of the area does not occur the same at every part of our region. In particular, for values where t is great than one, we are multiplying by a smaller factor, which means the region is expanded as part of the transformation from xy to st coordinates. For values where t is less than one, we are multiplying by a larger factor, which means the region is shrunk as part of the transformation from xy to st coordinates. ▲

To summarize our work on change of variables and double integrals, the preceding change of variable formula we have derived is stated here.

Change of Variables in a Double Integral.

Suppose a change of variables $x = x(s,t)$ and $y = y(s,t)$ transforms a closed and bounded region R in the st-plane into a closed and bounded region R' in the xy-plane. Under modest conditions (that are studied

in advanced calculus), it follows that

$$\iint_{R'} f(x,y)\,dA = \iint_{R} f(x(s,t),y(s,t)) \left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\,dt.$$

Activity 12.9.3 In this activity, we will look at a few familiar Jacobians and their interpretations.

(a) Find the Jacobian for the transformation given by $x = 3s-1$ and $y = \frac{t}{3}+2$. Write a couple of sentences about why the area of a region will not change with respect to this change of coordinates.

(b) Find the Jacobian when changing from rectangular to polar coordinates. Write a couple of sentences to connect to our work on double integrals in polar coordinates.

Given a particular double integral, it is natural to ask, "how can we find a useful change of variables?" There are two general factors to consider: if the integrand is particularly difficult, we might choose a change of variables that would make the integrand easier; or, given a complicated region of integration, we might choose a change of variables that transforms the region of integration into one that has a simpler form. These ideas are illustrated in the next activities.

Activity 12.9.4 Consider the problem of finding the area of the region D' defined by the ellipse $x^2+\frac{y^2}{4} = 1$. In this activity, we will look at what happens when we use a change of variables so that the pre-image of the domain is a circle.

(a) Let $x(s,t) = s$ and $y(s,t) = 2t$. Explain why the pre-image of the original ellipse (which lies in the xy plane) is the circle $s^2+t^2 = 1$ in the st-plane.

(b) Recall that the area of the ellipse D' is determined by the double integral $\iint_{D'} 1\,dA$. Explain why

$$\iint_{D'} 1\,dA = \iint_{D} 2\,ds\,dt$$

where D is the disk bounded by the circle $s^2+t^2 = 1$. In particular, explain the source of the "2" in the st integral.

(c) Without evaluating any of the integrals present, explain why the area of the original elliptical region D' is 2π.

Activity 12.9.5 Let D' be the region in the xy-plane bounded by the lines $y = 0$, $x = 0$, and $x+y = 1$. We will evaluate the double integral

$$\iint_{D'} \sqrt{x+y}(x-y)^2\,dA \tag{12.9.4}$$

with a change of variables.

(a) Sketch the region D' in the xy-plane.

(b) We would like to make a substitution that makes the integrand easier to antidifferentiate. Let $s = x+y$ and $t = x-y$. Explain why this should make antidifferentiation easier by making the corresponding substitutions and writing the new integrand in terms of s and t.

(c) Solve the equations $s = x + y$ and $t = x - y$ for x and y. (Doing so determines the standard form of the transformation, since we will have x as a function of s and t, and y as a function of s and t.)

(d) To actually execute this change of variables, we need to know the st-region D that corresponds to the xy-region D'.

i. What st equation corresponds to the xy equation $x + y = 1$?

ii. What st equation corresponds to the xy equation $x = 0$?

iii. What st equation corresponds to the xy equation $y = 0$?

iv. Sketch the st region D that corresponds to the xy domain D'.

(e) Make the change of variables indicated by $s = x + y$ and $t = x - y$ in the double integral (12.9.4) and set up an iterated integral in st variables whose value is the original given double integral. Finally, evaluate the iterated integral.

12.9.4 Change of Variables in a Triple Integral

The details of an argument for the change of variable formula for triple integrals are more complicated the two dimension case we have just discussed and we omit those here. The general process, though, is the same as the two-dimensional case; Namely, we will use partial derivatives and the determinant of a matrix of the partial derivatives to give the appropriate volume element for the new coordinate.

Given a solid S' in the xyz-coordinate system in $\mathbb{R}^3$, a change of variables transformation $x = x(s,t,u)$, $y = y(s,t,u)$, and $z = z(s,t,u)$ transforms S' into a region S in stu-coordinates. Any function $f = f(x,y,z)$ defined on S' can be considered as a function $f = f(x(s,t,u), y(s,t,u), z(s,t,u))$ in stu-coordinates defined on S. The volume element dV in xyz-coordinates cooresponds to a scaled volume element in stu-coordinates, where the scale factor is given by the absolute value of the Jacobian, $\frac{\partial(x,y,z)}{\partial(s,t,u)}$, which is the determinant of the 3×3 matrix

$$\begin{bmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} & \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial s} & \frac{\partial z}{\partial t} & \frac{\partial z}{\partial u} \end{bmatrix}.$$

(Recall that this determinant was introduced in Section 9.5.) That is, $\frac{\partial(x,y,z)}{\partial(s,t,u)}$ is given by

$$\frac{\partial x}{\partial s}\left[\frac{\partial y}{\partial t}\frac{\partial z}{\partial u} - \frac{\partial y}{\partial u}\frac{\partial z}{\partial t}\right] - \frac{\partial x}{\partial t}\left[\frac{\partial y}{\partial s}\frac{\partial z}{\partial u} - \frac{\partial y}{\partial u}\frac{\partial z}{\partial s}\right] + \frac{\partial x}{\partial u}\left[\frac{\partial y}{\partial s}\frac{\partial z}{\partial t} - \frac{\partial y}{\partial t}\frac{\partial z}{\partial s}\right].$$

Change of Variables in a Triple Integral.

Suppose a change of variables $x = x(s,t,u)$, $y = y(s,t,u)$, and $z = z(s,t,u)$ transforms a closed and bounded region S in stu-coordinates into a closed and bounded region S' in xyz-coordinates. Under modest conditions (that are studied in advanced calculus), the triple integral $\iiint_{S'} f(x,y,z)\,dV$ is equal to

$$\iiint_S f(x(s,t,u), y(s,t,u), z(s,t,u)) \left|\frac{\partial(x,y,z)}{\partial(s,t,u)}\right| ds\,dt\,du.$$

Activity 12.9.6 Consider the solid S' defined by the inequalities $0 \le x \le 2$, $\frac{x}{2} \le y \le \frac{x}{2} + 1$, and $0 \le z \le 6$. Consider the transformation defined by $s = \frac{x}{2}$, $t = \frac{x-2y}{2}$, and $u = \frac{z}{3}$. Let $f(x, y, x) = x - 2y + z$.

(a) The transformation turns the solid S' in xyz-coordinates into a box S in stu-coordinates. Apply the transformation to the boundries of the solid S' to find stu-coordinate descriptions of the box S.

(b) Compute and simplify the Jacobian $\frac{\partial(x,y,z)}{\partial(s,t,u)}$.

(c) Use the transformation to perform a change of variables and evaluate $\iiint_{S'} f(x, y, z)\, dV$ by evaluating

$$\iiint_S f(x(s,t,u), y(s,t,u), z(s,t,u)) \left|\frac{\partial(x,y,z)}{\partial(s,t,u)}\right| ds\, dt\, du$$

Summary

- If an integral is described in terms of one set of variables, we may write that set of variables in terms of another set of the same number of variables. If the new variables are chosen appropriately, the transformed integral may be easier to evaluate.
- The Jacobian is a scalar function that relates the area or volume element in one coordinate system to the corresponding element in a new system determined by a change of variables.

12.9.5 Exercises

1. Let $\Phi(u, v) = (7u+v, 4u+6v)$ be a map from the uv-plane to the xy-plane. Describe the image of the line $v = 4u$ under Φ in slope-intercept form.
$y =$ _____

2. Find the absolute value of the Jacobian, $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$, for the change of variables given by $x = 7s + 6t, y = 7s + 3t$
$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$_____

3. Find the Jacobian. $\frac{\partial(x,y,z)}{\partial(s,t,u)}$, where $x = 5s - t - 3u, y = 2t - 5s - 4u, z = 3t - 3s - 5u$.
$\frac{\partial(x,y,z)}{\partial(s,t,u)} =$ _____

4. Compute the Jacobian of :

$$\Phi(u, v) = (-(8u + 9v), 3u + 5v)$$

$\text{Jac}(\Phi) =$ _____

5. Compute the Jacobian of :

$$\Phi(r, \theta) = (3r\cos\theta, 7r\sin\theta)$$

$\text{Jac}(\Phi) =$ _____

6. Consider the transformation $T : x = \frac{24}{40}u - \frac{32}{40}v, \quad y = \frac{32}{40}u + \frac{24}{40}v$
A. Compute the Jacobian:
$\frac{\partial(x,y)}{\partial(u,v)} =$ _____
B. The transformation is linear, which implies that it transforms lines into lines. Thus, it transforms the square $S : -40 \le u \le 40, -40 \le v \le 40$

into a square $T(S)$ with vertices:

T(40, 40) = (_____, _____)
T(-40, 40) = (_____, _____)
T(-40, -40) = (_____, _____)
T(40, -40) = (_____, _____)

C. Use the transformation T to evaluate the integral $\iint_{T(S)} x^2+y^2\, dA$

7. Use the change of variables $s = y$, $t = y - x^2$ to evaluate $\int\int_R x\, dx\, dy$ over the region R in the first quadrant bounded by $y = 0$, $y = 36$, $y = x^2$, and $y = x^2 - 4$.

$\int\int_R x\, dx\, dy =$ _____

8. Use the change of variables $s = x+3y$, $t = y$ to find the area of the ellipse $x^2 + 6xy + 10y^2 \leq 1$.

area = _____

9. Use the change of variables $s = xy$, $t = xy^2$ to compute $\int_R xy^2\, dA$, where R is the region bounded by $xy = 3,\ xy = 5,\ xy^2 = 3,\ xy^2 = 5$.

$\int_R xy^2\, dA =$ _____

10. Find positive numbers a and b so that the change of variables $s = ax, t = by$ transforms the integral $\int\int_R dx\, dy$ into

$$\int\int_T \left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\, dt$$

for the region R, the rectangle $0 \leq x \leq 60$, $0 \leq y \leq 75$ and the region T, the square $0 \leq s, t \leq 1$.

$a =$ _____
$b =$ _____
What is $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$ in this case?
$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ _____

11. Find a number a so that the change of variables $s = x + ay, t = y$ transforms the integral $\int\int_R dx\, dy$ over the parallelogram R in the xy-plane with vertices $(0, 0)$, $(28, 0)$, $(-34, 11)$, $(-6, 11)$ into an integral

$$\int\int_T \left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\, dt$$

over a rectangle T in the st-plane.

$a =$_____
What is $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$ in this case?
$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ _____

12. In this problem we use the change of variables $x = 5s + t$, $y = s - 3t$ to compute the integral $\int_R (x + y)\, dA$, where R is the parallelogram with vertices $(x, y) = (0, 0)$, $(5, 1)$, $(7, -5)$, and $(2, -6)$.

First find the magnitude of the Jacobian, $\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ _____.
Then, with $a =$ _____, $b =$ _____,
$c =$ _____, and $d =$ _____,
$\int_R (x + y)\, dA = \int_a^b \int_c^d (\ ____\ s + ____\ t + ____\ )\ dt\, ds =$
_____

13. Find positive numbers a and b so that the change of variables $s = ax, t = by$ transforms the integral $\int\int_R dx\,dy$ into

$$\int\int_T \left|\frac{\partial(x,y)}{\partial(s,t)}\right| ds\,dt$$

for the region R, the elliptical region $x^2/4 + y^2/36 \le 1$ and the region T, the circle $s^2 + t^2 = 1$.

$a =$ _____

$b =$ _____

What is $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$ in this case?

$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ _____

14. Show that $T(u,v) = (u^2 - v^2, 2uv)$ maps the triangle $= \{(u,v) : 0 \le v \le u \le 1\}$ to the domain $\mathcal{D}$ bounded by $x = 0$, $y = 0$, and $y^2 = 4 - 4x$. Use T to evaluate

$$\iint_{\mathcal{D}} \sqrt{x^2 + y^2}\,dx\,dy$$

Answer : _____

15. Sketch the domain $\mathcal{D}$ bounded by $y = x^2$, $y = \frac{1}{2}x^2$, and $y = 4x$. Use a change of variables with the map $x = uv$, $y = u^2$ to calculate

$$\iint_{\mathcal{D}} y^{-1}\,dx\,dy$$

This is an improper integral since $f(x,y) = y^{-1}$ is undefined at $(0,0)$, but it becomes proper after changing variables.

$\iint_{\mathcal{D}} y^{-1}\,dx\,dy =$ _____

16. Let $I = \iint_{\mathcal{D}}(x^2 - y^2)\,dx\,dy$, where

$$\mathcal{D} = \{(x,y) : 2 \le xy \le 6, 0 \le x - y \le 5, x \ge 0, y \ge 0\}$$

Show that the mapping $u = xy$, $v = x - y$ maps $\mathcal{D}$ to the rectangle $\mathcal{R} = [2,6] \times [0,5]$.

(**a**) Compute $\partial(x,y)/\partial(u,v)$ by first computing $\partial(u,v)/\partial(x,y)$.

(**b**) Use the Change of Variables Formula to show that I is equal to the integral of $f(u,v) = v$ over $\mathcal{R}$ and evaluate.

(**a**) $\frac{\partial(x,y)}{\partial(u,v)} =$ _____

(**b**) $I =$ _____

17. Let D' be the region in the xy-plane that is the parallelogram with vertices $(3,3)$, $(4,5)$, $(5,4)$, and $(6,6)$.

(a) Sketch and label the region D' in the xy-plane.

(b) Consider the integral $\iint_{D'}(x + y)\,dA$. Explain why this integral would be difficult to set up as an iterated integral.

(c) Let a change of variables be given by $x = 2u + v$, $y = u + 2v$. Using substitution or elimination, solve this system of equations for u and v in terms of x and y.

(d) Use your work in (c) to find the pre-image, D, which lies in the uv-plane, of the originally given region D', which lies in the xy-plane.

For instance, what uv point corresponds to $(3, 3)$ in the xy-plane?

(e) Use the change of variables in (c) and your other work to write a new iterated integral in u and v that is equivalent to the original xy integral $\iint_{D'}(x + y)\,dA$.

(f) Finally, evaluate the uv integral, and write a sentence to explain why the change of variables made the integration easier.

18. Consider the change of variables

$$x(\rho, \theta, \phi) = \rho\sin(\phi)\cos(\theta) \quad y(\rho, \theta, \phi) = \rho\sin(\phi)\sin(\theta) \quad z(\rho, \theta, \phi) = \rho\cos(\phi),$$

which is the transformation from spherical coordinates to rectangular coordinates. Determine the Jacobian of the transformation. How is the result connected to our earlier work with iterated integrals in spherical coordinates?

19. In this problem, our goal is to find the volume of the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1$.

(a) Set up an iterated integral in rectangular coordinates whose value is the volume of the ellipsoid. Do so by using symmetry and taking 8 times the volume of the ellipsoid in the first octant where x, y, and z are all nonnegative.

(b) Explain why it makes sense to use the substitution $x = as$, $y = bt$, and $z = cu$ in order to make the region of integration simpler.

(c) Compute the Jacobian of the transformation given in (b).

(d) Execute the given change of variables and set up the corresponding new iterated integral in s, t, and u.

(e) Explain why this new integral is better, but is still difficult to evaluate. What additional change of variables would make the resulting integral easier to evaluate?

(f) Convert the integral from (d) to a new integral in spherical coordinates.

(g) Finally, evaluate the iterated integral in (f) and hence determine the volume of the ellipsoid.

Chapter 13

Vector Calculus

13.1 Integrating Multivariable Vector-Valued Functions

In Chapter 10, we looked at the calculus of functions with one scalar input and multiple outputs, usually organized as a vector. In Chapter 11 we looked at measuring changes involving functions with multiple input variables and a single scalar output. In Chapter 12 we looked at integration involving scalar valued multivariable functions. In this chapter, we will look at combining all of our multivariable function tools, as well as our vector tools, to look at the calculus of functions with multiple input variables and multiple output variables. This chapter is significantly longer than others in this text because there are many different measurements to investigate, as well as several fundamental thereoms used in applications of these new measurements.

The example of wind has been used several times in this text to describe a measurement that would require multiple inputs and outputs. In particular, wind will change over the points in some region (as well as changing over time). Additionally, wind requires a measurement of strength in each direction which is typically stored as a vector. Thus wind has multiple scalar inputs corresponding to the location of the measurement and multiple scalar outputs that measure the strength in coordinate directions.

13.2 Vector Fields

Motivating Questions

- What is a vector field?
- What are some familiar contexts in which vector fields arise?
- How do we draw a vector field?
- How do gradients of functions with partial derivatives connect to vector fields?

13.2.1 Introduction

Vectors have played a central role in our study of multivariable calculus. We know how to do operations on vectors (addition, scalar multiplication, dot

product, etc.), and we have seen how vectors can be used to describe curves in $\mathbb{R}^2$ and $\mathbb{R}^3$. The examples of using vectors to describe curves was our first example of a vector-valued function. In Definition 10.2.4 a curve is traced by the terminal point of $\vec{r}(t)$, a function that has a real number as an input and produces a vector in $\mathbb{R}^2$ or $\mathbb{R}^3$. In this section, we will expand our understanding of vector-valued functions to take a point (x, y) in $\mathbb{R}^2$ (or a point (x, y, z) in $\mathbb{R}^3$) as an input and produce a vector (typically in $\mathbb{R}^2$ or $\mathbb{R}^3$, respectively) as output.

Preview Activity 13.2.1 It's common when discussing weather to talk about the wind **speed**, but as any student who has gotten this far in the text will know, this nomenclature is imprecise. It's not terribly helpful to tell someone the wind is blowing at 10 $\frac{\text{km}}{\text{h}}$ without telling them the direction in which the wind is blowing. If you're trying to make a decision based on what the wind is doing, you need to know about the direction as well. For instance, if you are taking off in a hot air balloon, the wind direction will determine which direction the chase team should start going to keep track of you. Because of the swirling nature of wind, it makes sense to give the wind **velocity** at each point in a region (two-dimensional or three-dimensional).

(a) Suppose that given a point (x, y) in the plane, you know that the wind velocity at that point is given by the vector $\langle y, x\rangle$. For example, we'd then know that at the point $(1, -1)$, the wind velocity is $\langle -1, 1\rangle$. We will give the wind velocity as a function $\vec{F}$, where $\vec{F}(x, y) = \langle y, x\rangle$. In the table below, fill in the wind velocity vectors for the given points.

(x, y)	$(2, 1)$	$(0, 0)$	$(-1, 2)$	$(3, -1)$	$(-2, -1)$
$\vec{F}(x, y)$					

(b) Suppose that we associate the vector $\vec{G}(x, y) = -x\hat{j}$ to a point (x, y) in the plane. Complete the table below by giving the vector associated to each of the given points.

(x, y)	$(-2, 0)$	$(-1, 2)$	$(0, -2)$	$(2, 3)$	$(3, 2)$	$(-1, 0)$	$(1, 3)$
$\vec{G}(x, y)$							

(c) A table of values of these vector-valued functions is useful to understand the input vs. output nature of a vector field as a function, but perhaps even better is a method of visualizing the vector outputs. A good picture is worth a thousand words (or numbers). Returning to our analogy of the output vector for our vector field being wind velocity, if $\vec{F}(2, 1) = \langle 1, 2\rangle$, this means that at the location $(2, 1)$ the wind is moving in the direction given by $\langle 1, 2\rangle$. Thus, we draw the output vector $\langle 1, 2\rangle$ with its initial point at $(2, 1)$.

Using the first set of axes in Figure 13.2.1, plot the vectors $\vec{F}(x, y)$ for the five points in the table in part a. The example $\vec{F}(1, -1) = \langle -1, 1\rangle$ is drawn for you.

(d) Using the second set of axes in Figure 13.2.1, plot the vectors $\vec{G}(x, y)$ for the eight points in the table in part b.

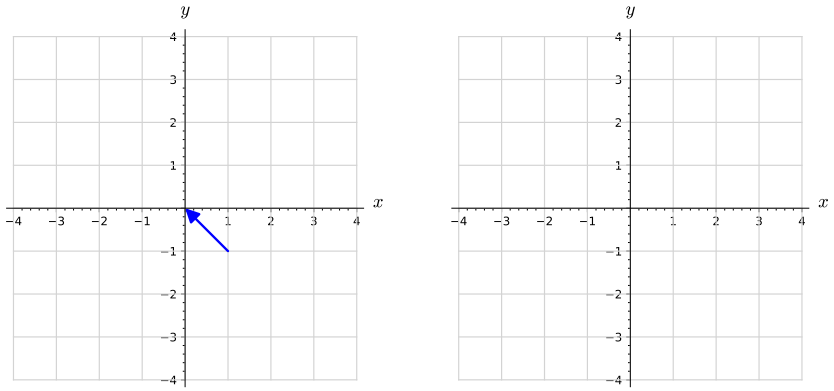


Figure 13.2.1 Axes for plotting some vectors from $\vec{F}(x, y)$ and $\vec{G}(x, y)$.

13.2.2 Examples of Vector Fields

As Preview Activity 13.2.1 showed, a **velocity vector field** is an example of a scenario where associating a vector to each point in a region is useful. We denote such a vector field by $\vec{F}(x, y)$ or $\vec{F}(x, y, z)$, where the vector associated to the point (x, y) or (x, y, z) is the velocity of something at that point. Wind velocity is one example, but another example would be the velocity of a flowing fluid. Figure 13.2.2 shows such a velocity vector field. Technically, it only shows some of the vectors in the vector field, since the figure would be unintelligible if all of the vectors were shown. This is illustrated by the inset in the upper left corner, which gives a better picture of what we would see if we zoomed in on the red square of the main figure.

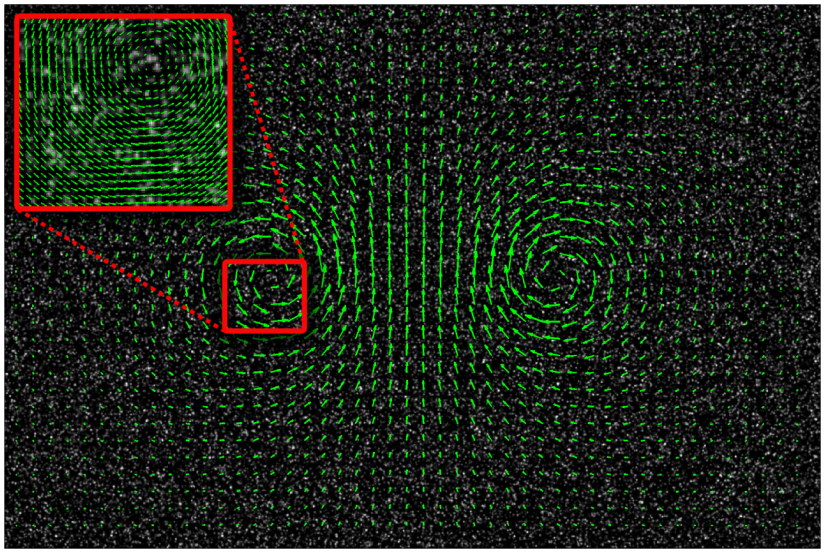
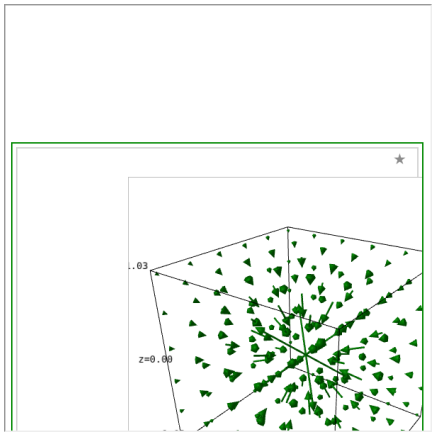


Figure 13.2.2 An illustration of some of the vectors in a fluid velocity vector field. "PIVlab multipass" by Willa Licensed under CC-BY-SA 3.0 via Wikimedia Commons.

Force fields, such as those created by gravity, are also examples of vector fields. For example, the earth exerts a gravitational force on objects which is directed from the center of the object to the center of the earth. The magnitude of the force vector is determined by the distance between the object and the earth (by an reciprocal squared relationship.) An illustration of this vector

field can be seen in Figure 13.2.3, where the earth is positioned at the origin, but not shown. Notice that the vectors get shorter as the distance from the origin increases, reflecting the fact that the gravitational force is weaker at larger distances from the origin (Earth).



Standalone
In Context
Embed

Figure 13.2.3 Gravitational vector field.

13.2.3 Mathematical Vector Fields

As suggested in the introduction and Preview Activity 13.2.1, vector fields can be specified using the notation of functions and vectors.

Definition 13.2.4 A **vector field** in 2-space is a function whose value at a point (x, y) is a 2-dimensional vector $\vec{F}(x, y)$. Similarly, in 3-space, a vector field is a function $\vec{F}(x, y, z)$ whose value at the point (x, y, z) is a 3-dimensional vector. ◆

Since $\vec{F}(x, y, z)$ is a vector, it has $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ components. Each of these components is a scalar function of the point (x, y, z), and so we will often write

$$\vec{F}(x, y, z) = F_1(x, y, z)\hat{\imath} + F_2(x, y, z)\hat{\jmath} + F_3(x, y, z)\hat{k}$$

For example, if $\vec{F}(x, y, z) = \langle x^2, xy\sin(z), y^3\rangle$, then the component functions of $\vec{F}$ would be $F_1(x, y, z) = x^2$, $F_2(x, y, z) = xy\sin(z)$, and $F_3(x, y, z) = y^3$. Any time we are considering a vector field $\vec{F}(x, y, z)$, the definitions of functions F_1, F_2, and F_3 should be inferred in this manner. (For a vector field $\vec{F}(x, y)$ in 2-space, we only have the functions F_1 and F_2, which are defined analogously.)

13.2.4 Plotting Vector Fields

Preview Activity 13.2.1 gave you a chance to plot some vectors in the vector fields $\vec{F}(x, y) = \langle y, x\rangle$ and $\vec{G}(x, y) = \langle 0, -x\rangle$. It would be impossible to sketch *all* of the vectors in these vector fields, since there is one for every point in the plane. In fact, even sketching many more of the vectors than you were asked to in the preview activity rapidly becomes tedious. Fortunately, computers can do a great job of making such sketches. One thing to keep in mind, however, is that the magnitudes of the vectors in computer plots are typically scaled, including plots of vector fields we will encounter later in this text. To illustrate this, consider the two plots of the vector field $\vec{F}(x, y) = y\hat{\imath} + x\hat{\jmath}$ in Figure 13.2.5.

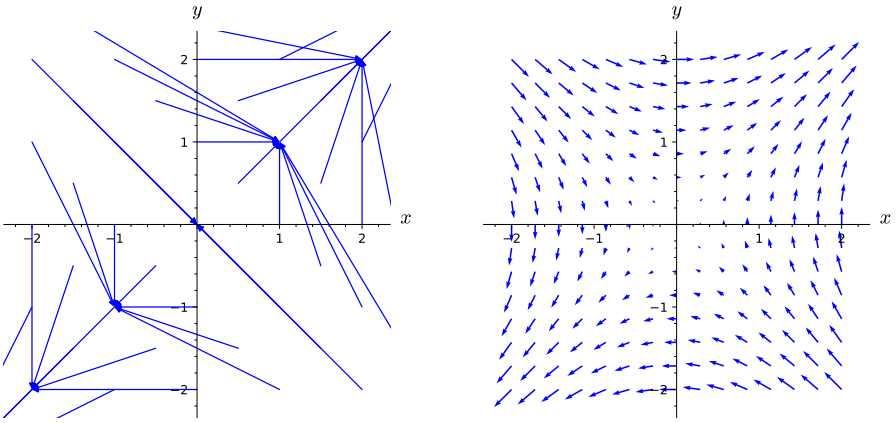


Figure 13.2.5 Two plots of $\vec{F}(x, y) = y\hat{\imath} + x\hat{\jmath}$ from Sage

The left plot shows some of the vectors and accurately depicts all of their magnitudes, making the figure very hard to understand, especially along the lines $y = x$ and $y = -x$. The plot on the right, however, uses a uniform rescaling to make the figure easier to read. As before, each vector's direction is completely accurate, but now the magnitudes are much smaller. However, the *relative* magnitudes are preserved, helping us to see that vectors farther from the origin have larger magnitude than those closer to the origin.

Activity 13.2.2 The plot in Figure 13.2.6 illustrates the vector field $\vec{F}(x, y) = y\hat{\imath} - x\hat{\jmath}$.

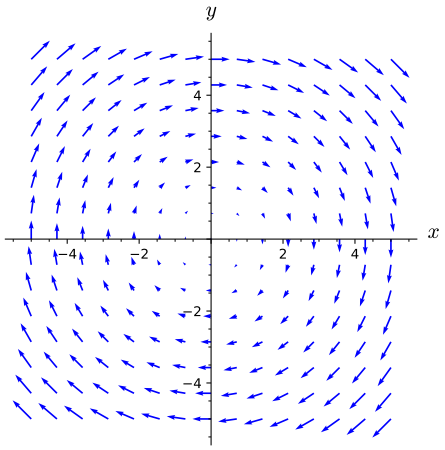


Figure 13.2.6 The vector field $y\hat{\imath} - x\hat{\jmath}$

(a) Starting with one of the vectors near the point $(2, 0)$, sketch a curve that follows the direction of the vector field $\vec{F}$. To help visualize what you are doing, it may be useful to think of the vector field as the velocity vector field for some flowing water and that you are imagining tracing the path that a tiny particle inserted into the water would follow as the water moves it around.

(b) Repeat the previous step for at least two other starting points not on the curve you previously sketched.

(c) What shape do the curves you sketched in the previous two steps form?

(d) Verify that $\vec{F}(x, y)$ is orthogonal to $\langle x, y \rangle$.

(e) Calculate the gradient of the function $f(x, y) = x^2 + y^2$ and write a sentence comparing your result to the vector $x\hat{\imath} + y\hat{\jmath}$.

(f) Write a sentence describing the geometric relationship between $\vec{F}(x, y)$ and a circle centered at the origin. What is the relationship between $\left\|\vec{F}(x, y)\right\|$ and the radius of that circle?

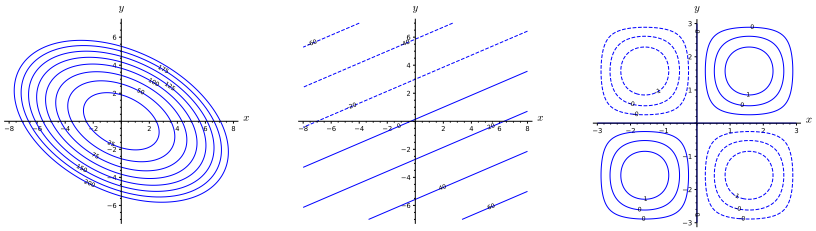
In the previous activity, we looked a special type of curve in a vector field, namely the curve that flows with the output of the vector field. Geometrically, the output vectors of the vector field will be tangent to the flow curves. These flow curves come up in several physical situations, including as solution curves to a system of differential equations. We will not discuss these applications here but you should should look out for uses of vector fields in the next few math courses.

13.2.5 Gradient Vector Fields

Without using the terminology, we've actually already encountered one very important family of vector fields a number of times. Given a function f of two or three (or more!) variables, the gradient of f is a vector field, since for any point where f has first-order partial derivatives, ∇f assigns a vector to that point (look at Subsection 11.8.4 for a review).

Activity 13.2.3

(a) In Figure 13.2.7 there are three sets of axes showing level curves for functions f, g, and h, respectively. Sketch at least six vectors in the gradient vector field for each function. In making your sketches, you don't have to worry about getting vector magnitudes precise, but you should ensure that the relative magnitudes (and directions) are correct for each function independently.



(a) f **(b)** g **(c)** h

Figure 13.2.7 Three sets of level curves

(b) Verify that $\vec{F}(x, y) = \langle 6xy, 3x^2 + 9\sqrt{y}\rangle$ is a gradient vector field by finding a function f such that $\nabla f(x, y) = \vec{F}(x, y)$. For reasons originating in physics, such a function f is called a **potential function** for the vector field $\vec{F}$.

(c) Is the function f found in part b unique? That is, can you find another function g such that $\nabla g(x, y) = \vec{F}(x, y)$ but $f \neq g$?

(d) Is the vector field $\vec{F}(x, y) = 6xy\hat{\imath} + (2x + 9\sqrt{y})\hat{\jmath}$ a gradient vector field? Why or why not?

In this text, we will most often refer to a vector field that can be written as the gradient of some scalar valued function to be a gradient field. In other sources and fields, a vector field $\vec{F}$ where $\vec{F} = \nabla f$ for some scalar valued

f is called a conservative vector field. We will describe this term more in Section 13.5 but mention this synonym term here as well.

Summary

- A 2-dimensional vector field is a function defined on part of $\mathbb{R}^2$ whose value is a 2-dimensional vector. A 3-dimensional vector field is a function defined on part of $\mathbb{R}^3$ whose value is a 3-dimensional vector.
- A vector field is typically described in terms of its multivariable component functions, $\vec{F}(x,y) = \langle f(x,y), g(x,y)\rangle = f(x,y)\hat{\imath} + g(x,y)\hat{\jmath}$ or in 3D

$$\begin{aligned}\vec{F}(x,y,z) &= \langle f(x,y,z), g(x,y,z), h(x,y,z)\rangle \\ &= f(x,y,z)\hat{\imath} + g(x,y,z)\hat{\jmath} + h(x,y,z)\hat{k}.\end{aligned}$$

- Vector fields arise in familiar contexts such as wind velocity, fluid velocity, and gravitational force.
- Vector fields are generally plotted in ways that ensure the direction and relative magnitudes of the vectors sketched are correct instead of ensuring that each vector's magnitude is depicted correctly.
- The gradient of a function f of two or three variables is a vector field defined wherever f has partial derivatives.

13.2.6 Exercises

1. Sketch the vector field $\vec{F}(\vec{r}) = 2\vec{r}$ in the plane, where $\vec{r} = \langle x, y\rangle$. Select all that apply.

- ☐ All the vectors point toward the origin.
- ☐ All the vectors point away from the origin.
- ☐ The length of each vector is 2.
- ☐ The vectors increase in length as you move away from the origin.
- ☐ All the vectors point in the same direction.
- ☐ The vectors decrease in length as you move away from the origin.

2. Sketch the vector field $\vec{F}(x,y) = 3\vec{\imath} + 5\vec{\jmath}$ in the plane. Select all that apply.

- ☐ The vectors decrease in length as you move away from the origin.
- ☐ All the vectors point away from the origin.
- ☐ All the vectors point toward the origin.
- ☐ All the vectors point in the same direction.
- ☐ The length of each vector is 34.
- ☐ The vectors increase in length as you move away from the origin.
- ☐ The length of each vector is $\sqrt{34}$.

3. Sketch the vector field $\vec{F}(x,y) = -\dfrac{\vec{r}}{||\vec{r}||^3}$ in the plane, where $\vec{r} = \langle x, y\rangle$. Select all that apply.

- ☐ The length of each vector is 1.
- ☐ The vectors increase in length as you move away from the origin.
- ☐ All the vectors point toward the origin.
- ☐ The vectors decrease in length as you move away from the origin.
- ☐ All the vectors point away from the origin.

4. Let $\vec{F} = x\,\vec{\imath} + (x+y)\,\vec{\jmath} + (x-y+z)\,\vec{k}$. Let the line l be $x = -(1+4t)$, $y = 1+5t$, $z = -(4+3t)$.

(a) Find a point $P = (x_0, y_0, z_0)$ where $\vec{F}$ is parallel to l.

$P =$ _____

Find a point $Q = (x_1, y_1, z_1)$ at which $\vec{F}$ and l are perpendicular.

$Q =$ _____

Give an equation for the set of all points at which $\vec{F}$ and l are perpendicular.

equation: _____

5. For each of the following vector fields $\boldsymbol{F}$, decide whether it is conservative or not by computing the appropriate first order partial derivatives. Type in a potential function f (that is, $\nabla f = \mathbf{F}$) with $f(0,0) = 0$. If it is not conservative, type N.

A. $\mathbf{F}(x,y) = (16x - 6y)\,\mathbf{i} + (-6x + 12y)\,\mathbf{j}$

$f(x,y) =$ _____

B. $\mathbf{F}(x,y) = 8y\mathbf{i} + 9x\mathbf{j}$

$f(x,y) =$ _____

C. $\mathbf{F}(x,y) = (8\sin y)\,\mathbf{i} + (-12y + 8x\cos y)\,\mathbf{j}$

$f(x,y) =$ _____

Note: Your answers should be either expressions of x and y (e.g. "3xy + 2y"), or the letter "N"

6. For each of the following vector fields $\boldsymbol{F}$, decide whether it is conservative or not by computing curl $\boldsymbol{F}$. Type in a potential function f (that is, $\nabla f = \mathbf{F}$). If it is not conservative, type N.

A. $\mathbf{F}(x,y) = (-14x + 5y)\,\mathbf{i} + (5x + 4y)\,\mathbf{j}$

$f(x,y) =$ _____

B. $\mathbf{F}(x,y) = -7y\mathbf{i} - 6x\mathbf{j}$

$f(x,y) =$ _____

C. $\mathbf{F}(x,y,z) = -7x\mathbf{i} - 6y\mathbf{j} + \mathbf{k}$

$f(x,y,z) =$ _____

D. $\mathbf{F}(x,y) = (-7\sin y)\,\mathbf{i} + (10y - 7x\cos y)\,\mathbf{j}$

$f(x,y) =$ _____

E. $\mathbf{F}(x,y,z) = -7x^2\mathbf{i} + 5y^2\mathbf{j} + 2z^2\mathbf{k}$

$f(x,y,z) =$ _____

Note: Your answers should be either expressions of x, y and z (e.g. "3xy + 2yz"), or the letter "N"

7. Match the following vector fields with the verbal descriptions of the level curves or level surfaces to which they are perpendicular by putting the letter of the verbal description to the left of the number of the vector field.

(a) $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$

(b) $\mathbf{F} = 2x\mathbf{i} + y\mathbf{j}$

(c) $\mathbf{F} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$

(d) $\mathbf{F} = y\mathbf{i} + x\mathbf{j}$

(e) $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$

(f) $\mathbf{F} = x\mathbf{i} + y\mathbf{j} - z\mathbf{k}$

(g) $\mathbf{F} = 2x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$

(h) $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$

(i) $\mathbf{F} = x\mathbf{i} - y\mathbf{j}$

(j) $\mathbf{F} = x\mathbf{i} + y\mathbf{j} - \mathbf{k}$

(k) $\mathbf{F} = 2\mathbf{i} + \mathbf{j}$

A. hyperboloids

B. ellipsoids

C. spheres

D. lines

E. circles

F. paraboloids

G. hyperbolas

H. planes

I. ellipses

8. Assume that $x, y \geq 0$ for all of the vector fields in this question. Select an answer for each question and explain your reasoning.

(a) Let $\vec{F}_1 = y\vec{i}$.

(a) The vector field $\vec{F}_1$ is

- parallel to the x-axis
- parallel to the y-axis
- neither

(b) As x increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(c) As y increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(b) Let $\vec{F}_2 = \langle y, 1 \rangle$.

(a) The vector field $\vec{F}_2$ is

- parallel to the x-axis
- parallel to the y-axis
- neither

(b) As x increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(c) As y increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(c) Let $\vec{F}_3 = (x + e^{1-y})\vec{j}$.

(a) The vector field $\vec{F}_3$ is

- parallel to the x-axis
- parallel to the y-axis
- neither

(b) As x increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(c) As y increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(d) Let $\vec{F}_4 = \nabla(y^4 + e^{2x})$.

(a) The vector field $\vec{F}_4$ is

- parallel to the x-axis
- parallel to the y-axis
- neither

(b) As x increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

(c) As y increases,

- the length of the vector field increases
- the length of the vector field decreases
- neither

13.3 The Idea of a Line Integral

Motivating Questions

- What is an oriented curve and how can we represent one algebraically?
- What is the meaning of the line integral of a vector-valued function along a curve and how can we estimate if its value is positive, negative, or zero?
- What are important properties of the line integral of a vector-valued functions along a curve?

13.3.1 Introduction

Velocity versus Force. Movement by things like air or water will exert a force on objects, so wind velocity vector fields are related to a force field but the relationship between a moving fluid and the force a fluid exerts on an object in the fluid can be complicated. Specifically, issues like drag/friction and cross sectional area are a critical part of this relationship but are not something that we will focus on in our discussion. In the exploration below, we will look at a simplified version of this situation.

As we discussed in Section 13.2, vector fields are often used to represent forces such as gravity or electromagnetism, as well as the velocity of movement for things like wind or flowing water. We learned in Section 9.4 that the dot product of a force vector and a displacement vector tells us how much work the force did on the object as it moved from the start of its displacement vector to the end. However, this calculation assumes that the force is constant in the region of movement and that the object moves in a straight line along the displacement vector. The situation is more complicated than a dot product calculation when an object's movement is not in a straight line and when the force is not uniform throughout the area in which the object moves.

The preview activity uses cardinal directions to specify the direction of displacement vectors. These directions can be described by a Copyright 2013-2026 . The Copyright 2013-2026 given in Figure 13.3.1 is an example of a sixteen point rose. Note that directions like ESE are read as "east-southeast" half way between east and southeast.

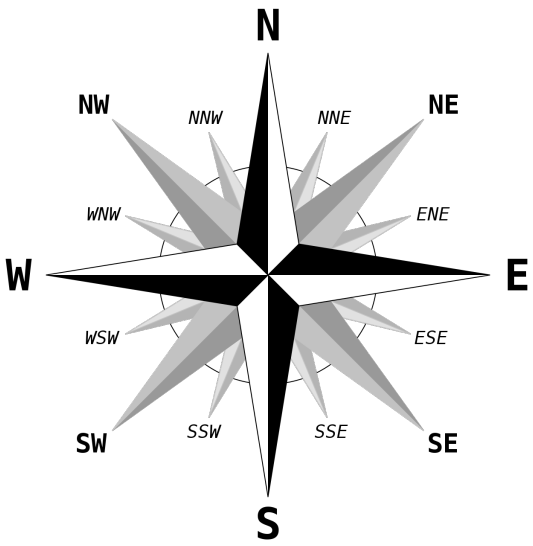


Figure 13.3.1 An example of a sixteen point Copyright 2013-2026 (from By Brosen~commonswiki - Own work, CC BY 2.5)

Preview Activity 13.3.1 Recall from Section 9.4 that the work done by a force $\vec{F}$ on an object that moves with displacement vector $\vec{v}$ is $\vec{F} \cdot \vec{v}$. In this Preview Activity, we consider the work done by wind on a helicopter at various stages of its journey.

(a) Our intrepid pilot flies for some time and finds that they are 30 km from where they started at a heading of 20° degrees east of due north. During this portion of the trip, the wind is exerting a force of 100 N on the helicopter in the due east direction. Find the work the wind has done on the helicopter during the flight.

(b) Our pilot sees a storm ahead and changes their direction. Some time later, the pilot determines that they are 25 km due north of where they previously checked their position. The wind is still exerting a force on the helicopter of 100 N in the due east direction. Find the work done by the wind on the helicopter during the second part of the flight.

(c) Find the helicopters's displacement from its original position after the first two parts of its flight and use that to find the work done by the wind on the helicopter during the first two parts of flight.

(d) How does your answer to part 13.3.1.c connect to the answers to part 13.3.1.a and part 13.3.1.b?

(e) In order to get further away from the storm, the pilot turns and flies 45° west of due north for 50 km. The storm the pilot was avoiding has caused the wind to change as well. For this portion of the flight, the wind is exerting a force on the helicopter of 125 N in the south direction. Find the work done by the wind on the helicopter during this part of the flight.

(f) Explain why you cannot take the total displacement of the three parts of the helicopter flight and calculate the total work done by the wind on the helicopter.

13.3.2 Orientations of Curves

Given our motivation for calculating the work that a force field does on an object as it moves through the field, it is natural to concern ourselves with *how* the object moves. In particular, in many circumstances it will be different if an object moves from the point $(0, 1)$ to the point $(4, 3)$ by first going up the y-axis to $(0, 3)$ and then moving horizontally to $(4, 3)$ (illustrated by C_1 in Figure 13.3.2) than if the object moves along the line segment from $(0, 1)$ directly to $(4, 3)$(illustrated by C_2 in Figure 13.3.2). Similarly, given a fixed force field, we would expect the work done to be different (in fact, opposite) if the object moves from $(4, 3)$ to $(0, 1)$ directly along a line segment (C_3 in Figure 13.3.2). We say that a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$ is **oriented** if we have specified the direction of travel along the curve. When a curve is given parametrically (including as a vector-valued function), our convention will be that the orientation follows from the smallest allowable value of the parameter to the largest.

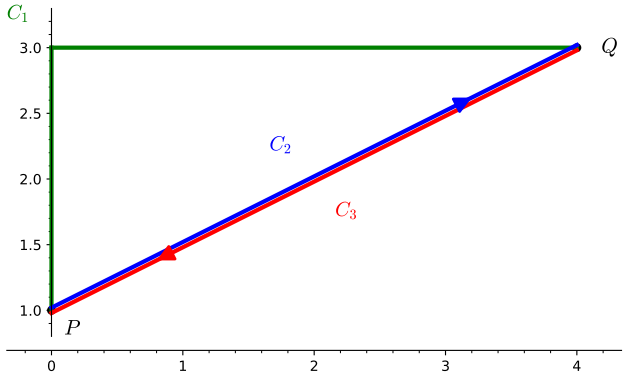


Figure 13.3.2 Curves C_1, C_2, and C_3 demonstrating different paths and orientations between P and Q

Activity 13.3.2 For each curve below, find a parametrization of the curve. Ensure that each curve's orientation matches the one specified.

(a) The line segment in $\mathbb{R}^3$ from $(0, 1, -2)$ to $(3, -1, 2)$.

(b) The line segment in $\mathbb{R}^3$ from $(3, -1, 2)$ to $(0, 1, -2)$.

(c) The circle of radius 3 (in $\mathbb{R}^2$) centered at the origin, beginning at the point $(0, -3)$ and proceeding clockwise around the circle.

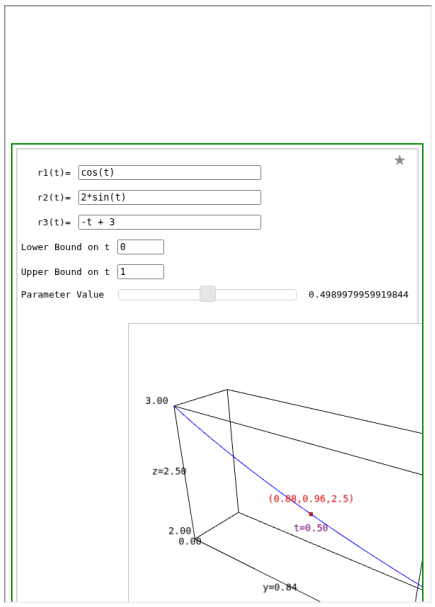
(d) In $\mathbb{R}^2$, the portion of the parabola $y^2 = x$ from the point $(4, 2)$ to the point $(1, -1)$.

In general, there are many ways to parametrize an oriented curve. With line segments, it is common to have the parameter range from 0 to 1, although there are sometimes good reasons to choose another method. For circles and ellipses, you may find it useful to interchange the placement of $\cos(t)$ and $\sin(t)$ to change the orientation, but then careful attention will need to be paid to the start and end points. The interactive graph below allows you to plot parametric curves. You should experiment with the graph below and try to make sense of how changing different elements affects the graph shown below. Remember that you can change the highlighted point using the slider. You should take time now to at least try the following:

1. Change the upper bound of the plot to 2π
2. Exchange the sine and cosine functions in r_1 and r_2 (but do not change

the coefficients)

3. Change the t in each of the component functions to $-t$
4. Change the $-t$ to e^t (you will need to use the exp(t) function in Sage)
5. Change the upper and lower bounds to get the same curve plotted as in the earlier parts



Standalone
In Context
Embed

Figure 13.3.3 This is a plot of a parametric curve of the form $\mathbf{r}(t) = \langle r_1(t), r_2(t), r_3(t) \rangle$. For two-dimensional curves, put 0 for $r_3(t)$.

13.3.3 Line Integrals

Preview Activity 13.3.1 showed how we can break up the work done along a path into a sum of work done on each piece. This will be a very helpful idea, especially if we consider the work done by a vector field that is not constant.

For example, let's consider how to measure the work done by $\vec{F}$, a vector field, along C, the curve shown below that goes from P to Q.

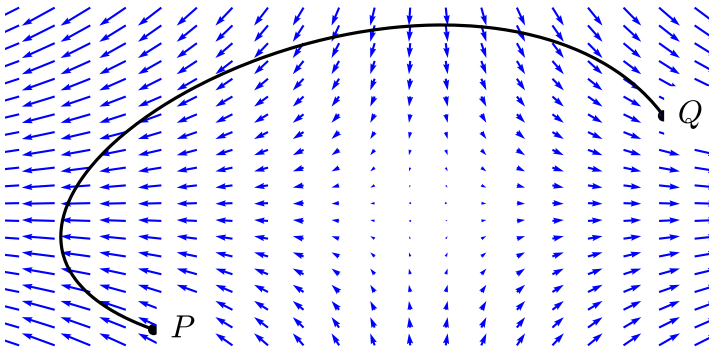


Figure 13.3.4 A curve C oriented from the point P to the point Q with the plot of a vector field $\vec{F}$

You can see that there will be parts of C such that the dot product of the direction of travel and the vector field will be positive and some parts where

the dot product is negative. We don't need to consider the output of the vector field except at the points on the curve. Thus, we will look at the following plot of the output of $\vec{F}$ at a collection of points on the curve C. Remember that when the vector field is plotted on the whole space, the lengths are rescaled to not be visually cluttered. In Figure 13.3.5 the actual output vectors are plotted for some points on C.

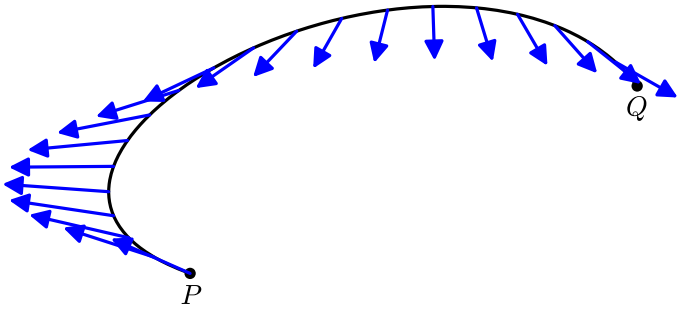


Figure 13.3.5 A curve C oriented from the point P to the point Q with the vector field, $\vec{F}$, plotted at points along the curve

Since the output of $\vec{F}$ is changing as you move along the path, we will use a classic calculus approach. For step 1 of our CCA, we will break our region up into smaller pieces to approximate the work done on each piece. Figure 13.3.6 shows how we can break C into n pieces, which we will call C_i. Note that C_i goes from the point $\mathbf{r}_{i-1}$ to $\mathbf{r}_i$.

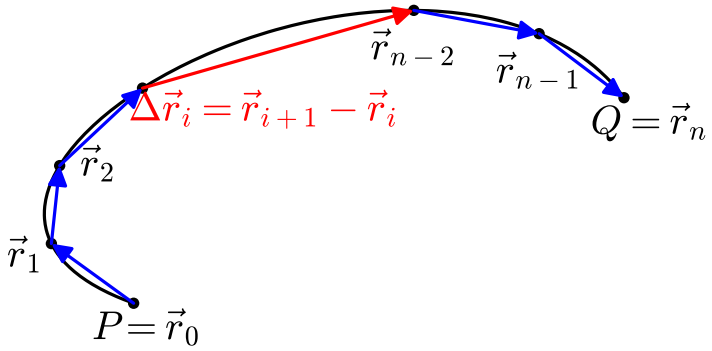


Figure 13.3.6 A curve C oriented from the point P to the point Q broken into n pieces

If we look at one of these smaller pieces (as show in Figure 13.3.7), we can see that vector field still changes at these points, but the output vectors are very similar. We can also see that the vector $\Delta\mathbf{r}_i = \mathbf{r}_i - \mathbf{r}_{i-1}$ is a good approximation of the curve piece C_i. Therefore we can approximate the work done by the vector field $\vec{F}$ on the piece C_i by $\vec{F}(\mathbf{r}_i) \cdot \Delta\mathbf{r}_i$. Remember that this is the same idea as in Section 9.4; Namely, the work done is the dot product of the force and the displacement, but this is done on many small pieces instead of the whole path.

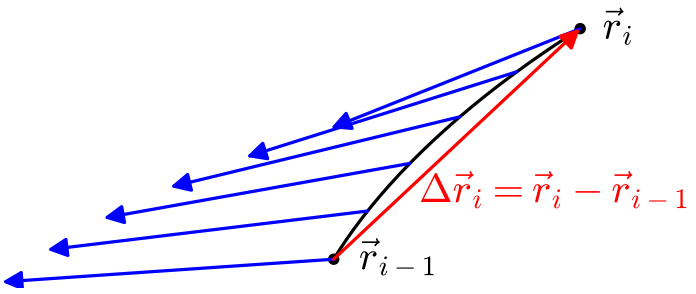


Figure 13.3.7 A plot of the piece C_i with the output of our vector field plotted at six poins on C_i

Our approximation with n pieces of C would be

$$\sum_{i=1}^{n} \vec{F}(\mathbf{r}_i) \cdot \Delta\mathbf{r}_i$$

which completes step 1 of the CCA.

Note that as we increase n, the number of pieces of C in our approximation, and make sure the length of each piece, C_i, goes to zero, the vector field $\vec{F}$ will be nearly constant on each piece. Additionally, the displacement vector, $\Delta\mathbf{r}_i$, will be a better approximation of C_i as n increases. This all combines to ensure that as we increase n, the number of smaller segments, our approximation will improve even more quickly.

We can calculate the exact amount of work done by $\vec{F}$ over C by taking the limit of this approximation as the number of pieces increases, while ensuring the size of each piece goes to zero. This should be a familiar Riemann sum argument from earlier definitions involving integration.

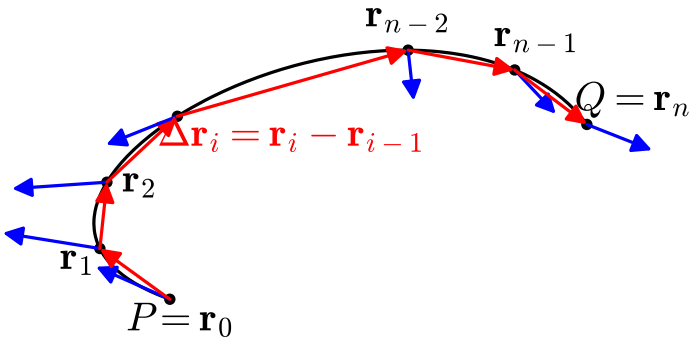


Figure 13.3.8 A curve C oriented from the point P to the point Q broken into n pieces with the output of $\vec{F}$ plotted at each intermediate point

Definition 13.3.9 Let C be an oriented curve and $\vec{F}$ a vector field defined in a region containing C. The **line integral** of $\vec{F}$ along C is

$$\int_C \vec{F} \cdot d\vec{r} = \lim_{\|\Delta\vec{r}_i\| \to 0} \sum_{i=1}^{n} \vec{F}(\vec{r}_i) \cdot \Delta\vec{r}_i,$$

provided the limit exists. ◆

Qualititatively, the line integral of a vector field measures how much of the vector field is in the same direction as the motion along the curve C. Physically, we can think of the line integral as measuring the work done by the vector field on an object moving along C.

We are able to say a bit here about conditions that guarantee the limit in Definition 13.3.9 exists. First, we require that C is a relatively "nice" curve. We say that a curve is **smooth** provided that it can be parameterized with functions that are infinitely differentiable. We do not require that C be smooth, but only that it be **piecewise smooth**, which means that C can be separated into a finite number of parts which are individually smooth. The other requirement is that $\vec{F}$ is a continuous vector field, by which we mean that each component function of $\vec{F}$ is continuous as a function of 2 or 3 variables (for a large enough region around our curve C).

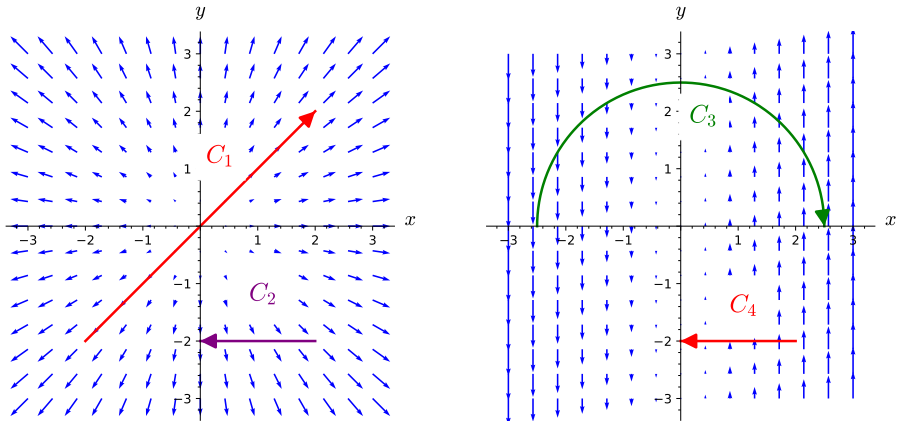
Because the dot products in the definition of the line integral $\int_C \vec{F} \cdot d\vec{r}$ can each be viewed as the work done by $\vec{F}$ as an object moves along the (very small) vector $\Delta\vec{r}_i$, the line integral gives the total work done by the vector field on an object that moves along C (in the direction of its orientation).

Alternative notation for line integrals. Our notation for line integrals is one of several common notations. This notation's strength is that it emphasizes the role of a vector field and dot product. Another common notation for the line integral of a vector field $\langle P, Q, R\rangle$ along a curve C is $\int_C P\,dx + Q\,dy + R\,dz$. This notation is called the differential form of line integrals and is commonly used in physics and engineering. We will limit ourselves to the $\int_C \vec{F} \cdot d\vec{r}$ notation in the body of the text, but some exercises may use alternative notation. You can find more info at Subsection 13.4.3.

If we are trying to determine how much a wind current helps or hinders an aircraft flying along a path determined by the curve, then calculating the dot product $\vec{F}(\vec{r}_i) \cdot \Delta\vec{r}_i$ makes sense for the local amount of help or hindrance. Note here that our vector $\Delta\vec{r}_i$ is the displacement from $\vec{r}_{i-1}$ to $\vec{r}_i$. If the displacement vector, $\Delta\vec{r}_i$, and the force field vector, $\vec{F}(\vec{r}_i)$, point in directions with an acute angle between them, the dot product will be positive. [1] On the other hand, if the angle between them is obtuse, the dot product will be negative. In this case, we also note that the force field is hindering the aircraft's progress. This interpretation carries through with our limit argument above. Specifically, the line integral over a curve will be positive if more of the vector field is in the direction of travel than against the direction of travel. You will need to consider the value of the dot product and not just the sign when assessing whether a line integral will be positive/neagtive/zero.

Activity 13.3.3 Shown in Figure 13.3.10 are two vector fields, $\vec{F}$ and $\vec{G}$ and four oriented curves, as labeled in the plots. For each of the line integrals below, determine if its value should be positive, negative, or zero. Do this by thinking about if the vector field is helping or hindering a particle moving along the oriented curve, rather than by doing calculations.

[1] We are abusing notation here a tiny bit, since technically the domain of $\vec{F}$ is points in $\mathbb{R}^2$ or $\mathbb{R}^3$, and $\vec{r}_i$ is a vector. By $\vec{F}(\vec{r})$, we mean $\vec{F}(r_1, r_2)$, where $\vec{r} = \langle r_1, r_2\rangle$.



(a) A plot of $\vec{F}$ with paths C_1 and C_2 **(b)** A plot of $\vec{G}$ with paths C_3 and C_4

Figure 13.3.10 Vector fields and oriented curves

(a) $\displaystyle\int_{C_1} \vec{F} \cdot d\vec{r}$

(b) $\displaystyle\int_{C_2} \vec{F} \cdot d\vec{r}$

(c) $\displaystyle\int_{C_3} \vec{G} \cdot d\vec{r}$

(d) $\displaystyle\int_{C_4} \vec{G} \cdot d\vec{r}$

The next several sections will be devoted to determining ways to efficiently calculate line integrals. As with the limits in the definition of every other type of integral we've studied so far, the limit in the definition of the line integral is is cumbersome to work with in most cases. However, in the case where the oriented curve C is composed of horizontal and vertical line segments, we can make a rather quick reduction to a single-variable integral, as the following example shows.

Example 13.3.11 Consider the constant vector field $\vec{F}(x, y) = \langle 2, 1 \rangle$. Let C be the curve that follows the horizontal line segment from $(1, 1)$ to $(4, 1)$ and then continues down the vertical line segment to $(4, -2)$. Figure 13.3.12 shows $\vec{F}$ and C, including the orientation. We will calculate $\displaystyle\int_C \vec{F} \cdot d\vec{r}$.

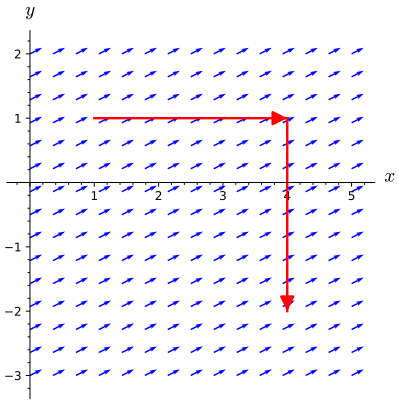


Figure 13.3.12 An oriented curve from $(1, 1)$ to $(4, -2)$ in a vector field $\vec{F}$.

To calculate $\int_C \vec{F} \cdot d\vec{r}$, we start by working with the horizontal line segment. Along that part of C, notice that $d\vec{r} \approx \Delta\vec{r} = \Delta x\hat{\imath}$. Thus, the Riemann sum that calculates the line integral along this portion of C consists of terms of the form $\langle 2, 1\rangle \cdot (\Delta x\hat{\imath}) = 2\Delta x$. Along this part of C, x ranges from 1 to 4, and thus we can turn the Riemann sum here into the definite integral $\int_1^4 2\,dx = 6$. Since the vectors are generally pointing in a direction that agrees with the orientation of C, we are not surprised to have a positive value here.

Now we turn our attention to the vertical portion of C. Here $d\vec{r} \approx \Delta\vec{r} = \Delta y\hat{\jmath}$, which means that $\vec{F} \cdot d\vec{r} \approx 1\Delta y$. Hence, our Riemann sum can be calculated by the definite integral $\int_1^{-2} 1\,dy = -3$. Notice that the limits of integration here were set up to match the orientation of C. Also, the negative value should not be unexpected, since C is oriented in a direction for which the vectors of $\vec{F}$ point in a direction that would hinder motion along C.

Combining these two calculations, we find that $\int_C \vec{F} \cdot d\vec{r} = 6 - 3 = 3$. ▲

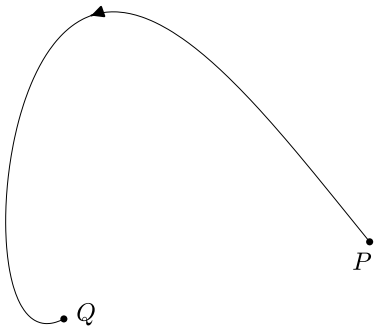
13.3.4 Properties of Line Integrals

In Example 13.3.11, we implicitly made use of the idea that if C can be broken up into two curves C_1 and C_2 such that the terminal point of C_1 is the initial point of C_2, then the line integral of $\vec{F}$ along C is the sum of the line integrals of $\vec{F}$ along C_1 and along C_2. This is a generalization of the property for definite integrals that tells us if $c \in [a, b]$, then

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

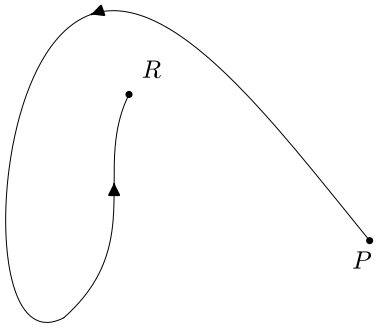
We next describe some common ways of breaking line integrals into pieces.

Convention 13.3.13 Describing Paths in Line Integrals. Before stating some useful properties of line integrals, we will establish some convenient notation. If C_1 and C_2 are oriented curves, with C_1 from a point P to a point Q and C_2 from Q to a point R, we denote by $C_1 + C_2$ the oriented curve from P to R that follows C_1 to Q and then continues along C_2 to R. Also, if C_1 is an oriented curve, $-C_1$ denotes the same curve but with the opposite orientation. The list below summarizes some other properties of line integrals, each of which has a familiar analog amongst the properties of definite integrals.

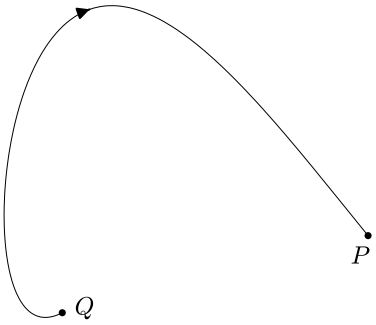


(a) C_1

(b) C_2



(c) $C_1 + C_2$



(d) $-C_1$

Figure 13.3.14 Plots of two oriented curves as well has the combination of them and one with the opposite orientation

Properties of Line Integrals.

For a constant scalar k, vector fields $\vec{F}$ and $\vec{G}$, and oriented curves C, C_1, and C_2, the following properties hold:

a. $\displaystyle\int_C (k\vec{F}) \cdot d\vec{r} = k \int_C \vec{F} \cdot d\vec{r}$

b. $\displaystyle\int_C (\vec{F} + \vec{G}) \cdot d\vec{r} = \int_C \vec{F} \cdot d\vec{r} + \int_C \vec{G} \cdot d\vec{r}$

c. $\displaystyle\int_{-C} \vec{F} \cdot d\vec{r} = -\int_C \vec{F} \cdot d\vec{r}$

d. $\displaystyle\int_{C_1+C_2} \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{C_2} \vec{F} \cdot d\vec{r}.$

Activity 13.3.4 Figure 13.3.15 shows a vector field $\vec{F}$ as well as six oriented curves, as labeled in the plot.

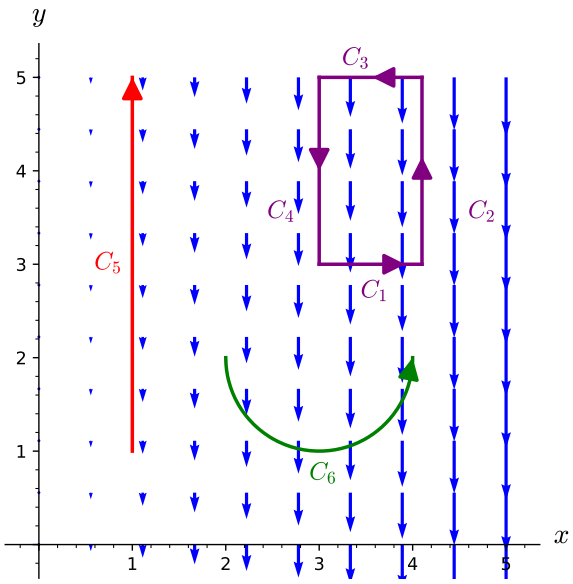


Figure 13.3.15 A vector field $\vec{F}$ and six oriented curves.

(a) Is $\int_{C_6} \vec{F} \cdot d\vec{r}$ positive, negative, or zero? Explain.

(b) Let $C = C_1 + C_2 + C_3 + C_4$. Determine if $\displaystyle\int_C \vec{F} \cdot d\vec{r}$ is positive, negative, or zero.

(c) Order the line integrals below from smallest to largest.

$$\int_{C_1} \vec{F} \cdot d\vec{r} \quad \int_{C_2} \vec{F} \cdot d\vec{r} \quad \int_{C_3} \vec{F} \cdot d\vec{r} \quad \int_{C_4} \vec{F} \cdot d\vec{r} \quad \int_{C_5} \vec{F} \cdot d\vec{r}$$

13.3.5 The Circulation of a Vector Field

If an oriented curve C ends at the same point where it started, we say that C is **closed**. The line integral of a vector field $\vec{F}$ along a closed curve C is called the **circulation** of $\vec{F}$ around C. To emphasize the fact that C is closed, we sometimes write $\oint_C \vec{F} \cdot d\vec{r}$ for $\int_C \vec{F} \cdot d\vec{r}$. Circulation serves as a measure of a vector field's tendency to rotate in a manner consistent with the orientation of the (closed) curve and is measured by looking at whether the vector field is working with or against the motion along the path.

Activity 13.3.5 Determine if the circulation of the vector field around each of the closed curves shown in Figure 13.3.16 is positive, negative, or zero.

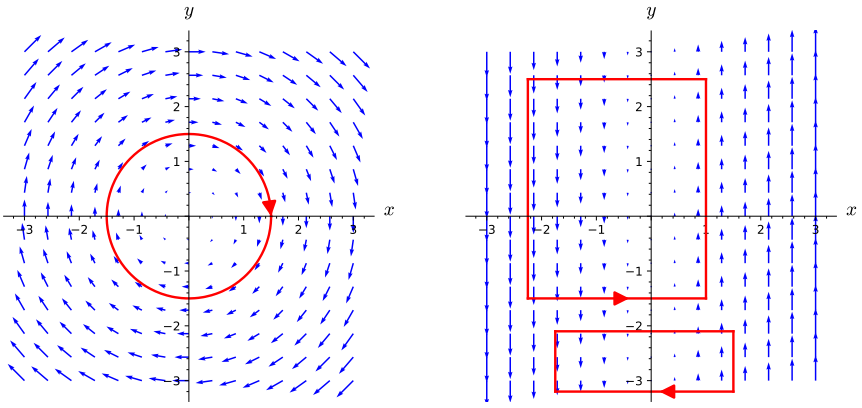


Figure 13.3.16 Vector fields and closed curves

Summary

- An oriented curve can be represented by a vector-valued function of one variable $\vec{r}(t)$ where we interpret the initial and terminal values of the domain of $\vec{r}$ as giving an orientation to the curve. A curve that ends at the same point where it started is said to be closed.
- A line integral (of a vector field) measures the extent to which the vector field points in a direction consistent with the orientation of the curve.
- Line integrals have many properties that are analogous to those of definite integrals of functions of a single variable.
- The line integral of a vector field along a closed curve is called the circulation of the vector field along the curve.

13.3.6 Exercises

1. Let C be the counter-clockwise planar circle with center at the origin and radius r > 0. Without computing them, determine for the following vector fields $\mathbf{F}$ whether the line integrals $\int_C \mathbf{F} \cdot d\mathbf{r}$ are positive, negative, or zero and type P, N, or Z as appropriate.

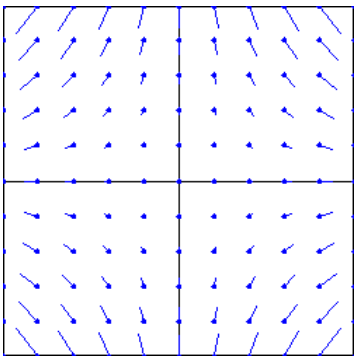
A. $\mathbf{F}$ = the radial vector field = $x\mathbf{i} + y\mathbf{j}$: _____

B. $\mathbf{F}$ = the circulating vector field = $-y\mathbf{i} + x\mathbf{j}$: _____

C. $\mathbf{F}$ = the circulating vector field = $y\mathbf{i} - x\mathbf{j}$: _____

D. $\mathbf{F}$ = the constant vector field = $\mathbf{i} + \mathbf{j}$: _____

2. Consider the vector field shown below.



*(**Note:** For the vector field, vectors are shown with a dot at the **tail** of the vector.)*

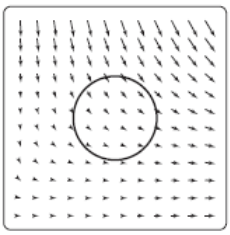
Let C_1 be the line from $(0,0)$ to $(-1,1)$, C_2 be the line from $(0,1)$ to $(-1,0)$, and C_3 be the semi-circle centered on the origin that goes from $(0,1)$ to $(0,-1)$ in a counterclockwise direction.

Is the line integral of the vector field along each of these curves positive, negative or zero?

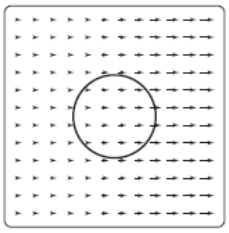
Along C_1, the integral is (☐ positive ☐ negative ☐ zero)
Along C_2, the integral is (☐ positive ☐ negative ☐ zero)
Along C_3, the integral is (☐ positive ☐ negative ☐ zero)

3. The three figures below show three vector fields. In each case, determine whether the line integral around the circle (oriented counterclockwise) is positive, negative, or zero.

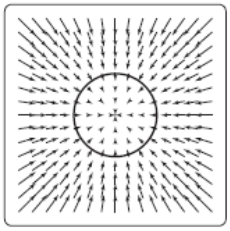
Note: Use "0" for zero, "P" for positive, and "N" for negative.



(A) __

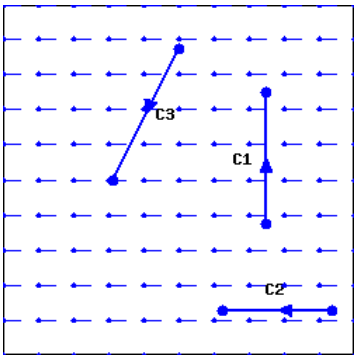


(B) __



(C) __

4. Consider the vector field $\vec{F}$ shown in the figure below together with the paths C_1, C_2, and C_3.

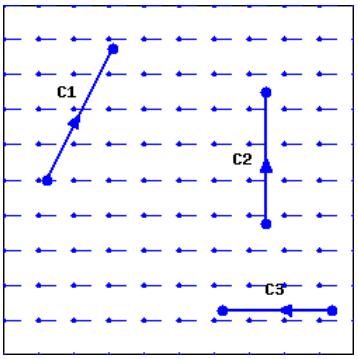


*(**Note:** For the vector field, vectors are shown with a dot at the **tail** of the vector.)*

Arrange the line integrals $\int_{C_1} \vec{F}\cdot d\vec{r}$, $\int_{C_2} \vec{F}\cdot d\vec{r}$ and $\int_{C_3} \vec{F}\cdot d\vec{r}$ in ascending order:

(☐ integral on C1 ☐ integral on C2 ☐ integral on C3) < (☐ integral on C1 ☐ integral on C2 ☐ integral on C3) < (☐ integral on C1 ☐ integral on C2 ☐ integral on C3) .

5. Consider the vector field $\vec{F}$ shown in the figure below together with the paths C_1, C_2, and C_3.

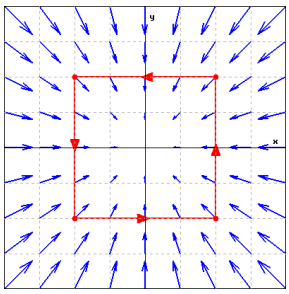


*(**Note:** For the vector field, vectors are shown with a dot at the **tail** of the vector.)*

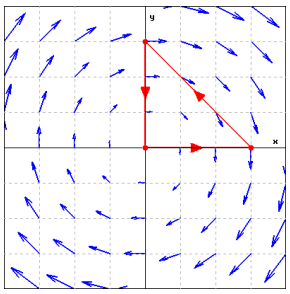
Arrange the line integrals $\int_{C_1} \vec{F}\cdot d\vec{r}$, $\int_{C_2} \vec{F}\cdot d\vec{r}$ and $\int_{C_3} \vec{F}\cdot d\vec{r}$ in ascending order:

(☐ integral on C1 ☐ integral on C2 ☐ integral on C3) < (☐ integral on C1 ☐ integral on C2 ☐ integral on C3) < (☐ integral on C1 ☐ integral on C2 ☐ integral on C3) .

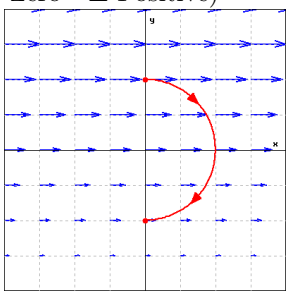
6. Determine whether the line integral of each vector field (in blue) along the oriented path (in red) is positive, negative, or zero.



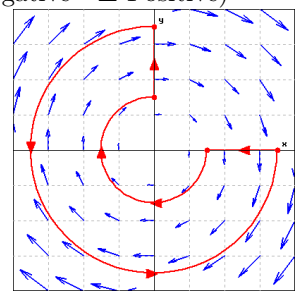
(☐ Negative ☐ Zero ☐ Positive)



(☐ Zero ☐ Negative ☐ Positive)



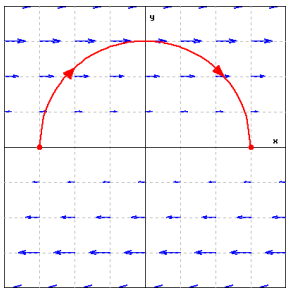
(☐ Positive ☐ Negative ☐ Zero)



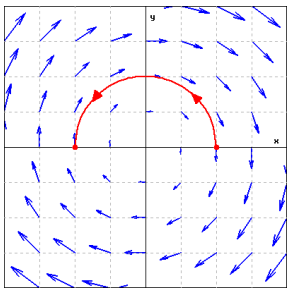
(☐ Positive ☐ Zero ☐ Negative)

(Click on a graph to enlarge it)

7. Determine whether the line integral of each vector field (in blue) along the semicircular, oriented path (in red) is positive, negative, or zero.

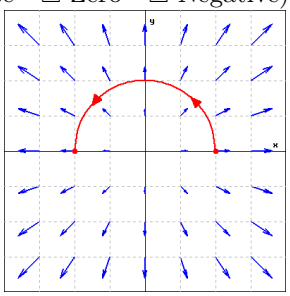


(☐ Positive ☐ Choose ☐ Zero ☐ Negative)

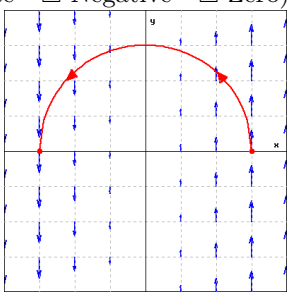


(☐ Positive ☐ Choose ☐ Negative ☐ Zero)

(☐



(☐ Positive ☐ Choose ☐ Negative ☐ Zero)



(☐ Zero ☐ Negative ☐ Choose ☐ Positive)

(☐

(Click on a graph to enlarge it)

8. If C is the z-axis from the origin to the point $(0, 0, 2)$, then $\int_C (-2\vec{j} - 2\vec{k}) \cdot d\vec{r} =$ _____

9. Sketch the vector field $\vec{F}(x, y) = -2\boldsymbol{i}$ and calculate the line integral of $\vec{F}$ along the line segment from $(-3, -2)$ to $(-2, -5)$.

10. Sketch the vector field $\vec{F}(x,y) = x\vec{i} + y\vec{j}$ and calculate the line integral of $\vec{F}$ along the line segment from $(7,3)$ to $(7,7)$.

11. Let C be the path given below from P to Q with pieces C_1, C_2, and C_3 as labeled. Let $\vec{F}$ be a vector field such that $\int_C \vec{F} \cdot d\vec{r} = 9$, $\int_{C_1} \vec{F} \cdot d\vec{r} = 6$,and $\int_{C_3} \vec{F} \cdot d\vec{r} = 7$.

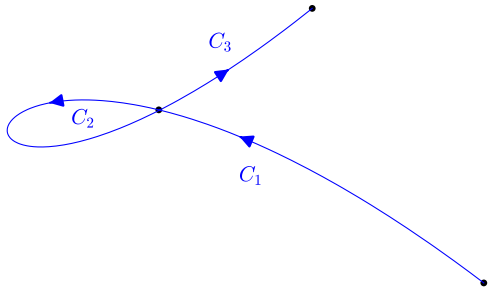


Figure 13.3.17 An oriented path broken into three pieces

Find the following:

(a) $\displaystyle\int_{-C_3} \vec{\vec{F}} \cdot d\vec{\vec{r}}$

(b) $\displaystyle\int_{C_2} \vec{\vec{F}} \cdot d\vec{\vec{r}}$

(c) $\displaystyle\int_{-C_1-C_3} \vec{\vec{F}} \cdot d\vec{\vec{r}}$

13.4 Using Parameterizations to Calculate Line Integrals

Motivating Questions

- How can we use a parametrization of an oriented curve C to calculate $\int_C \vec{F} \cdot d\vec{r}$?
- How does the parametrization chosen for an oriented curve C affect the value of the line integral $\int_C \vec{F} \cdot d\vec{r}$?
- Under what conditions will changing the path taken from P to Q change the the value of the line integral?

13.4.1 Introduction

We begin this section by taking a look at how to calculate a line integral of a vector field along different line segments. We will use this calculation as inspiration to see how treating oriented curves as vector-valued functions will allow us to quickly turn a line integral of a vector field into a single variable integral.

Preview Activity 13.4.1 Let $\vec{F} = \langle xy, y^2 \rangle$, let C_1 be the line segment from $(1,1)$ to $(4,1)$, let C_2 be the line segment from $(4,1)$ to $(4,3)$, and let C_3 be the line segment from $(1,1)$ to $(4,3)$. Also let $C = C_1 + C_2$. This vector field and the curves are shown in Figure 13.4.1.

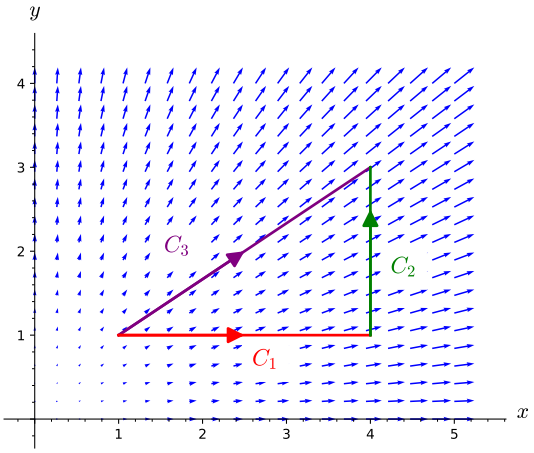


Figure 13.4.1 A vector field $\vec{F}$ and three oriented curves

(a) Every point along C_1 has $y = 1$. Therefore, along C_1, the vector field $\vec{F}$ can be viewed purely as a function of x. In particular, along C_1, we have $\vec{F}(x, 1) = \langle x, 1\rangle$. Since every point along C_2 has the same x-value, write $\vec{F}$ as a function of y only (for the points on C_2).

(b) Recall that $d\vec{r} \approx \Delta\vec{r}$, and along C_1, we have that $\Delta\vec{r} = \Delta x\hat{\imath} \approx dx\hat{\imath}$. Thus, $d\vec{r} = \langle dx, 0\rangle$. We know that along C_1, $\vec{F} = \langle x, 1\rangle$.

(i) Write $\vec{F} \cdot d\vec{r}$ along C_1 without using a dot product.

(ii) What interval of x-values describes C_1?

(iii) Write $\int_{C_1} \vec{F} \cdot d\vec{r}$ as an integral of the form $\int_a^b f(x)\, dx$ and evaluate the integral.

(c) Use an analogous approach to write $\int_{C_2} \vec{F} \cdot d\vec{r}$ as an integral of the form $\int_c^d g(y)\, dy$ and evaluate the integral.

(d) Use the previous parts and a property of line integrals to calculate $\int_C \vec{F} \cdot d\vec{r}$ without having to evaluate any additional integrals.

13.4.2 Parameterizations in the Definition of $\int_C \vec{F} \cdot d\vec{r}$

In Preview Activity 13.4.1, you saw how line integrals along vertical and horizontal line segments can be done as integrals of a single variable. Before moving on to the general case, let us consider in the next example how we might tackle $\int_{C_3} \vec{F} \cdot d\vec{r}$ in Figure 13.4.1.

Example 13.4.2 Since C_3 is from $(1, 1)$ to $(4, 3)$, we can determine that the line segment has slope $2/3$, so we can write an equation for the line as $y - 1 = (2/3)(x - 1)$ or $y = \frac{2}{3}x + \frac{1}{3}$. Thus, along the curve C_3, we can write

$$\vec{F} = \left\langle \frac{2}{3}x^2 + \frac{1}{3}x, (\frac{2}{3}x + \frac{1}{3})^2 \right\rangle.$$

Thinking of the slope of C_3 as $\Delta y/\Delta x$, we can write $\Delta y/\Delta x = 2/3$, which can be rearranged to $\Delta y = \frac{2}{3}\Delta x$. We may view $\Delta\vec{r}$ as $\langle \Delta x, \Delta y\rangle$. Since $\Delta x \approx dx$ and $\Delta y \approx dy$, we use the fact that $\Delta y = \frac{2}{3}\Delta x$ to write $d\vec{r} = \langle dx, \frac{2}{3}\, dx\rangle = \left\langle 1, \frac{2}{3}\right\rangle dx$. Along C_3, we have $1 \le x \le 4$, so having rewritten $\vec{F}$ in terms of x and $\vec{r}$ in

terms of dx, we can now write

$$\int_{C_3} \vec{F} \cdot d\vec{r} = \int_1^4 \left\langle \frac{2}{3}x^2 + \frac{1}{3}x, (\frac{2}{3}x + \frac{1}{3})^2 \right\rangle \cdot \left\langle 1, \frac{2}{3} \right\rangle dx$$
$$= \int_1^4 \frac{2}{3}x^2 + \frac{1}{3}x + \frac{2}{3}\left(\frac{2}{3}x + \frac{1}{3}\right)^2 dx = \frac{151}{6}.$$

Notice that this result is different than what you obtained for $\int_C \vec{F} \cdot d\vec{r}$ in Preview Activity 13.4.1, even though C and C_3 both start at $(1, 1)$ and end at $(4, 3)$. ▲

A recurring theme in this chapter will be the consideration of whether or not a vector field is a gradient vector field. Before moving on to generalize Example 13.4.2 to curves that are not line segments, it is worth examining this question for the vector field we have been investigating.

Activity 13.4.2 Is $\vec{F} = \langle xy, y^2 \rangle$ a gradient vector field? Why or why not?

Hint. What would Clairaut's Theorem tell you about a potential function f such that $\vec{F} = \nabla f$?

Preview Activity 13.4.1 and Example 13.4.2 have shown us that it is possible to evaluate line integrals without needing to work with Riemann sums directly. However, the approaches taken there seem rather cumbersome to use for oriented curves that are not line segments. It was not critical that the paths in Figure 13.4.1 were straight lines, but rather that the paths had a description where both x and y could be expressed in terms of one variable. Fortunately, parameterizing the oriented curve along which a line integral is calculated provides exactly the tool we are looking for. Namely a parameterization gives a way to express the points (x, y, z) of the oriented curve in terms of a single variable (the parameter). Note that given a parameterization, we can also translate the vector field component equations into functions of the parameter as well.

Example 13.4.3 In this example, we will look at the particulars of applying a parameterization of C (given by $\vec{r}(t) = \langle f(t), g(t) \rangle$ with $t \in [a, b]$) to Definition 13.3.9. In particular, we will use this parameterization and a set of points along C that are equally spaced in terms of parameter values. To make this discussion a bit easier to read, we will work with a curve in two dimensions. However, all the ideas here generalize directly to curves in $\mathbb{R}^3$.

Suppose that C is an oriented curve traced out by the vector-valued function $\vec{r}(t)$ for $a \leq t \leq b$, and let $\vec{F}$ be a continuous vector field defined on a region containing C. We can divide the interval $[a, b]$ into n sub-intervals, each of length $\Delta t = (b - a)/n$, by letting $t_i = a + i\Delta t$ for $i = 0, 1, \ldots, n$. Dividing the interval $[a, b]$ into n pieces using t_i then can be used to break C up into n pieces by letting C_i be the part of the curve from $\vec{r}(t_i)$ to $\vec{r}(t_{i-1})$ for $i = 1, \ldots, n$.

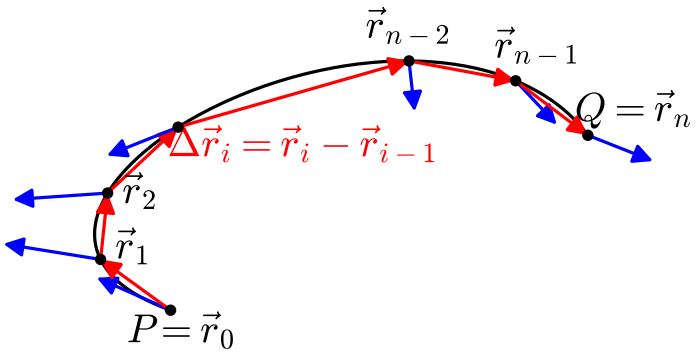


Figure 13.4.4 A curve C oriented from the point P to the point Q broken into n pieces evenly spaced in terms of the parameter t

We can approximate the path C_i with the displacement vector $\Delta\vec{r}_i = \vec{r}(t_i) - \vec{r}(t_{i-1})$ as we did in Figure 13.3.6. Now notice that

$$\Delta\vec{r}_i = \vec{r}(t_i) - \vec{r}(t_{i-1}) = \vec{r}(t_i) - \vec{r}(t_i - \Delta t) = \left(\frac{\vec{r}(t_i) - \vec{r}(t_i - \Delta t)}{\Delta t}\right)\Delta t.$$

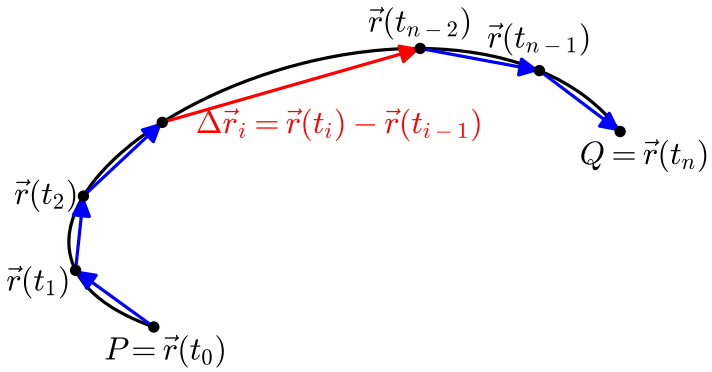


Figure 13.4.5 A curve C oriented from the point P to the point Q broken into n pieces evenly spaced in terms of the parameter t with displacement vectors Δr_i

Our vector field for $\int_C \vec{F} \cdot d\vec{r}$ can also be transformed into a function of t by substituting the x and y inputs with the corresponding components of the parameterization. If $\vec{F} = \langle F_1(x,y), F_2(x,y)\rangle$, then the parameterization given by $\vec{r}(t) = \langle f(t), g(t)\rangle$ means we can rewrite $\vec{F}$ as a function of t:

$$\vec{F}(t) = \vec{F}(r(t)) = \langle F_1(f(t),g(t)), F_2(f(t),g(t))\rangle$$

We can now substitute these pieces into Definition 13.3.9 to obtain the following

$$\begin{aligned}\int_C \vec{F} \cdot d\vec{r} &= \lim_{\|\Delta\vec{r}_i\|\to 0} \sum_{i=1}^{n} \vec{F}(\vec{r}_i) \cdot \Delta\vec{r}_i \\ &= \lim_{\Delta t \to 0} \sum_{i=1}^{n} \vec{F}(\vec{r}(t_i)) \cdot \left(\frac{\vec{r}(t_i) - \vec{r}(t_i - \Delta t)}{\Delta t}\right)\Delta t\end{aligned}$$

When evaluating the limit as $\Delta t \to 0$, the expression in the parentheses will be $\vec{r}'(t_i)$. We then have a Riemann sum that changes the evaluation of a line

integral of a vector field along an oriented curve to a definite integral of a function of one variable. In particular,

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\, dt$$

Note that after evaluating the dot product, $\vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)$ is (scalar) function of t. ▲

We now state the general form of the preceding example as a theorem that will allow us to evaluate line integrals of vector fields in many contexts.

Theorem 13.4.6 *Let C be a smooth, oriented curve traced out by the vector-valued function $\vec{r}(t)$ for $a \le t \le b$. If $\vec{F}$ is a continuous vector field defined near C, then*

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\, dt.$$

To illustrate how useful Theorem 13.4.6 is for evaluating line integrals, consider the following example.

Example 13.4.7 Let $\vec{F}(x, y) = x\hat{\imath} + y^2\hat{\jmath}$ and let C be the quarter of the circle of radius 3 from $(0, 3)$ to $(3, 0)$. This vector field and curve are shown in Figure 13.4.8. By properties of line integrals, we know that $\int_C \vec{F} \cdot d\vec{r} = -\int_{-C} \vec{F} \cdot d\vec{r}$, and we will use this property since $-C$ is the usual clockwise orientation of a circle, meaning we can parametrize $-C$ by $\vec{r}(t) = \langle 3\cos(t), 3\sin(t)\rangle$ for $0 \le t \le \pi/2$.

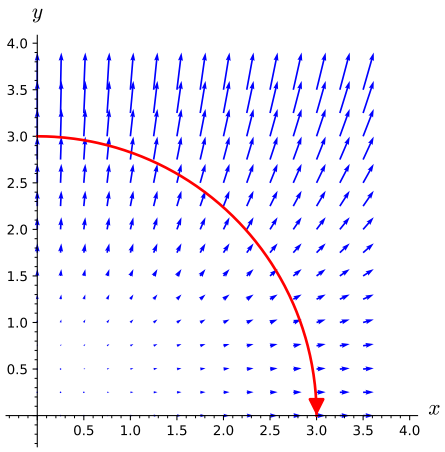


Figure 13.4.8 The vector field $\vec{F} = x\hat{\imath} + y^2\hat{\jmath}$ and an oriented curve C

To evaluate $\int_{-C} \vec{F} \cdot d\vec{r}$ using this parametrization, we need to note that

$$\vec{F}(\vec{r}(t)) = \langle 3\cos(t), 9\sin^2(t)\rangle \qquad \text{and} \qquad \vec{r}'(t) = \langle -3\sin(t), 3\cos(t)\rangle$$

Thus, we have

$$\begin{aligned}\int_C \vec{F} \cdot d\vec{r} &= -\int_{-C} \vec{F} \cdot d\vec{r} \\ &= -\int_0^{\pi/2} \langle 3\cos(t), 9\sin^2(t)\rangle \cdot \langle -3\sin(t), 3\cos(t)\rangle\, dt\end{aligned}$$

$$
\begin{aligned}
&= -\int_0^{\pi/2} \left(-9\sin(t)\cos(t) + 27\sin^2(t)\cos(t)\right)\, dt \\
&= -\int_0^1 (-9u + 27u^2)\, du = -\left[-\frac{9}{2}u^2 + 9u^3\right]_0^1 \\
&= -\left(-\frac{9}{2} + 9\right) = -\frac{9}{2}.
\end{aligned}
$$

Note that we have used the substitution $u = \sin(t)$ in evaluating the definite integral here. ▲

As Example 13.4.7 shows, Theorem 13.4.6 allows us to reduce the problem of calculating a line integral of a vector-valued function along an oriented curve to one of finding a suitable parametrization for the curve. Once we have such a parametrization, evaluating the line integral becomes evaluating a single-variable integral, which is something you have done many times before. This example also illustrates that using the properties of line integrals can allow us to use a more "natural" parametrization. You may find it interesting to use the parametrization $\langle 3\sin(t), 3\cos(t)\rangle$ for $0 \le t \le \pi/2$ to evaluate the line integral. Do you get the same result?

Activity 13.4.3

(a) Find the work done by the vector field $\vec{F}(x, y, z) = 6x^2z\hat{\imath} + 3y^2\hat{\jmath} + x\hat{k}$ on a particle that moves from the point $(3, 0, 0)$ to the point $(3, 0, 6\pi)$ along the helix given by $\vec{r}(t) = \langle 3\cos(t), 3\sin(t), t\rangle$.

(b) Let $\vec{F}(x, y) = \langle 0, x\rangle$. Let C be the closed curve consisting of the top half of the circle of radius 2 centered at the origin and the portion of the x-axis from $(2, 0)$ to $(-2, 0)$, oriented clockwise. Find the circulation of $\vec{F}$ around C.

Activity 13.4.4 will have you look at line integrals of the same vector field over several different types of curves. This will be an important, recurring theme as we study a variety of different integrals and vector fields in this chapter. In particular, we will use this approach of varying the region of integration for a fixed function several times in later activities.

Activity 13.4.4 Let $\vec{F}(x, y) = \langle y^2, 2xy + 3\rangle$.

(a) Let C_1 be the portion of the graph of $y = 2x^3 + 3x^2 - 12x - 15$ from $(-2, 5)$ to $(3, 30)$. Calculate $\int_{C_1} \vec{F} \cdot d\vec{r}$.

(b) Let C_2 be the line segment from $(-2, 5)$ to $(3, 30)$. Calculate $\int_{C_2} \vec{F} \cdot d\vec{r}$.

(c) Let C_3 be the circle of radius 3 centered at the origin, oriented counter-clockwise. Calculate $\oint_{C_3} \vec{F} \cdot d\vec{r}$.

(d) To connect the previous parts of this activity, use a graphing utility to plot the curves C_1 and C_2 on the same axes.

(i) What type of curve is $C_1 - C_2$?

(ii) What is the value of $\oint_{C_1 - C_2} \vec{F} \cdot d\vec{r}$?

(iii) What does your answer to part c allow you to say about the value of the line integral of $\vec{F}$ along the top half of C_3 compared to the line integral of $\vec{F}$ from $(3, 0)$ to $(-3, 0)$ along the bottom half of the circle of radius 3 centered at the origin?

13.4.3 Alternative Notation for Line Integrals

In contexts where the fact that the quantity we are measuring via a line integral is best measured via a dot product (such as calculating work), the notation we have used thus far for line integrals is fairly common. However, sometimes the vector field is such that the units on x, y, and z are not distances. In this case, a dot product may not have quite the same physical meaning, and an alternative notation using differentials can be common. Specifically, if $\vec{F}(x,y,z) = F_1(x,y,z)\hat{\imath} + F_2(x,y,z)\hat{\jmath} + F_3(x,y,z)\hat{k}$, then

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \langle F_1, F_2, F_3\rangle \cdot \langle dx, dy, dz\rangle = \int_C F_1\,dx + F_2\,dy + F_3\,dz.$$

A line integral in the form of $\int_C F_1\,dx + F_2\,dy + F_3\,dz$ is called the **differential form of a line integral**. (If $\vec{F}$ is a vector field in $\mathbb{R}^2$, the $F_3\,dz$ term is omitted.) For example, if $\vec{F}(x,y,z) = \langle x^2y, z^3, x\cos(z)\rangle$ and C is some oriented curve in $\mathbb{R}^3$, then

$$\int_C \vec{F} \cdot d\vec{r} = \int_C x^2y\,dx + z^3\,dy + x\cos(z)\,dz$$

It is important to recognize that the integral on the right-hand side is *still a line integral* and must be evaluated using techniques for evaluating line integrals. We cannot simply try to treat the line integral of the form $\int_C F_1\,dx + F_2\,dy + F_3\,dz$ as if it were a definite integral of a function of one variable. Because the notation $\int_C \vec{F} \cdot d\vec{r}$ provides a reminder that this is a line integral and *not* a definite integral of the types calculated earlier in your study of calculus, we will only use the vector notation for line integrals in the body of the text. However, some exercises may require the use of the differential form, and you may see the differential form used frequently in fields such as physics and engineering.

13.4.4 Independence of Parametrization for a Fixed Curve

Up to this point, we have chosen whatever parametrization of an oriented curve C was most convenient, and our argument for how we can use parameterizations to calculate line integrals did not depend on the specific choice of parametrization. However, it is not immediately obvious that different parameterizations don't result in different values of the line integral. Our next example explores this question.

Example 13.4.9 Let $\vec{F} = x\hat{\imath}$. We consider two different oriented curves from $(0,1)$ to $(3,3)$. The first oriented curve C travels horizontally to $(3,1)$ and then proceeds vertically to $(3,3)$. The second oriented curve C_3 is the line segment from $(0,1)$ to $(3,3)$. Notice that, as depicted in Figure 13.4.10, we can break C up into two oriented curves C_1 (the horizontal portion) and C_2 (the vertical portion) so that $C = C_1 + C_2$.

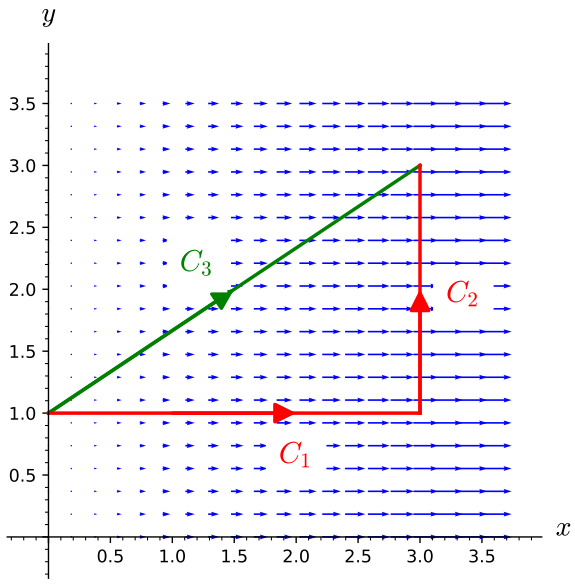


Figure 13.4.10 The vector field $\vec{F} = x\hat{\imath}$ and some oriented curves.

We first note that since $\vec{F}$ is orthogonal to C_2, $\int_{C_2} \vec{F} \cdot d\vec{r} = 0$; therefore $\int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r}$. We can parametrize C_1 as $t\hat{\imath} + \hat{\jmath}$ for $0 \le t \le 3$ (t is treated like the coordinate x), which leads to

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_0^3 \langle t, 0\rangle \cdot \langle 1, 0\rangle \, dt = \int_0^3 t \, dt = \frac{9}{2}.$$

Thus, $\int_C \vec{F} \cdot d\vec{r} = 9/2$.

Now we look at $\int_{C_3} \vec{F} \cdot d\vec{r}$, but we parametrize C_3 in a nonstandard way by letting $\vec{r}(t) = \langle 3\sin(t), 1 + 2\sin(t)\rangle$ for $0 \le t \le \frac{\pi}{2}$. (You should use a graphing utility to plot this parametrization to help convince yourself that it really does give C_3.) This gives $\vec{r}\,'(t) = \langle 3\cos(t), 2\cos(t)\rangle$, and

$$\begin{aligned}\int_{C_3} \vec{F} \cdot d\vec{r} &= \int_0^{\pi/2} \langle 3\sin(t), 0\rangle \cdot \langle 3\cos(t), 2\cos(t)\rangle \, dt \\ &= \int_0^{\pi/2} 9\sin(t)\cos(t) \, dt = \frac{9}{2}.\end{aligned}$$

In the next activity, you are asked to consider the more typical parametrization of C_3 and verify that using it gives the same value for the line integral.

It's also worth observing here that $\int_C \vec{F} \cdot d\vec{r} = \int_{C_3} \vec{F} \cdot d\vec{r}$, so at least two (very different) paths from $(0, 1)$ to $(3, 3)$ give the same value of the line integral here. The next section will further investigate when line integrals over different paths (with the same initial point and final point) will evaluate to the same value. ▲

As promised, the final activity of this section (Activity 13.4.5) asks you to look at another parametrization of the curve C_3 from the previous example. It also asks you to look at two different oriented curves between a pair of points, similarly to what you did in Activity 13.4.4.

Activity 13.4.5

(a) The typical parametrization of the line segment from $(0, 1)$ to $(3, 3)$ (the

oriented curve C_3 in Example 13.4.9) is $\vec{r}(t) = \langle 3t, 1+2t\rangle$ where $0 \le t \le 1$. Use this parametrization to calculate $\int_{C_3} \vec{F} \cdot d\vec{r}$ for the vector field $\vec{F} = x\hat{i}$ and compare your answer to the result of Example 13.4.9.

(b) Calculate $\int_C \langle (3xy + e^z), x^2, (4z + xe^z)\rangle \cdot d\vec{r}$ where C is the oriented curve consisting of the line segment from the origin to $(1, 1, 1)$ followed by the line segment from $(1, 1, 1)$ to $(0, 0, 2)$.

(c) Calculate $\int_{C'} \langle 3xy + e^z, x^2, 4z + xe^z\rangle \cdot d\vec{r}$ where C_3 is the line segment from $(0, 0, 0)$ to $(0, 0, 2)$.

(d) Is the vector field you considered in the previous two parts a gradient vector field? Why or why not? How does this compare to the vector field $\vec{F}$ of Activity 13.4.4?

Although we have not given a proof or even an intuitive argument, the phenomenon you observed in part a of Activity 13.4.5 is not particular to this curve or this vector field. The value of $\int_C \vec{F} \cdot d\vec{r}$ does not depend on the parametrization of C used to calculate the line integral when using Theorem 13.4.6.

Summary

- Line integrals of vector fields along oriented curves can be evaluated by parameterizing the curve in terms of t and then calculating the integral of $\vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)$ on the interval $[a, b]$.
- The parametrization chosen for an oriented curve C when calculating the line integral $\int_C \vec{F} \cdot d\vec{r}$ using the formula $\int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\, dt$ does not impact the value of the line integral.
- If C_1 and C_2 are different paths from P to Q, it is possible for $\int_{C_1} \vec{F} \cdot d\vec{r}$ to have a different value to $\int_{C_2} \vec{F} \cdot d\vec{r}$.

13.4.5 Exercises

1. If C is the line segment from $(2, 3)$ to, $(0, 0)$, find the value of the line integral:

$\int_C (4y^2\,\vec{i} + 4x\,\vec{j}) \cdot d\vec{r} =$ _____

2. Suppose $\vec{F}(x, y) = \langle e^x, e^y\rangle$ and C is the portion of the ellipse centered at the origin from the point $(0, 1)$ to the point $(6, 0)$ centered at the origin oriented clockwise.

(a) Find a vector parametric equation $\vec{r}(t)$ for the portion of the ellipse described above for $0 \le t \le \pi/2$.

$\vec{r}(t) =$ _____

(b) Using your parametrization in part (a), set up an integral for calculating the circulation of $\vec{F}$ around C.

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\, dt = \int_a^b _____\, dt$$

with limits of integration $a =$ _____ and $b =$ _____

(c) Find the circulation of $\vec{F}$ around C.

Circulation = _____

3. Evaluate the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$, where $\mathbf{F}(x, y, z) = x\mathbf{i} + 5y\mathbf{j} - 5z\mathbf{k}$ and C is given by the vector function $\mathbf{r}(t) = \langle \sin t, \cos t, t\rangle, \quad 0 \le t \le 3\pi/2$.

4. Evaluate the line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$, where $\mathbf{F}(x,y,z) = 3\sin x\mathbf{i} + \cos y\mathbf{j} - xz\mathbf{k}$ and C is given by the vector function $\mathbf{r}(t) = t^5\mathbf{i} - t^4\mathbf{j} + t^3\mathbf{k}$, $0 \le t \le 1$.

5. If C is the curve given by $\mathbf{r}(t) = (1+5\sin t)\,\mathbf{i} + \left(1+2\sin^2 t\right)\mathbf{j} + \left(1+5\sin^3 t\right)\mathbf{k}$, $0 \le t \le \frac{\pi}{2}$ and $\boldsymbol{F}$ is the radial vector field $\mathbf{F}(x,y,z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, compute the work done by $\boldsymbol{F}$ on a particle moving along C.

6. Find the work done by the force field $\mathbf{F}(x,y,z) = 2x\mathbf{i} + 2y\mathbf{j} + 1\mathbf{k}$ on a particle that moves along the helix $\mathbf{r}(t) = 7\cos(t)\mathbf{i} + 7\sin(t)\mathbf{j} + 7t\mathbf{k}, 0 \le t \le 2\pi$.

---

7. Let C be the curve which is the union of two line segments, the first going from (0, 0) to (-4, 2) and the second going from (-4, 2) to (-8, 0).

Compute the line integral $\displaystyle\int_C -4dy - 2dx$.

8. Calculate the line integral of the vector field $\vec{F} = -3\vec{\imath} - 4\vec{\jmath}$ along the line from the point $(1,0)$ to the point $(8,0)$.

The line integral = _____

9. Find $\int_C ((x^2+2y)\vec{i} + y^3\vec{j})\cdot d\vec{r}$ where C consists of the three line segments from $(2,0,0)$ to $(2,5,0)$ to $(0,5,0)$ to $(0,5,3)$.

$\int_C ((x^2+2y)\vec{i} + y^3\vec{j})\cdot d\vec{r}$ = _____

10. Find $\int_C \vec{F}\cdot d\vec{r}$ for $\vec{F} = 4y\vec{i} - (\sin y)\vec{j}$ on the curve counterclockwise around the unit circle C starting at the point $(1,0)$.

$\int_C \vec{F}\cdot d\vec{r}$ = _____

11. Find $\int_C \vec{F}\cdot d\vec{r}$ for $\vec{F} = -4y\,\vec{i} - 2x\,\vec{j} + 2z\,\vec{k}$ if C is the helix $x = \cos t, y = \sin t, z = t$, for $0 \le t \le 4\pi$.

$\int_C \vec{F}\cdot d\vec{r}$ = _____

12. Evaluate the line integral $\int_C y\,dx + x\,dy$ where C is the parameterized path $x = t^3$, $y = t^2$, $3 \le t \le 8$.

$\int_C y\,dx + x\,dy$ = _____

13. Evaluate the line integral $\int_C 6y\,dx + 4x\,dy$ where C is the straight line path from $(1,1)$ to $(3,4)$.

$\int_C 6y\,dx + 4x\,dy$ = _____

14. Curves C_1 and C_2 are parametrized as follows:

$$C_1 \text{ is } (x(t), y(t)) = (t, 0) \quad \text{for } -1 \le t \le 1$$

and

$$C_2 \text{ is } (x(t), y(t)) = (\cos t, \sin t) \quad \text{for } 0 \le t \le \pi.$$

Sketch, on a separate sheet of paper, the curves C_1 and C_2 with arrows showing their orientation.

Next, suppose that $\vec{F} = 2x\,\vec{i} + (5x+7y)\,\vec{j}$. Calculate $\int_C \vec{F}\cdot d\vec{r}$, where C is the curve given by $C = C_1 + C_2$.

$\int_C \vec{F}\cdot d\vec{r}$ = _____

15. Suppose C is the line segment from the point $(0,0)$ to the point $(3,24)$ and $\vec{F} = xy\,\vec{i} + 3x\,\vec{j}$.

(a) Is $\int_C \vec{F}\cdot d\vec{r}$ greater than, less than, or equal to zero? (☐ greater than zero ☐ less than zero ☐ equal to zero)

(Be sure you are able to give a geometric explanation for your answer.)

(b) A parameterization of C is $(x(t), y(t)) = (t, 8t)$ for $0 \le t \le 3$. Use this to compute $\int_C \vec{F}\cdot d\vec{r}$.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

(c) Suppose a particle leaves the point $(0, 0)$, moves along the line toward the point $(3, 24)$, stops before reaching it and backs up, stops again and reverses direction, then completes its journey to the endpoint. All travel takes place along the line segment joining the point $(0, 0)$ to the point $(3, 24)$. Call this path C'. How is $\int_{C'} \vec{F} \cdot d\vec{r}$ related to $\int_C \vec{F} \cdot d\vec{r}$?

$\int_{C'} \vec{F} \cdot d\vec{r}$ (☐ is greater than ☐ is less than ☐ is equal to) $\int_C \vec{F} \cdot d\vec{r}$

(Be sure you can explain why this is.)

(d) A parameterization for a path like C' is given by

$$(x(t), y(t)) = \left(\frac{1}{27} \left(16\,t^3 - 72\,t^2 + 99\,t\right), \frac{8}{27} \left(16\,t^3 - 72\,t^2 + 99\,t\right)\right).$$

For what value of t is this parameterization

at $(0, 0)$? $t =$ _____

at $(3, 24)$? $t =$ _____

What equation must $(x(t), y(t))$ satisfy?

(Enter your equation in terms of x and y: thus, if $y(t) = 3x(t)^2 - 6$, you would enter $y = 3x^2 - 6$.).

By hand, on a separate sheet of paper, verify that this equation is in fact satisfied by this parameterization.

At what values of t does the particle change direction?

At $t =$ _____ and _____

(e) Set up and find $\int_{C'} \vec{F} \cdot d\vec{r}$ using the parameterization in part ***(d)***.

With $a =$ _____ and $b =$ _____,

$\int_{C'} \vec{F} \cdot d\vec{r} = \int_a^b$ _____ $dt =$ _____

*(Note whether you get the same answer as you did in part **(b)**.)*

16.

(a) Compute $\int_{C_1} \vec{F} \cdot d\vec{r}$ when $\vec{F} = \langle x^2, xy \rangle$ and C_1 is the line segment from $(0, 0)$ to $(2, 2)$.

(b) Compute $\int_{C_2} \vec{F} \cdot d\vec{r}$ when $\vec{F} = \langle x^2, xy \rangle$ and C_2 is the line segment from $(2, 2)$ to $(0, 0)$.

17. If the wind in a region of space is given by $\vec{F} = \langle y + z, z - x, -z \rangle$ and a helicopter flies along the path given by $\vec{r}(t) = \langle 10\sin(t), 10\cos(t), (10-t)^2 \rangle$ as $0 \leq t \leq 4\pi$. Calculate the work done by the wind on the helicopter.

Hint. Set up your integral carefully and then use either integration by parts or an algebraic solver to compute the definite integral.

18. Let C_3 be the circle of radius 7 centered at the origin traveled counterclockwise. Compute $\int_{C_3} \langle M, N \rangle \cdot d\vec{r}$ when:

(a) $\langle M, N \rangle = \langle x, y \rangle$

(b) $\langle M, N \rangle = \langle -y, x \rangle$

(c) $\langle M, N \rangle = \langle 3, x \rangle$

19. Let C_4 be the curve given by traveling along the path given by $y = x^3 - x$ on the surface given by $z = xy$ as x goes from -1 to 2. What is the work done by $\langle x, z, x + y \rangle$?

Hint. Parametrize y in terms of x first, then use that relationship to give z as a function x.

13.5 Path-Independent Vector Fields and the Fundamental Theorem of Calculus for Line Integrals

Motivating Questions

- What characteristic of a vector field $\vec{F}$ will make $\int_C \vec{F} \cdot d\vec{r}$ have the same value for every oriented curve from a point P to a point Q?
- What special properties do gradient vector fields have?
- Given a gradient vector field $\vec{F}$, how can we efficiently find a potential function f so that $\vec{F} = \nabla f$?

13.5.1 Introduction

In Activity 13.4.4, Example 13.4.9, and Activity 13.4.5, we encountered situations where C_1 and C_2 are different oriented curves from a point P to a point Q and $\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$. In this section, we explore vector fields which have the property that *for all* points P and Q, if C_1 and C_2 are oriented paths from P to Q, then $\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$.

Preview Activity 13.5.1 In Activity 13.4.4, we considered the vector field $\vec{F}(x, y) = \langle y^2, 2xy+3 \rangle$ and two different oriented curves from $(-2, 5)$ to $(3, 30)$. We found that the value of the line integral of $\vec{F}$ was the same along those two oriented curves.

(a) Verify that $\vec{F}(x, y) = \langle y^2, 2xy + 3 \rangle$ is a gradient vector field by showing that $\vec{F} = \nabla f$ for the function $f(x, y) = xy^2 + 3y$.

(b) Calculate the change in the output of the scalar function f over the curves C_1 and C_2. In other words, what is the difference in the output of f at the start of the curve and the end of the curve? How does this value compare to the value of the line integral $\int_{C_1} \vec{F} \cdot d\vec{r}$ you found in Activity 13.4.4?

(c) Let C_3 be the line segment from $(1, 1)$ to $(3, 4)$. Calculate $\int_{C_3} \vec{F} \cdot d\vec{r}$ as well as $f(3, 4) - f(1, 1)$. Write a sentence that compares your answer to this part to your result for part 13.5.1.b.

13.5.2 Path-Independent Vector Fields

Hopefully Preview Activity 13.5.1 has prompted you to wonder about the phenomenon of the value of a line integral depending only on the initial and terminal points of the oriented path (rather than the oriented path itself) and how a potential function comes into play. We say that a vector field $\vec{F}$ defined on a region D is **path-independent** if $\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$ whenever C_1 and C_2 are oriented paths in D such that both curves start at point P and end at point Q.

In Activity 13.4.4 and Example 13.4.9, we encountered situations where we had evidence that a vector field was path-independent. However, since the definition of path-independence requires that the value of the line integral be the same for *every* possible path from one point to the other (regardless of

choice for the initial and final points), it doesn't appear that verifying a vector field is path-independent is an easy task.

Fortunately, one familiar class of vector fields can be shown to be path-independent. Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be a function for which ∇f is continuous on a region D. Suppose that P and Q are points in D and let C be a smooth oriented path from P to Q that is also contained in our region D. We consider $\int_C \nabla f \cdot d\vec{r}$ by fixing an arbitrary parametrization $\vec{r}(t)$ of C, $a \le t \le b$. Since we can write x, y, and z in terms of t, along C, the gradient of f is given by

$$\nabla f(\vec{r}(t)) = \langle f_x(\vec{r}(t)), f_y(\vec{r}(t)), f_z(\vec{r}(t))\rangle.$$

Hence, we can write the line integral of the vector field $\vec{F} = \nabla f$ over C in the following way:

$$\begin{aligned}\int_C \nabla f \cdot d\vec{r} &= \int_a^b \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)\, dt \\ &= \int_a^b \langle f_x(\vec{r}(t)), f_y(\vec{r}(t)), f_z(\vec{r}(t))\rangle \cdot \vec{r}'(t)\, dt\end{aligned}$$

If $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$, then the integrand above is

$$\begin{aligned}&\langle f_x(\vec{r}(t)), f_y(\vec{r}(t)), f_z(\vec{r}(t))\rangle \cdot \langle x'(t), y'(t), z'(t)\rangle \\ &= f_x\left(x(t), y(t), z(t)\right)\, x'(t) + f_y\left(x(t), y(t), z(t)\right)\, y'(t) \\ &\quad + f_z\left(x(t), y(t), z(t)\right)\, z'(t).\end{aligned}$$

Or more simply,

$$\nabla f \cdot \frac{d\vec{r}}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = \frac{d}{dt}\left(f(\vec{r}(t))\right).$$

Notice that this is exactly what the chain rule tells us $\frac{d}{dt} f(\vec{r}(t))$ is equal to. Therefore, we may apply the Fundamental Theorem of Calculus to obtain

$$\int_C \nabla f \cdot d\vec{r} = \int_a^b \frac{d}{dt} f(\vec{r}(t))\, dt = f(\vec{r}(b)) - f(\vec{r}(a)) = f(Q) - f(P).$$

In other words, gradient vector fields are path-independent vector fields, and we can evaluate line integrals of gradient vector fields by using a potential function. Remember that what we are calling gradient vector fields are also commonly called conservative vector fields.

Technically the argument above assumed that C was smooth, but we can replace C by a piecewise smooth curve by splitting the line integral up into the sum of finitely many line integrals along smooth curves.

This result is so important that it is frequently called the Fundamental Theorem of Calculus for Line Integrals, because of its similarity to the Fundamental Theorem of Calculus, which can be written as

$$\int_a^b \frac{df}{dx}(x)\, dx = f(b) - f(a).$$

Theorem 13.5.1 Fundamental Theorem of Calculus for Line Integrals. *Let f be a function for which ∇f is continuous on a region D. If P and Q are points in D and C is a piecewise smooth oriented path from P to Q in D, then*

$$\int_C \nabla f \cdot d\vec{r} = f(Q) - f(P).$$

Example 13.5.2

(a) Suppose we need to calculate the line integral of $\vec{F} = \langle 2xy - 2, x^2 + 3e^y \rangle$ over the path from $(2, 0)$ to $(4, -2)$ along the curve given by $(x - 2)^2 = -y/2$. If we use Theorem 13.4.6 to solve this, we would need to parameterize the curve, take a derivative of the parameterization, substitute into the line integral formula, and finally integrate the result. Fortunately, Theorem 13.5.1 allows us to be more efficient in how we approach this.

In order to use Theorem 13.5.1, we first need to verify that there exists a potential function $f(x, y)$ such that $\nabla f = \langle 2xy - 2, x^2 + 3e^y \rangle$. We will soon learn techniques in Activity 13.5.3 for finding a potential function, but for now, you should be able to verify for yourself that f will be of the form $f(x, y) = x^2y - 2x + 3e^y + k$, where k is a constant. Note that $f(4, -2) - f(2, 0) = (-32 - 8 + 3e^{-2} + k) - (0 - 4 + 3 + k) = 3e^{-2} - 41$, regardless of what value we choose for k. By Theorem 13.5.1,

$$\int_C \vec{F} \cdot d\vec{r} = f(4, -2) - f(2, 0) = 3e^{-2} - 41$$

If we changed the path in our line integral to be any other path between $(2, 0)$ and $(4, -2)$, then the result would still be $3e^{-2} - 41$.

(b) If we changed the path in our line integral to be some path C_1 from $(1, -1)$ to $(-1, 4)$ but did not change the vector field being used, then we would only need to evaluate $f(-1, 4) - f(1, -1)$. So for our new path C_1,

$$\begin{aligned} \int_{C_1} \vec{F} \cdot d\vec{r} &= f(-1, 4) - f(1, -1) \\ &= (4 - (-2) + 3e^4) - (-1 - 2 + 3e^{-1}) \\ &= 9 + 3(e^4 - e^{-1}). \end{aligned}$$

▲

Sometimes, we know that the vector field we need to compute a line integral in is the gradient vector field of a given function, in which case evaluation of the line integral is very efficient. Before investigating how to find a potential function, we have an activity to practice applying Theorem 13.5.1 when we are given the potential function.

Activity 13.5.2 Calculate each of the following line integrals.

(a) $\int_C \nabla f \cdot d\vec{r}$ if $f(x, y) = 3xy^2 - \sin(x) + e^y$ and C is the top half of the unit circle oriented from $(-1, 0)$ to $(1, 0)$.

(b) $\int_C \nabla g \cdot d\vec{r}$ if $g(x, y, z) = xz^2 - 5y^3 \cos(z) + 6$ and C is the portion of the helix $\vec{r}(t) = \langle 5\cos(t), 5\sin(t), 3t \rangle$ from $(5, 0, 0)$ to $(0, 5, 9\pi/2)$.

(c) $\int_C \nabla h \cdot d\vec{r}$ if $h(x, y, z) = 3y^2e^{y^3} - 5x\sin(x^3z) + z^2$ and C is the curve consisting of the line segment from $(0, 0, 0)$ to $(1, 1, 1)$, followed by the line segment from $(1, 1, 1)$ to $(-1, 3, -2)$, followed by the line segment from $(-1, 3, -2)$ to $(0, 0, 10)$.

In Activity 13.4.2, we used Clairaut's Theorem to argue that a vector field in $\mathbb{R}^2$ is not a gradient vector field when $\partial F_1/\partial y \neq \partial F_2/\partial x$. In Preview Activity 13.5.1, you verified that a given vector field was the gradient of a particular function of two variables. Clairaut's Theorem holds for functions of three vari-

ables. However, in that case there are six mixed partials to calculate, and thus it can be rather tedious. Activity 13.5.3 suggest a process for determining if a vector field in $\mathbb{R}^3$ is a gradient vector field as well as finding a potential function for the vector field.

Activity 13.5.3 Let $\vec{G}(x,y,z) = \langle 3e^{y^2} + z\sin(x), 6xye^{y^2} - z, 3z^2 - y - \cos(x)\rangle$ and $\vec{H}(x,y,z) = \langle 3x^2y, x^3 + 2yz^3, xz + 3y^2z^2\rangle$.

(a) If $\vec{G}$ and $\vec{H}$ are to be gradient vector fields, then there are functions g and h for which $\vec{G} = \nabla g$ and $\vec{H} = \nabla h$. If such functions g and h exist, what would g_y, g_z, h_x, h_y, and h_z be?

(b) Let $g_1(x,y,z) = 3xe^{y^2} + xyz - z\sin(x)$. Calculate $\partial g_1/\partial x$. Could g_1 be a potential function for the vector field $\vec{G}$?

(c) Find a function g so that $\partial g/\partial x = 3e^{y^2} + z\sin(x)$. Find a function h so that $\partial h/\partial x = 3x^2y$.

(d) When finding the most general anti-derivative for a function of one variable, we add a constant of integration (usually denoted by C) to capture the fact that any constant will become 0 through differentiation.

(i) When taking the partial derivative with respect to x of a function of x, y, and z, what variables can appear in terms that become 0 in the partial derivative because they are treated as constants?

(ii) What does this tell you should be added to g and h in the previous part to make them the most general possible functions with the desired partial derivatives with respect to x?

(e) Now calculate $\partial g/\partial y$ and $\partial h/\partial y$ based on your choices for part 13.5.3.c. Write a few sentences to explain why this tells you that we must have

$$g(x,y,z) = 3xe^{y^2} - z\cos(x) - yz + m_1(z)$$

and

$$h(x,y,z) = x^3y + y^2z^3 + m_2(z)$$

for some functions m_1 and m_2 depending only on z.

(f) Calculate $\frac{\partial g}{\partial z}$ and $\frac{\partial h}{\partial z}$ for the functions in the part above. Notice that m_1 and m_2 are functions of z alone, so taking a partial derivative with respect to z is the same as taking an ordinary derivative, and thus you may use the notation $m_1'(z)$ and $m_2'(z)$.

(g) Explain why $\vec{G}$ is a gradient vector field but $\vec{H}$ is not a gradient vector field. Find a potential function for $\vec{G}$.

By following the methodology laid out in Activity 13.5.3 to show that a given vector field is a gradient vector field, the Fundamental Theorem of Calculus for Line Integrals makes evaluating some line integrals much simpler now.

Activity 13.5.4 Calculate each of the following line integrals.

(a) $\int_C \vec{F}\cdot d\vec{r}$ if $\vec{F}(x,y) = \langle 2x, 2y\rangle$ and C is the line segment from $(1,2)$ to $(-1,0)$.

Hint. Find $f(x,y)$ such that $\nabla f = \vec{F}$.

(b) $\int_C \vec{G}\cdot d\vec{r}$ if $\vec{G}(x,y) = \langle 4x^3 - 12y\cos(xy), 9y^2 - 12x\cos(xy)\rangle$ and C is the

portion of the unit circle from $(0,-1)$ to $(0,1)$.

(c) $\int_C \vec{H}\cdot d\vec{r}$ if $\vec{H}(x,y,z)=\langle H_1,H_2,H_3\rangle$ with

$$\begin{aligned}H_1(x,y,z)&=e^{z^2}+2xy^3z+\cos(x)-y^3\sin(x)\\H_2(x,y,z)&=2ye^{y^2}+3x^2y^2z+3y^2z^2+3y^2\cos(x)\\H_3(x,y,z)&=x^2y^3+2xze^{z^2}+2y^3z-4z^3\end{aligned}$$

and C is the curve consisting of the line segment from $(1,1,1)$ to $(3,0,3)$, followed by the line segment from $(3,0,3)$ to $(1,5,-1)$, followed by the line segment from $(1,5,-1)$ to $(0,0,0)$.

13.5.3 Line Integrals Along Closed Curves

Recall that an oriented curve C is **closed** if the curve has the same initial and terminal point. A typical example of a closed curve would be a circle (with an orientation of which way to go around), but we could also consider something like the square with vertices $(1,1)$, $(-1,1)$, $(-1,-1)$, and $(1,-1)$, oriented clockwise (or counterclockwise). Recall that we sometimes use the symbol $\oint$ for a line integral when the curve is closed and that if $C=C_1+C_2$, then $\int_C \vec{F}\cdot d\vec{r}=\int_{C_1}\vec{F}\cdot d\vec{r}+\int_{C_2}\vec{F}\cdot d\vec{r}$.

Activity 13.5.5 Suppose that $\vec{F}$ is a continuous path-independent vector field (in $\mathbb{R}^2$ or $\mathbb{R}^3$) on some region D.

(a) Let P and Q be points in D and let C_1 and C_2 be oriented curves from P to Q. What can you say about $\int_{C_1}\vec{F}\cdot d\vec{r}$ and $\int_{C_2}\vec{F}\cdot d\vec{r}$?

(b) Let $C=C_1-C_2$. Explain why C is a closed curve.

(c) Calculate $\oint_C \vec{F}\cdot d\vec{r}$.

(d) Write a sentence that summarizes what we can conclude about line integrals of $\vec{F}$ at this point in the activity.

(e) Now let us suppose that $\vec{G}$ is a continuous vector field on a region D for which $\oint_C \vec{G}\cdot d\vec{r}=0$ for all closed curves C. Pick two points P and Q in D. Let C_1 and C_2 be oriented curves from P to Q. What type of curve is $C=C_1-C_2$?

(f) What is $\oint_C \vec{G}\cdot d\vec{r}$? Why?

(g) What does that tell you about the relationship between $\int_{C_1}\vec{G}\cdot d\vec{r}$ and $\int_{C_2}\vec{G}\cdot d\vec{r}$?

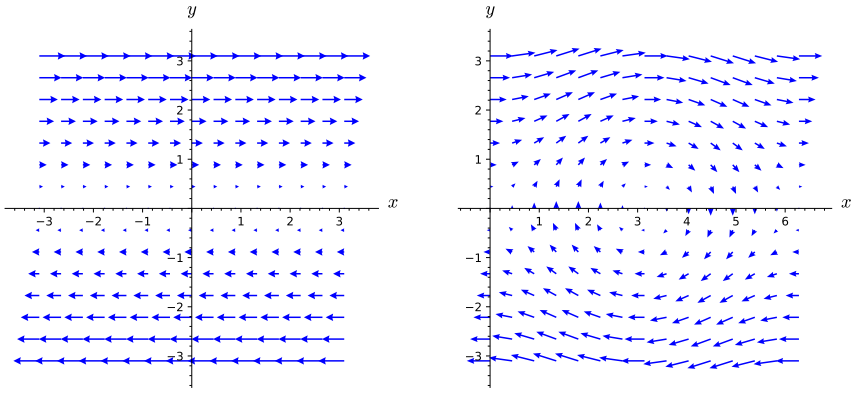
(h) Explain why this shows that $\vec{G}$ is path-independent.

We summarize the result of Activity 13.5.5 with the theorem below. Although this theorem is not a terribly useful way to show that a vector field is path-independent, it can be a useful way to show that a vector field is *not* path-independent: If you can find a closed curve around which the circulation is not zero, then the vector field is ***not*** path independent.

Theorem 13.5.3 *Let $\vec{F}$ be a continuous vector field on a region D. Suppose that C is a closed curve in D. The circulation of $\vec{F}$ along C, given by $\oint_C \vec{F}\cdot d\vec{r}$, is zero if and only if $\vec{F}$ is path-independent.*

The following activity gives you a chance to reason about path-independence based purely on a graphical representation of a vector field.

Activity 13.5.6 Explain why neither of the vector fields in Figure 13.5.4 is path-independent.



(a) $\vec{F}$ **(b)** $\vec{G}$

Figure 13.5.4 Two vector fields that are not path-independent.

13.5.4 What other vector fields are path-independent?

Recall that in single variable calculus, The Second Fundamental Theorem of Calculus tells us that given a constant c and a continuous function f, there is a unique function $A(x)$ for which $A(c) = 0$ and $\frac{dA}{dx}(x) = f(x)$. In particular, $A(x) = \int_c^x f(t)\,dt$ is this function. We are about to investigate an analog of this result for path-independent vector fields, but first we require two additional definitions.

If D is a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$, we say that D is **open** provided that for every point in D, there is a disc (in $\mathbb{R}^2$) or ball (in $\mathbb{R}^3$) centered at that point such that every point of the disc/ball is contained in D. For example, the set of points (x, y) in $\mathbb{R}^2$ for which $x^2 + y^2 < 1$ is open, since we can always surround any point in this set by a tiny disc contained in the set (as illustrated by point P in Figure 13.5.5). However, if we change the inequality to $x^2 + y^2 \leq 1$, then the set is not open, as any point on the circle $x^2 + y^2 = 1$ cannot be surrounded by a disc contained in the set; any disc surrounding a point on that circle will contain points outside the set, that is with $x^2 + y^2 > 1$ (as illustrated by the point Q in Figure 13.5.5). We will also say that a region D is **path-connected** provided that for every pair of points in D, there is a path from one to the other contained in D.

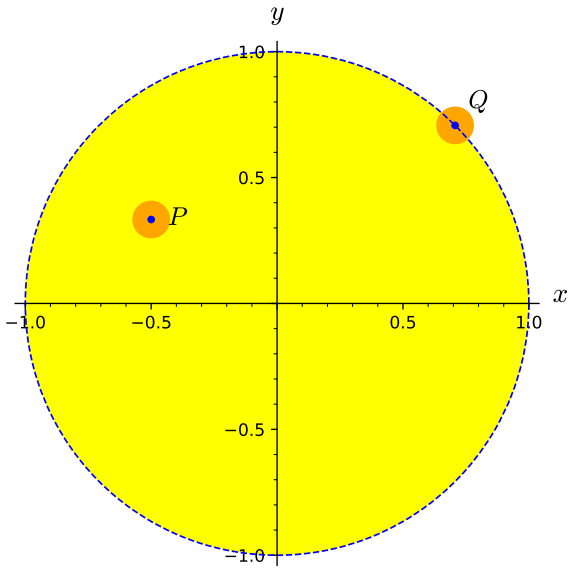


Figure 13.5.5 The open set $\{(x,y)|x^2+y^2<1\}$ is plotted in yellow with a point Q on the boundary that shows why $\{(x,y)|x^2+y^2\leq 1\}$ is not open

Activity 13.5.7 Let $\vec{F}=\langle F_1,F_2\rangle$ be a continuous, path-independent vector field on an open, path-connected region D. We will assume that D is in $\mathbb{R}^2$ and $\vec{F}$ is a two-dimensional vector field, but the ideas below generalize completely to $\mathbb{R}^3$. We want to define a function f on D by using the vector field $\vec{F}$ and line integrals, much like the Second Fundamental Theorem of Calculus allows us to define an antiderivative of a continuous function using a definite integral. To that end, we assign $f(x_0,y_0)$ an arbitrary value. (Setting $f(x_0,y_0)=0$ is probably convenient, but we won't explicitly tie our hands. Just assume that $f(x_0,y_0)$ is defined to be some number.) Now for any other point (x,y) in D, define

$$f(x,y)=f(x_0,y_0)+\int_C \vec{F}\cdot d\vec{r},$$

where C is any oriented path from (x_0,y_0) to (x,y). Since D is path-connected, such an oriented path must exist. Since $\vec{F}$ is path-independent, f is well-defined. If different paths from (x_0,y_0) to (x,y) gave different values for the line integral, then we would not be sure what $f(x,y)$ really is.

To better understand this mysterious function f we've now defined, let's start looking at its partial derivatives.

(a) Since D is open, there is a disc (perhaps very small) surrounding (x,y) that is contained in D, so fix a point (a,b) in that disc. Since D is path-connected, there is a path C_1 from (x_0,y_0) to (a,b). Let C_y be the line segment from (a,b) to (a,y) and let C_x be the line segment from (a,y) to (x,y). (See Figure 13.5.6.) Rewrite $f(x,y)$ as a sum of $f(x_0,y_0)$ and line integrals along C_1, C_y, and C_x.

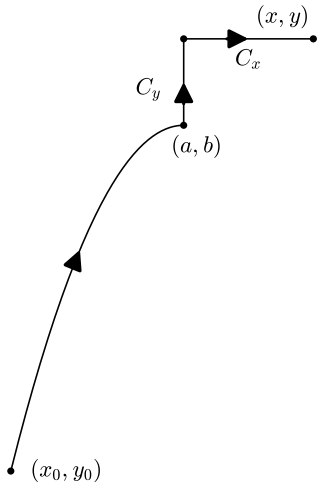


Figure 13.5.6 A piecewise smooth oriented curve from (x_0, y_0) to (x, y).

(b) Notice that we can parametrize C_y by $\langle a, t\rangle$ for $b \le t \le y$. Find a similar parametrization for C_x.

(c) Use the parametrization from above to write $\int_{C_y} \vec{F} \cdot d\vec{r}$ and $\int_{C_x} \vec{F} \cdot d\vec{r}$ as single variable integrals in the manner of Section 13.4. Use the fact that $\vec{F}(x, y) = \langle F_1(x, y), F_2(x, y)\rangle$ to express your integrals in terms of F_1 and F_2 without any dot products.

(d) Rewrite your expression for $f(x, y)$ using a line integral along C_1 and the single variable integrals above.

(e) Notice that your expression for $f(x, y)$ from the previous part only depends on x as the upper limit of an single variable integral. Use the Second Fundamental Theorem of Calculus to calculate $f_x(x, y)$.

(f) To calculate $f_y(x, y)$, we continue to consider a path C_1 from (x_0, y_0) to (a, b), but now let L_x be the line segment from (a, b) to (x, b) and let L_y be the line segment from (x, b) to (y, b). Modify the process you used to find $f_x(x, y)$ to find $f_y(x, y)$.

(g) What can you conclude about the relationship between ∇f and $\vec{F}$? What does this tell you about $\vec{F}$ beyond that it is path-independent and continuous?

We summarize the result of Activity 13.5.7 below. Much like the Second Fundamental Theorem of Calculus, which tells us that a function is an antiderivative for another function, but leaves the antiderivative in terms of a definite integral, this theorem tells us that a function is a potential function for a vector field, but the definition of the potential function is in terms of a line integral.

Theorem 13.5.7 Path-Independent Vector Fields. *If $\vec{F}$ is a path-independent vector field on an open, path-connected region D, then $\vec{F}$ is a gradient vector field on D. If the circulation of $\vec{F}$ around any closed loop in an open, path-connected region D is zero, then $\vec{F}$ is a path-independent vector field on D. Furthermore, if P is a point in D and $f(P)$ is fixed, then for a*

point Q in D and an oriented curve C from P to Q in D, the function

$$f(Q) = f(P) + \int_C \vec{F} \cdot d\vec{r}$$

is a potential function for $\vec{F}$.

Summary

- Gradient vector fields are path-independent, and if C is an oriented curve from (x_1, y_1) to (x_2, y_2), then

$$\int_C \nabla f \cdot d\vec{r} = f(x_2, y_2) - f(x_1, y_1)$$

  with the analogous result holding if f is a function of three variables.

- A vector field is path-independent if and only if the circulation around every closed curve in its domain is 0.

- If a vector field $\vec{F}$ is path-independent, then there exists a function f such that $\nabla f = \vec{F}$. That is, $\vec{F}$ is a conservative or gradient vector field.

13.5.5 Exercises

1. For each of the following decide whether the vector field could be a gradient vector field. Be sure that you can justify your answer.

(a) $\vec{F}(x, y, z) = \frac{z}{\sqrt{x^2+z^2}}\,\vec{i} + \frac{y}{\sqrt{x^2+z^2}}\,\vec{j} + \frac{x}{\sqrt{x^2+z^2}}\,\vec{k}$

$\vec{F}(x, y, z)$ is (☐ a gradient vector field ☐ not a gradient vector field)

(b) $\vec{F}(x, y, z) = 2x\,\vec{i} + 2z\,\vec{j}$

$\vec{F}(x, y, z)$ is (☐ a gradient vector field ☐ not a gradient vector field)

(c) $\vec{F}(x, y, z) = \langle 2yz, 2xz, 2xy \rangle$

$\vec{F}(x, y, z)$ is (☐ a gradient vector field ☐ not a gradient vector field)

(d) $\vec{F}(x, y, z) = 4y\,\vec{i} + 4x\,\vec{j}$

$\vec{F}(x, y, z)$ is (☐ a gradient vector field ☐ not a gradient vector field)

2. Find f if $\nabla f = 2xy\,\vec{i} + (x^2 + 5y^4)\,\vec{j}$

$f =$ _____

3. Let $\nabla f = -4xe^{-x^2}\sin(5y)\,\vec{i} + 10e^{-x^2}\cos(5y)\,\vec{j}$. Find the change in f between $(0, 0)$ and $(1, \pi/2)$ in two ways.

(a) First, find the change by computing the line integral $\int_C \nabla f \cdot d\vec{r}$, where C is a curve connecting $(0, 0)$ and $(1, \pi/2)$.

The simplest curve is the line segment joining these points. Parameterize it:

with $0 \le t \le 1$, $\vec{r}(t) =$ _____$\vec{i}$+ _____$\vec{j}$

So that $\int_C \nabla f \cdot d\vec{r} = \int_0^1$ _____ dt

Note that this isn't a very pleasant integral to evaluate by hand (though we could easily find a numerical estimate for it). It's easier to find $\int_C \nabla f \cdot d\vec{r}$ as the sum $\int_{C_1} \nabla f \cdot d\vec{r} + \int_{C_2} \nabla f \cdot d\vec{r}$, where C_1 is the line segment from $(0, 0)$ to $(1, 0)$ and C_2 is the line segment from $(1, 0)$ to $(1, \pi/2)$. Calculate these integrals to find the change in f.

$\int_{C_1} \nabla f \cdot d\vec{r} =$ _____

$\int_{C_2} \nabla f \cdot d\vec{r} =$ _____

So that the change in $f = \int_C \nabla f \cdot \vec{r} = \int_{C_1} \nabla f \cdot d\vec{r} + \int_{C_2} \nabla f \cdot d\vec{r} =$ _____

(b) By computing values of f. To do this,
First find $f(x, y) =$ _____
Thus
$f(0, 0) =$ _____ and $f(1, \pi/2) =$ _____,
and the change in f is _____.

4. Consider the vector field $\mathbf{F}(x, y, z) = (2z + 4y)\,\mathbf{i} + (3z + 4x)\,\mathbf{j} + (3y + 2x)\,\mathbf{k}$.
a) Find a function f such that $\mathbf{F} = \nabla f$ and $f(0, 0, 0) = 0$.
$f(x, y, z) =$ _____
b) Suppose C is any curve from $(0, 0, 0)$ to $(1, 1, 1)$. Use part a) to compute the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$.

5. Consider the vector field $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.
a) Find a function f such that $\mathbf{F} = \nabla f$ and $f(0, 0, 0) = 0$.
$f(x, y, z) =$ _____
b) Use part a) to compute the work done by $\boldsymbol{F}$ on a particle moving along the curve C given by $\mathbf{r}(t) = (1 + 3\sin t)\,\mathbf{i} + \left(1 + \sin^2 t\right)\mathbf{j} + \left(1 + \sin^3 t\right)\mathbf{k}$, $0 \le t \le \frac{\pi}{2}$.

6. Let $\mathbf{F}$ be the radial force field $\mathbf{F} = x\mathbf{i} + y\mathbf{j}$. Find the work done by this force along the following two curves, both which go from (0, 0) to (7, 49). (Compare your answers!)

A. If C_1 is the parabola: $x = t,\ y = t^2,\ 0 \le t \le 7$, then $\displaystyle\int_{C_1} \mathbf{F} \cdot d\mathbf{r} =$ _____

B. If C_2 is the straight line segment: $x = 7t^2,\ y = 49t^2,\ 0 \le t \le 1$, then $\displaystyle\int_{C_2} \mathbf{F} \cdot d\mathbf{r} =$ _____

7. Consider the vector field $\mathbf{F}(x, y, z) = (2z + 5y)\,\mathbf{i} + (4z + 5x)\,\mathbf{j} + (4y + 2x)\,\mathbf{k}$.
a) Find a function f such that $\mathbf{F} = \nabla f$ and $f(0, 0, 0) = 0$.
$f(x, y, z) =$ _____
b) Suppose C is any curve from $(0, 0, 0)$ to $(1, 1, 1)$. Use part a) to compute the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$.

8. Suppose $\mathbf{F} = \mathbf{F}(x, y, z)$ is a gradient field with $\mathbf{F} = \nabla f$, S is a level surface of f, and C is a curve on S. What is the value of the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$?

9. If $\vec{F} = \nabla(5x^2 + 4y^4)$, find $\int_C \vec{F} \cdot d\vec{r}$ where C is the quarter of the circle $x^2 + y^2 = 1$ in the first quadrant, oriented counterclockwise.
$\int_C \vec{F} \cdot d\vec{r} =$ _____

10. Use the Fundamental Theorem of Line Integrals to calculate $\int_C \vec{F} \cdot d\vec{r}$ exactly, if $\vec{F} = x^{2/3}\,\vec{i} + e^{y/3}\,\vec{j}$, and C is the quarter of the unit circle in the first quadrant, traced counterclockwise from $(1, 0)$ to $(0, 1)$.
$\int_C \vec{F} \cdot d\vec{r} =$ _____

11. For the vector field $\vec{G} = (ye^{xy} + \cos(x + y))\,\vec{i} + (xe^{xy} + \cos(x + y))\,\vec{j}$, find the line integral of $\vec{G}$ along the curve C from the origin along the x-axis to the point $(5, 0)$ and then counterclockwise around the circumference of the circle $x^2 + y^2 = 25$ to the point $(5/\sqrt{2}, 5/\sqrt{2})$.
$\int_C \vec{G} \cdot d\vec{r} =$ _____

12. The domain of $f(x, y)$ is the xy-plane, and values of f are given in the table below.

$y\backslash x$	0	1	2	3	4
0	55	56	57	56	56
1	56	56	56	56	58
2	58	60	61	61	63
3	59	59	59	59	61
4	62	62	63	63	63

Find $\int_C \text{grad } f \cdot d\vec{r}$, where C is

(a) A line from $(0, 2)$ to $(2, 4)$.

$\int_C \text{grad } f \cdot d\vec{r} =$ _____

(b) A circle of radius 1 centered at $(1, 1)$ traversed counterclockwise.

$\int_C \text{grad } f \cdot d\vec{r} =$ _____

13. The path C is a line segment of length 10 in the plane starting at $(3, 5)$. For $f(x, y) = 6x + 8y$, consider

$$\int_C \nabla f \cdot d\vec{r}.$$

(a) Where should the other end of the line segment C be placed to maximize the value of the integral?

At $x =$ _____, $y =$ _____

(b) What is the maximum value of the integral?

maximum value = _____

14. Let $\vec{F} = 5y\vec{i} + 6x\vec{j}$, $\phi = \frac{8}{3}x^3 + 5xy$, and $h = y - 4x^2$.

(a) Find each of the following:

$\vec{F} - \nabla\phi =$ _____

$\nabla h =$ _____

How are $\vec{F} - \nabla\phi$ and ∇h related?

$\vec{F} - \nabla\phi =$ _____ ∇h

(Note that this shows that $\vec{F} - \nabla\phi$ is parallel to ∇h.)

(b) Use ϕ and the Fundamental Theorem of Calculus for Line Integrals to evaluate $\int_C \vec{F} \cdot d\vec{r}$, where C is the oriented path on a contour of h from $P(0, 8)$ to $Q(5, 108)$.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

15. For each of the following, decide if the given vector field is a gradient of a function f. If so, give find the function f and enter it as your answer; if not, enter the word *none* for your answer and be sure that you are able to explain why this is the case.

(a) $[4x\cos(y^2)]\,\vec{i} + [4y\cos(x^2)]\,\vec{j}$: $f =$ _____

(b) $[8x]\,\vec{i} + [6y]\,\vec{j}$: $f =$ _____

(c) $[8xy]\,\vec{i} + [4x^2]\,\vec{j}$: $f =$ _____

(d) $[8\cos(x^2 + 4y^2) - 16x^2\sin(x^2 + 4y^2)]\,\vec{i} + [-64xy\sin(x^2 + 4y^2)]\,\vec{j}$: $f =$ _____

16. Let $\mathbf{F}(x, y) = \frac{-y\mathbf{i} + x\mathbf{j}}{x^2 + y^2}$ and let C be the circle $\mathbf{r}(t) = (\cos t)\,\mathbf{i} + (\sin t)\,\mathbf{j}$, $0 \le t \le 2\pi$.

A. Compute $\frac{\partial Q}{\partial x}$

Note: Your answer should be an expression of x and y; e.g. "3xy - y"

B. Compute $\frac{\partial P}{\partial y}$

Note: Your answer should be an expression of x and y; e.g. "3xy - y"

C. Compute $\int_C \mathbf{F} \cdot d\mathbf{r}$

Note: Your answer should be a number
D. Is $\boldsymbol{F}$ conservative? Type Y if yes, type N if no.

17. For each of the following line integrals:

- Determine if the Fundamental Theorem of Calculus for Line Integrals applies.
- If the Fundamental Theorem of Calculus for Line Integrals applies, then find the potential function and use this to evaluate the line integral
- If the Fundamental Theorem of Calculus for Line Integrals does not apply, then describe where the process laid out in Preview Activity 13.5.1 fails.

(a) The line integral of $\vec{F} = \langle yz, xz, y \rangle$ along the helix of radius of 3 given by $\vec{r}(t) = \langle 3\sin(t), 3\cos(t), \frac{4}{\pi}t \rangle$ as $-\pi \leq t \leq \pi$.

(b) The line integral of $\vec{F} = \langle \sin(yz), xz\cos(yz) - z\sin(y), xy\cos(yz) + \cos(y) \rangle$ along the line segment from $(0, \pi, 3)$ to $(2, -1, 2\pi)$.

(c) The integral $\int_C y^2\,dx + 2xy\,dy$ where C is the parabolic path along $y = x^2$ from $(-1, 1)$ to $(4, 16)$.

18. Let $\vec{F} = \langle F_1, F_2, F_3 \rangle$, where $F_1 = 3xy^2 + z$ and $F_3 = yz^2 + \cos(z)$.

(a) Give a component function F_2 such that $\vec{F}$ is a gradient vector field.

(b) Give a component function F_2 such that $\vec{F}$ is *not* a gradient vector field.

19. This exercise focuses on reasoning graphically about line integrals and vector fields.

(a) Find a closed curve on which the line integral of the vector field pictured below will not be zero. Be sure to clearly define your curve and explain why the line integral over your curve is non-zero.

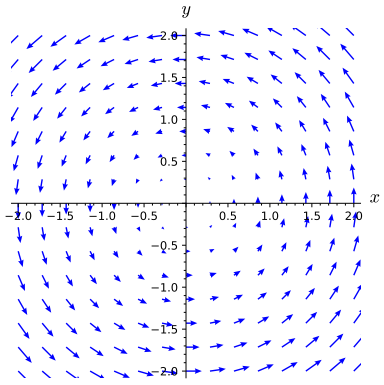


Figure 13.5.8 A vector field

(b) Explain why you think the following vector field is path independent or not.

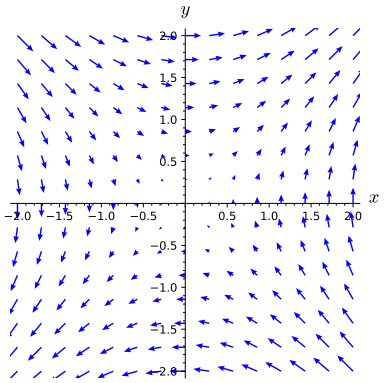


Figure 13.5.9 Another vector field

20. Compute $\int_C ye^z\,dx + xe^z\,dy + xye^z\,dz$ where C is given by $\langle t^2, t^3, t-1\rangle$ for $1 \le t \le 2$.

21. Let $\vec{F}(x,y,z)$ be a vector field such that $\vec{F} = \nabla f$ for a smooth function $f(x,y,z)$. Let S be a level surface of f and let C_1 be curve that stays on the surface S. Let C_2 be a curve that begins and ends on S but does have points not on S.

(a) What is the value of $\int_{C_1} \vec{F} \cdot d\vec{r}$?

(b) What is the value of $\int_{C_2} \vec{F} \cdot d\vec{r}$?

13.6 Line Integrals of Scalar Functions

Motivating Questions

- How can you measure the accumulation of a scalar function over a curve in space?
- How can you efficiently calculate the scalar line integral over a curve in space?

13.6.1 Introduction

In Section 13.3, the idea of a line integral was introduced by looking at the work done by a vector field when traveling along C, a path in space. In particular, the line integral measured the accumulated amount of the vector field that is along the path C (in the direction of travel.) Definition 13.3.9 shows how to use a Riemann sum to measure the accumulation of the vector field in the direction of travel along the given curve. Additionally, Theorem 13.4.6 shows how the line integral $\int_C \vec{F} \cdot d\vec{r}$ can be efficiently calculated in terms of the vector field and the derivative of *any* parameterization of C. Philosophically, the Riemann sum in Definition 13.3.9 is adding up the scalar function given by the dot product of the vector field and the unit vector in the direction of travel (along the curve).

We will devote the rest of this section to answering these questions about generalizing the ideas of line integrals as fully as possible.

- How can we measure the accumulation of a scalar valued function along a curve in space?

- What would that accumulation measure physically?
- What is the most efficient way to calculate this accumulation?

Preview Activity 13.6.1 In order to pay for tuition, you take a job driving a mining machine that collects a very valuable mineral called copium. Copium is only produced on the surface and is mined by scooping up the soil at the front of your machine, so the amount of copium ore collected depends on the density of the ore and the distance driven by the mining machine. The plot of land you are mining has been surveyed for the density of copium ore and is presented in the contour plot below.

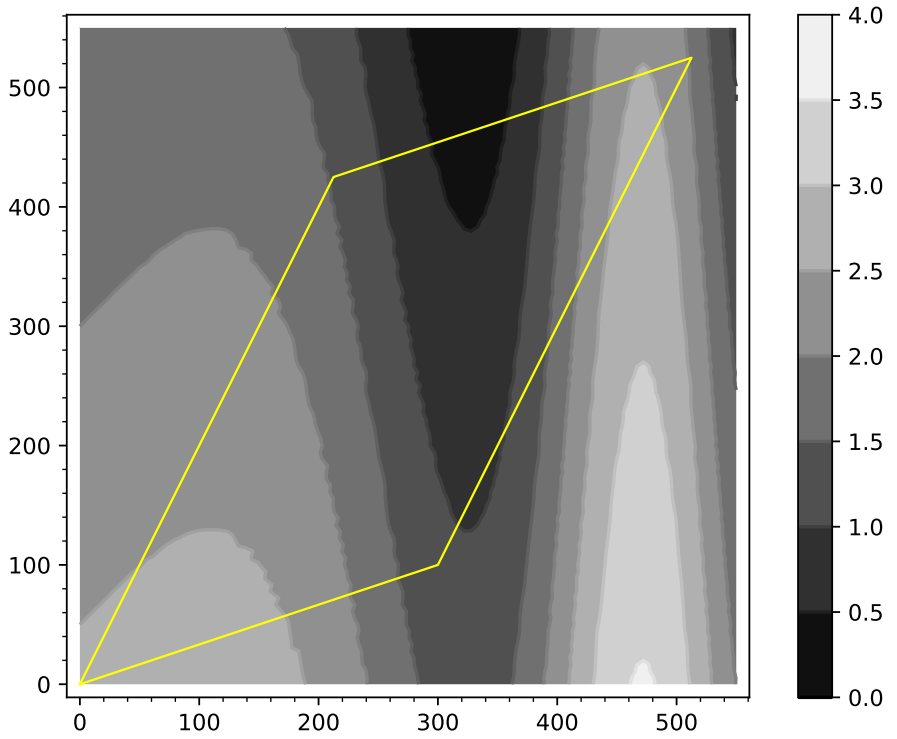


Figure 13.6.1 A plot of land with density of copium deposits

(a) Estimate the amount of copium that would be mined from driving along the left side of plot. You should write a few sentences about how you got your estimate based on the copium density and length of the path. (Did you use more than one piece?)

(b) Estimate the amount of copium that would be mined by driving along the entire outer edge of the plot. You should write a few sentences about how you got your estimate based on the copium density and length of the paths.

(c) Estimate the amount of copium that would be mined from the scraping the curved path shown in Figure 13.6.2. You should use at least 3 segments in your estimate. You should write a few sentences about how you got your estimate based on the copium density and length of the paths.

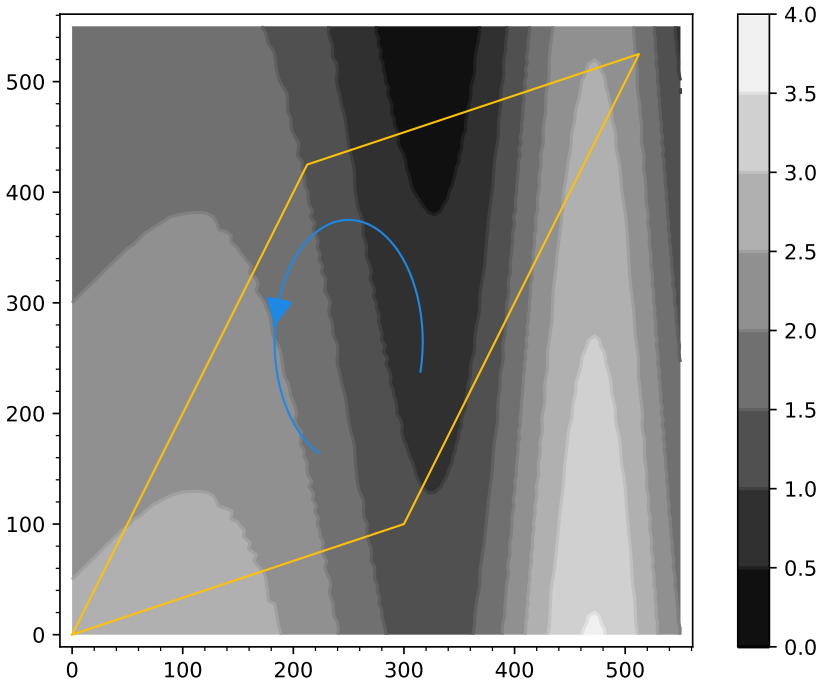


Figure 13.6.2 A plot of land with copium Density and a path plotted in blue

(d) Estimate the amount of copium that would be mined from the scraping the curved path shown below. You should use at least 3 segments in your estimate. You should also explain how and why your answer to this question is different or similar to the previous task.

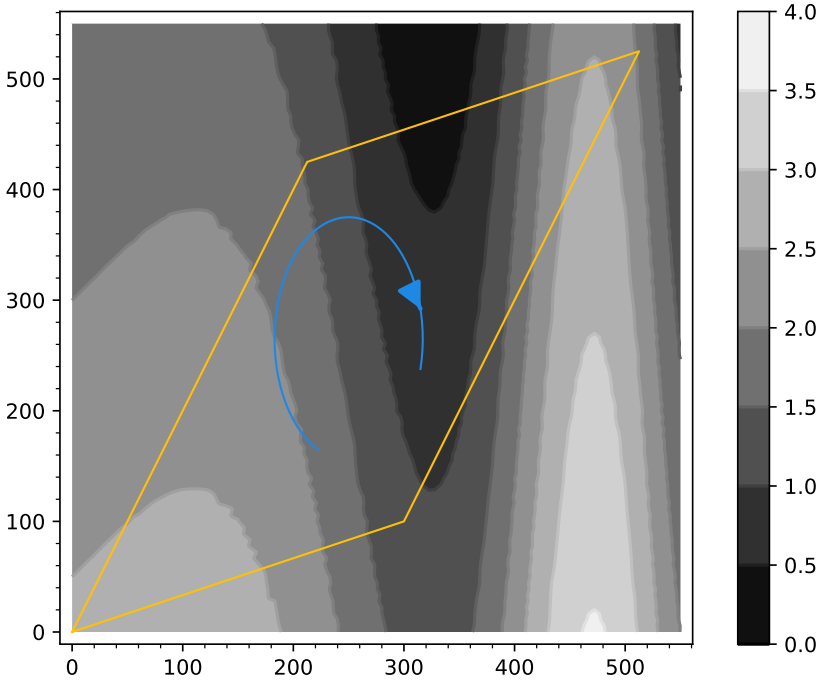


Figure 13.6.3 A plot of land with copium density and a path plotted in blue

13.6.2 Defining line integrals of scalar functions

In Preview Activity 13.6.1, you approximated the distance traveled for various paths and multiplied by the density of the copium on each piece of the path. In contrast to the line integral of a vector field, the calculations of the ore mined does not depend on what direction the path was traveled. We will now use these same ideas to give precise meaning to the measurement of the accumulation of a scalar function's output over a path in space. As you might guess, we will use a classic calculus approach of 1) making an approximation, 2) describe how this approximation changes under smaller scales, and 3) use a limit to find the exact value. In the preceeding discussion, we will briefly cover this CCA and return to the details of step 2 later in this section.

Let f be a continuous function of x, y, and z for some open set around C, a curve from a point P to a point Q. We will begin to approximate the accumulation of the output of f over C by breaking C into pieces with boundary points $P = \vec{r}_0, \vec{r}_1, \ldots, \vec{r}_n - 1, \vec{r}_n = Q$. The curve C_i is the part of C that goes from $\vec{r}_{i-1}$ to $\vec{r}_i$ and $\Delta\vec{r}_i$ is the displacement vector from $\vec{r}_{i-1}$ to $\vec{r}_i$

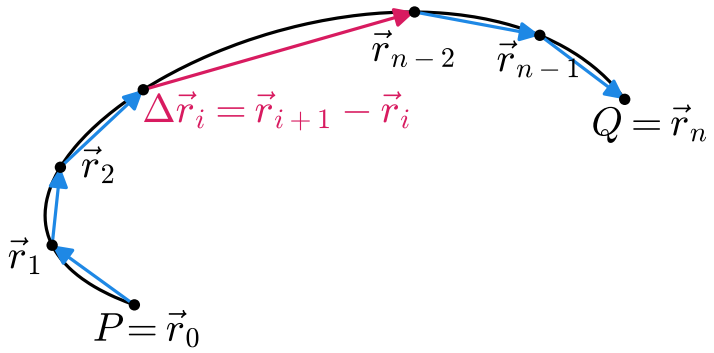


Figure 13.6.4 A curve in space with segments given by $\mathbf{r}_i$

For Step 1 of the CCA, we can approximate the accumulation of f over C with the following sum

$$\sum_{i=1}^{n} f(\vec{r}_i^*) \left\|\vec{r}_i - \vec{r}_{i-1}\right\|,$$

where $f(r_i^*)$ is the output of f for some $\vec{r}_i^* \in C_i$. Similar to our work on line integrals of a vector field, as this sum uses more pieces and all of the lengths of the pieces goes to zero (i.e. $\|\Delta\vec{r}_i\| \to 0$), we would expect that the sum will approach the actual accumulation of f over C. Notice that it won't matter how we select the point $\vec{r}_i^*$ that is used in each piece to evaluate the output of f since evaluating the limit as the length of C_i gets smaller will ensure that the output value chosen will be within a shrinking error from the average value on each piece. Evaluating the limit of the sum above as the size of all of the pieces goes to zero will transform our Riemann sum into an integral that will measure the accumulation of the output of f over C.

Definition 13.6.5 The Line Integral of a Scalar Function. Let C be a curve from a point P to a point Q in space. Let f be a continuous function of x, y, and z for some open set around C. The **line integral of f over C** is defined as

$$\int_C f ds = \lim_{n\to\infty} \sum_{i=1}^{n} f(\vec{r}_i^*) \left\|\vec{r}_i - \vec{r}_{i-1}\right\|$$

where $\vec{r}_i$ are points such that $P = \vec{r}_0, \vec{r}_1, \ldots, \vec{r}_n - 1, \vec{r}_n = Q$ and as $n \to \infty$ the distance between $\vec{r}_{i-1}$ and $\vec{r}_i$ goes to zero.

The integral

$$\int_C f\,ds$$

is called the **scalar line integral of f over the curve C.** ◆

The notation for a scalar line integral ($\int_C f\,ds$) may not immediately make sense. As with the other types of integration we have done (single variable integration, double integrals, line integrals of vector fields, etc.), the subscript of the integral symbol denotes the region of integration. In the case of a scalar line integral, the region of integration is a collection of points given by a curve in space. The function we are integrating is f, a scalar-valued function of multiple variables. The differential ds may seem unusual to you. If you remember from Section 10.4, s is the arc length of a curve in space. So the differential ds in the scalar line integral notation means that we are adding up the output of f over steps in arc length. This should make sense in terms of how we set up our Riemann sums. We did not set up the pieces of our curve as steps in x, y, or z, but rather as steps in arc length (estimated by $\|\vec{r}_{i+1} - \vec{r}_i\|$). This may feel similar to situations such as double integrals, where we generically used dA for the differential, but the different contexts called for different values of dA. For instance, in polar coordinates, we use $dA = r\,dr\,d\theta$

Example 13.6.6 Before getting into the details of computing line integrals of scalar functions, we will first make arguments about when the line integral of a scalar function is positive, negative, or zero. For all of this example, we will use the curve C given by $y = x^2$ for $-1 \leq x \leq 2$. We also will define the following three scalar-valued functions on $\mathbb{R}^2$.

- $f(x, y) = x$
- $g(x, y) = -y$
- $h(x, y) = y - x^2$

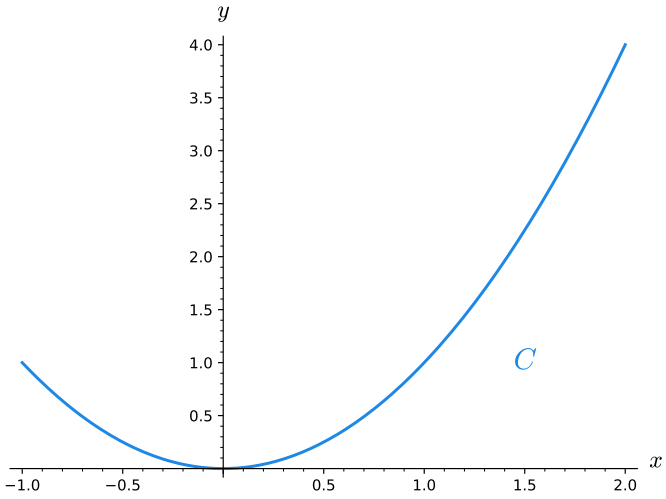


Figure 13.6.7 The curve C used in the line integral for Example 13.6.6

(a) For our first case, we will consider the line integral

$$\int_C f\,ds.$$

Remember that the line integral of f over C will measure the accumulation of the output of f over the points on the curve C. Figure 13.6.8

shows that the blue branch of the curve C (for $0 \leq x \leq 2$) will yield positive output values for f, while the magenta part, which has $-1 \leq x \leq 0$, will produce negative output values for f. Intuitively, there is more of the curve with positive f outputs than negative f outputs, so we would expect $\int_C f\,ds$ to be positive. However, the magnitude of the f-values produced must also be accounted for.

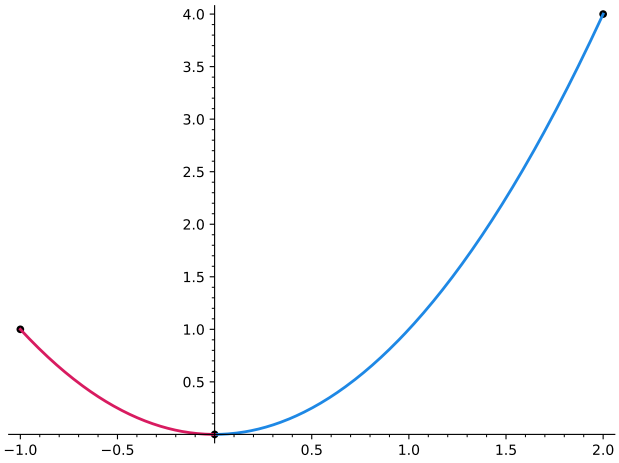


Figure 13.6.8 The curve C split into regions with positive (blue) and negative (magenta) outputs of f

If we break our curve C into three pieces by dividing at $x = 0$ and $x = 1$ as shown in Figure 13.6.9, we can make a rigorous argument about why the result of $\int_C f\,ds$ will be positive. We will call the magenta, green, and blue sections of Figure 13.6.9 C_1, C_2, and C_3, respectively. Using the notation of Properties of Line Integrals, we have $C = C_1 + C_2 + C_3$. No matter how we break up the curve C_1 to set up a Riemann sum used in Definition 13.6.5 we can do the symmetric version on curve C_2. The only difference between the Riemann sums and their associated line integrals for C_1 and C_2 will be the sign on the output of f. This means that

$$\int_{C_1} f\,ds = -\int_{C_2} f\,ds$$

and thus

$$\begin{aligned}\int_C f\,ds &= \int_{C_1} f\,ds + \int_{C_2} f\,ds + \int_{C_3} f\,ds \\ &= -\int_{C_2} f\,ds + \int_{C_2} f\,ds + \int_{C_3} f\,ds \\ &= \int_{C_3} f\,ds.\end{aligned}$$

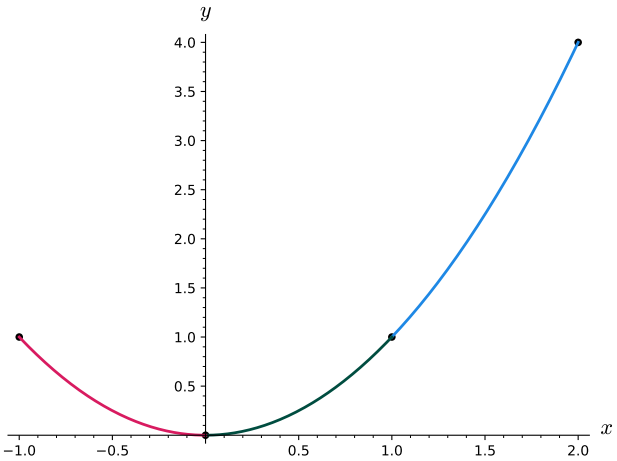


Figure 13.6.9 The curve C split into three regions with positive (blue and green) and negative (magenta) outputs of f

The argument for why $\int_{C_1} f\,ds = -\int_{C_2} f\,ds$ requires that both the points on the curves C_1 and C_2 and the output of f on those parts is symmetric. In Activity 13.6.3 you will see examples where having only one part of this symmetry is not sufficient to make this type of argument.

We have shown that $\int_C f\,ds = \int_{C_3} f\,ds$ and want to demonstrate why $\int_{C_3} f\,ds$ will be positive. Because the output of f is positive for all of the points on the curve C_3, all elements of the associated Riemann sum in Definition 13.6.5 will be positive (the length of the displacement vectors and corresponding arc lengths are always positive). Thus $\int_C f\,ds = \int_{C_3} f\,ds > 0$.

(b) We now want to consider whether $\int_C g\,ds$ will be positive, negative, or zero. This is a simpler argument than in the previous case because the output of g will be negative for all points in C (except for the origin). Each term in the Riemann sum used to define the scalar line integral will be the product of a negative value ($g(\vec{r}_i^*)$) and a positive value ($\|\vec{r}_i - \vec{r}_{i-1}\|$). Therefore the Riemann sums will be negative and the limit as you take more terms in this sum will also be negative. Thus, $\int_C g\,ds < 0$.

(c) Finally, we want to make an argument whether $\int_C h\,ds$ will be positive, negative, or zero. While we cannot lean on the intuitive nature of the where coordinates are positive or negative, we can make a very precise argument about the output of h for the points on C. While $g(x, y) = y - x^2$ will have a range that includes all real numbers, the output of g is ***always*** zero for the points on our curve C because for points on C, the y-coordinate is equal to the x-coordinate squared. This means that no matter how we break up our segments for the Riemann sum, the $h(\vec{r}_i^*)$ terms will always be zero. No matter how many terms are used the Riemann sum will be zero, and the limit of the Riemann sums will also be zero. Thus, $\int_C h\,ds = 0$.

▲

Before delving into the exact computation of line integrals of scalar-valued functions, here is an activity that allows you to practice arguments similar to Example 13.6.6.

Activity 13.6.2 In this activity, we will be making sense of scalar line integrals by examining a few common functions and justifying whether the scalar line integrals given are positive, negative, or zero. Let the functions f_1, f_2, f_3, and f_4 be defined as

- $f_1(x, y, z) = y$
- $f_2(x, y, z) = z$
- $f_3(x, y, z) = x^2$
- $f_4(x, y, z) = x - y$

(a) For each of the paths given below, sketch (in either 2D or 3D) the curve and label at least three points on the curve including the end points (if they exist).

(i) C_1 is the part of the unit circle in the xy-plane centered at the origin that is above the line $y = -x$.

(ii) C_2 is the part of the curve at the intersection of the cylinder given by $x^2 + y^2 = 1$ and the plane $z = x$ such that $y \geq -x$; You may want to consider the circle that is the intersection of $x^2 + y^2 = 1$ and $z = x$, then think about which half of this circle satisfies the inequality $y \geq -x$

(iii) C_3 is the part of the helix given by $\vec{r}(t) = \langle \cos(t), \sin(t), \frac{t}{2\pi} \rangle$ with $t \in [0, \pi]$

(b) For each of the functions f_1, f_2, and f_3 defined above, state whether $\int_{C_1} f_i \, ds$ is positive, negative, or zero. Be sure to justify your answer in terms of the function being integrated *and* the particulars of the curve of integration.

(c) For each of the functions f_1, f_2, and f_3, defined above, state whether $\int_{C_2} f_i \, ds$ is positive, negative, or zero. Be sure to justify your answer in terms of the function being integrated *and* the particulars of the curve of integration.

(d) For each of the functions f_1, f_2, and f_3, defined above, state whether $\int_{C_3} f_i \, ds$ is positive, negative, or zero. Be sure to justify your answer in terms of the function being integrated *and* the particulars of the curve of integration.

(e) For the function f_4, defined above, state each of the following integrals is positive, negative, or zero. Be sure to justify your answer in terms of the function being integrated *and* the particulars of the curve of integration. You should consider which parts of the curve being integrated will have positive/negative/zero output for the function f_4.

$$\int_{C_1} f_4 \, ds$$
$$\int_{C_2} f_4 \, ds$$
$$\int_{C_3} f_4 \, ds,$$

13.6.3 Using Parameterizations to Calculate Scalar Line Integrals

Definition 13.6.5 defined $\int_C f\,ds$ in terms of a limit of a Riemann sum which is often useful for understanding what is being measured and not very useful when it comes to efficiently calculating the value of a given integral. A scalar line integral is presented algebraically in terms of three variables because the curve is given in terms of points in three coordinates and the function to be integrated is dependent on those same coordinate values. Geometrically, the scalar line integral is a one dimensional problem because we only have one dimension to travel; namely, we can travel along the curve in steps of arc length. Remember that a parameterization of a curve in space is a description of how to travel through the points (given as three coordinates) of the curve in terms of a parameter (usually given as t.) Parameterizations are very useful converting the three-variable (x, y, z) algebra of a scalar line integral problem into a one dimensional integral. Once we have done that, we can use all of the tools of single-variable calculus to evaluate the scalar line integral. To understand this conversion, we will look at the details of the CCA using a parameterization.

Let's look at applying a parameterization for C given by a vector-valued function of one variable $\vec{r}(t) = \langle f(t), g(t), h(t)\rangle$ for t in some interval $[a, b]$ to Definition 13.6.5. Instead of thinking in terms of pieces of the curve C, the parameterization allows us to break the interval $[a, b]$ into pieces $a = t_0, t_1, \ldots, t_n = b$ where $t_i = a + i(\Delta t)$ and $\Delta t = \frac{b-a}{n}$. While these pieces will be equally spaced in terms of the parameter t, the corresponding points on the curve C given by $r(t_i)$ may not be equally spaced.

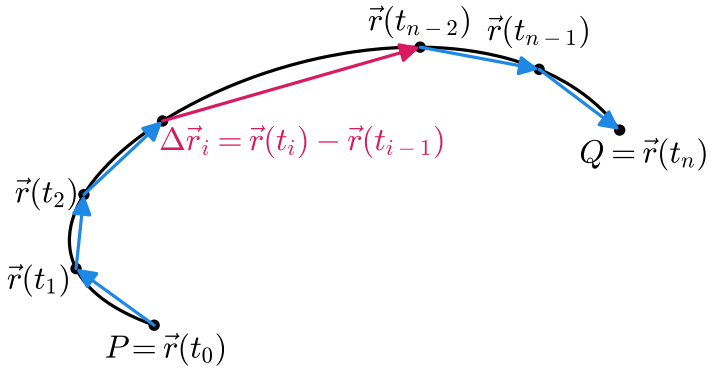


Figure 13.6.10 The curve C split into segments defined by equally spaced parameter values

To simplify the notation of our function evaluation, we will use the following:

$$f(\vec{r}_i^*) = f(\vec{r}(t_i^*)) = f(t_i^*)$$

where $t_{i-1} \le t_i^* \le t_i$. Remember that the parameterization will allow us to write all parts of the scalar line integral as a function of t, so we will simplify $\|\Delta\vec{r}_i\|$ as

$$\|\vec{r}_i - \vec{r}_{i-1}\| = \|\vec{r}(t_i) - \vec{r}(t_{i-1})\| = \frac{\|\vec{r}(t_i) - \vec{r}(t_{i-1})\|}{\Delta t}\Delta t$$

Applying our parameterization and corresponding points to Definition 13.6.5

gives

$$\begin{aligned}\int_C f\,ds &= \lim_{n\to\infty}\sum_{i=1}^{n} f(\vec{r}_i^*)\,\|\vec{r}_i - \vec{r}_{i-1}\| \\ &= \lim_{n\to\infty}\sum_{i=1}^{n} f(t_i^*)\frac{\|r(t_{i-1}+\Delta t) - r(t_{i-1})\|}{\Delta t}\Delta t.\end{aligned}$$

This Riemann sum corresponds to the definite integral of a scalar function of t, specifically $f(t)v(t)$ where $v(t) = \|\vec{r}'(t)\|$. We may think of $v(t)$ as the speed of the parameterization $\vec{r}$.

Theorem 13.6.11 Calculating Scalar Line Integrals with Parameterizations. *Let C be a curve in space parameterized by $r(t)$ for $a \le t \le b$. If $f(x, y, z)$ is a multivariable function that is continuous for a region around C, then*

$$\int_C f\,ds = \int_a^b f\left(\vec{r}\,(t)\right)\|\vec{r}\,'(t)\|\,dt.$$

With Theorem 13.6.11 established, we now consider a couple of examples that allow us to find the exact value of line integrals of scalar functions considered in Activity 13.6.2.

Example 13.6.12 Let us return to some of the problems from Activity 13.6.2. Specifically, let C_1 be part of the unit circle on the xy-plane that is centered at the origin and is above the line given by $y = -x$. We will also consider $f_1(x, y, z) = y$. We can parameterize C_1 by $\vec{r}(t) = \langle\cos(t), \sin(t), 0\rangle$ with $-\frac{\pi}{4} \le t \le \frac{3\pi}{4}$. Note here that $\|\vec{r}\,'(t)\| = \|\langle-\sin(t), \cos(t), 0\rangle\| = 1$. We can rewrite the value of f_1 along C_1 using the parameterization as $f_1(x(t), y(t), z(t)) = f_1(\cos(t), \sin(t), 0) = \sin(t)$. Applying Theorem 13.6.11, we have

$$\int_{C_1} f_1\,ds = \int_{-\frac{\pi}{4}}^{\frac{3\pi}{4}} \sin(t)(1)\,dt = -\cos\left(\frac{3\pi}{4}\right) - \left(-\cos\left(-\frac{\pi}{4}\right)\right) = \sqrt{2}.$$

This result should make sense from your earlier work in part 13.6.2.b. Specifically, we can break the line integral of f_1 along C_1 into three parts according to the plot shown in Figure 13.6.13.

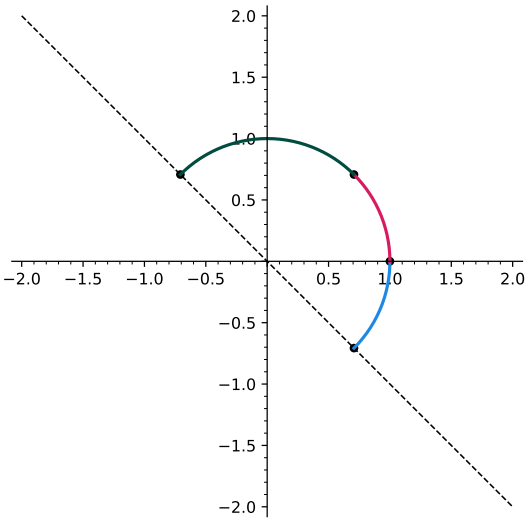


Figure 13.6.13 A subdivided plot of C_1

The scalar line integral on the blue and the magenta paths will cancel each

other out exactly since the paths are symmetric about x-axis and the output of f_1 will be opposite in sign on the blue and magenta portions. This means that $\int_{C_1} f_1\,ds$ will be the same as the line integral of $f_1(x,y,z) = y$ on the green path. Since the output of f_1 is positive on the green path, our result for the scalar line integral should be positive. ▲

When working through Activity 13.6.2, you may have found the line integrals involving $f_3(x,y,z) = x - y$ to be more challenging to reason through without computational tools. We next work through the details of evaluating the line integral of this function along the same curve as in the previous example.

Example 13.6.14 Let C_1 be part of the unit circle on the xy-plane that is centered at the origin and is above the line given by $y = -x$. We will also consider $f_4(x,y,z) = x - y$. We can parameterize C_1 by $\vec{r}(t) = \langle \cos(t), \sin(t), 0\rangle$ with $-\frac{\pi}{4} \leq t \leq \frac{3\pi}{4}]$. Note here that $\|\vec{r}'(t)\| = \|\langle -\sin(t), \cos(t), 0\rangle\| = 1$. We can write the value of f_4 along C_1 using the parameterization as $f_4(x(t), y(t), z(t)) = f_4(\cos(t), \sin(t), 0) = \cos(t) - \sin(t)$. Applying Theorem 13.6.11, we have

$$\begin{aligned}\int_{C_1} f_4\,ds &= \int_{-\frac{\pi}{4}}^{\frac{3\pi}{4}} (\cos(t) - \sin(t))\,(1)dt \\ &= \sin(\frac{3\pi}{4}) - \cos(\frac{3\pi}{4}) - \left(\sin(-\frac{\pi}{4}) - \cos(-\frac{\pi}{4})\right) = 0\end{aligned}$$

This result should make sense from your earlier work in part 13.6.2.b. Specifically, we can break the line integral of f_4 into two parts according to the plot in Figure 13.6.15.

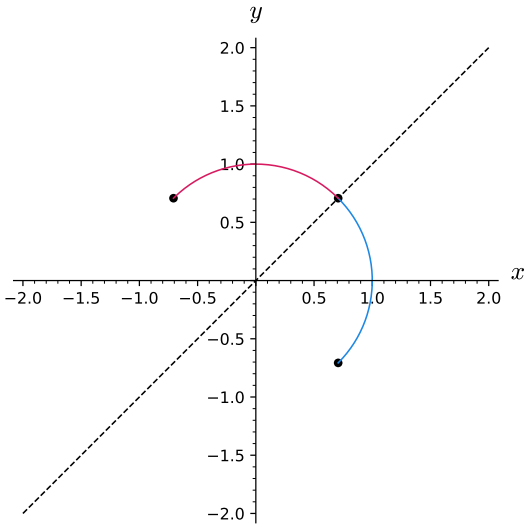


Figure 13.6.15 A subdivided plot of C_1

The scalar line integral on the blue and the magenta paths will cancel each other out exactly since the paths are symmetric and the output of f_4 will be opposite in sign on the blue and magenta. Thus the line integral of f_4 over C_1 will be zero. ▲

Activity 13.6.3 In this activity, we will examine why we must be careful when using symmetry to make arguments about scalar line integral. Let C_1 and C_2 be the paths shown in Figure 13.6.16. We will consider the function $f(x,y) = x$ for this activity.

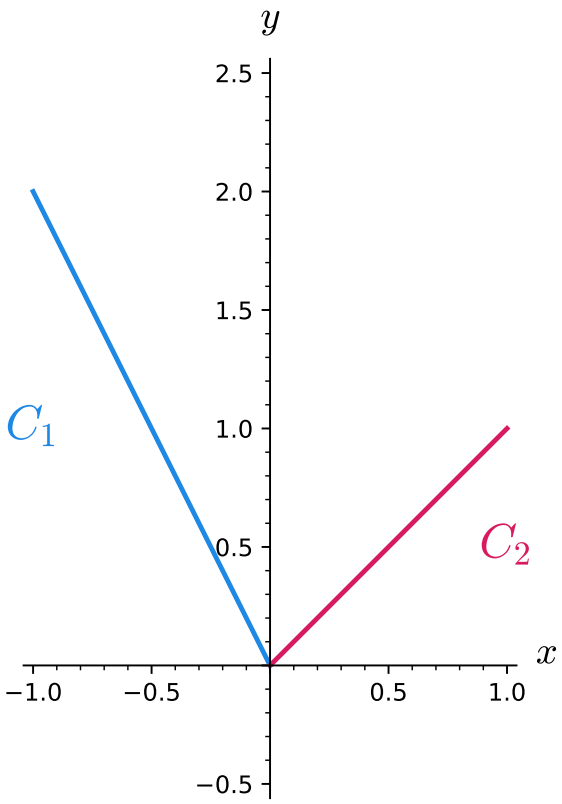


Figure 13.6.16 A plot of paths C_1 and C_2

(a) Parameterize C_1 and C_2 as $\vec{r}_1(t)$ and $\vec{r}_2(t)$. (It is fine to have $0 \leq t \leq 1$ for both of your parameterizations.)

(b) Use Theorem 13.6.11 to compute $\int_{C_1} f\,ds$ and $\int_{C_2} f\,ds$.

(c) As with line integrals of vector fields, we have that $\int_{C_1+C_2} f\,ds = \int_{C_1} f\,ds + \int_{C_2} f\,ds$. Use this property to compute $\int_{C_1+C_2} f\,ds$.

In Activity 13.6.3, it may have been tempting to note that output of the function being integrated is simply the x-coordinate of the points on the curve to try to make an argument by symmetry. Since we are adding up the x-values of the points on C_1 and C_2 and the x-coordinates on C_1 go from -1 to 0 and the x-coordinates on C_2 go from 0 to 1, you may have wanted to argue that the values of $\int_{C_1} x\,ds$ and $\int_{C_2} x\,ds$ will cancel out to zero. However, as you found, this is not correct because the scalar line integrals are not just adding up the output of our scalar function. The Riemann sum used in the scalar line integral is adding up the product of the scalar function's output with the displacement of the segment used. Because the displacement used in C_1 and C_2 will not be symmetric, we cannot make the geometric cancellation argument as in part 13.6.6.a.

13.6.4 Properties of Scalar Line Integrals

Before stating some useful properties of scalar line integrals, we will recall some convenient notation from Properties of Line Integrals. If C_1 and C_2 are oriented curves, with C_1 from a point P to a point Q and C_2 from Q to a point R, we denote by $C_1 + C_2$ the oriented curve from P to R that follows C_1 to Q and then continues along C_2 to R. Also, if C is an oriented curve, $-C$ denotes the same curve but with the opposite orientation. The list below summarizes some other properties of line integrals, each of which has a familiar in definite integrals.

Properties of Scalar Line Integrals.

For a constant scalar k, scalar valued functions f and g, and oriented curves C, C_1, and C_2, the following properties hold:

a. $\displaystyle\int_C (kf)\,ds = k\int_C f\,ds$

b. $\displaystyle\int_C (f+g)\,ds = \int_C f\,ds + \int_C g\,ds$

c. $\displaystyle\int_{-C} f\,ds = \int_C f\,ds$

d. $\displaystyle\int_{C_1+C_2} f\,ds = \int_{C_1} f\,ds + \int_{C_2} f\,ds$

The biggest difference between Properties of Line Integrals and Properties of Scalar Line Integrals is part c. The orientation of the curve does not change the value of the scalar line integral. Activity 13.6.4 will have you make sense of these properties for scalar line integrals.

Activity 13.6.4 Explaining Properties of Scalar Line Integrals. In this activity, we will be explaining each of the Properties from Properties of Scalar Line Integrals in the context of our copium mining analogy from Preview Activity 13.6.1. Remember that the curve in our scalar line integral corresponds to the path the mining rig will take and the function in the scalar line integral measures the density of copium at that point on the surface.

(a) Explain in your own words what $\displaystyle\int_C f\,ds$ means in the copium analogy and what exactly would be measured by this scalar line integral.

(b) Explain in your own words what $\displaystyle\int_C (kf)\,ds = k\int_C f\,ds$ means in the copium analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

(c) Explain in your own words what $\displaystyle\int_C (f+g)\,ds = \int_C f\,ds + \int_C g\,ds$ means in the copium analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy. You may also find it helpful to consider f and g to measure density of different materials.

(d) Explain in your own words what $\displaystyle\int_{-C} f\,ds = \int_C f\,ds$ means in the copium analogy. It may be helpful to describe each side of the equation separately

and say why they are equal in the analogy.

(e) Explain in your own words what $\int_{C_1+C_2} f\,ds = \int_{C_1} f\,ds + \int_{C_2} f\,ds$ means in the copium analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

13.6.5 Visualizations of Scalar Line Integrals as Area Under a Curve

We will spend the last part of this section talking about a way to try to visualize the scalar line integral as an area under a curve, much as we visualized integrals when we first encountered them. Let's return to our copium analogy from Preview Activity 13.6.1. In particular, we can look the left side of the mining area.

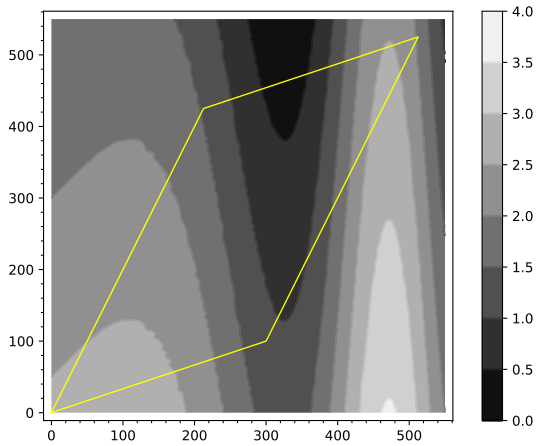


Figure 13.6.17 A plot of land with density of copium deposits

We could visualize the linear density of copium along the left side of the area using a plot like Figure 13.6.18.

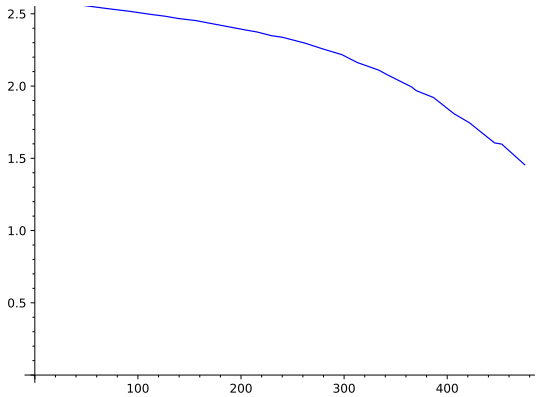
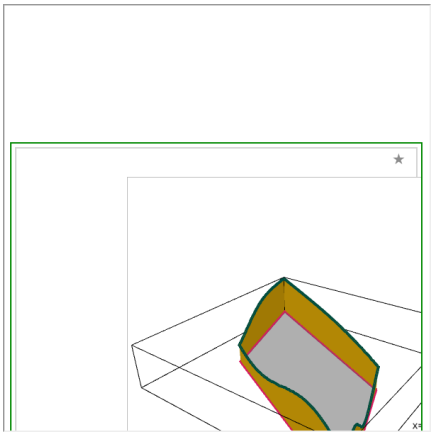


Figure 13.6.18 A plot of the copium density on the left side of the mine shown above

In Figure 13.6.18, the horizontal axis gives the distance traveled along the left side of Figure 13.6.1. Because this is a stright path, we could plot the density above the path of the copium mining plot. In fact, we could plot the density above the plot for all of the sides of the mining plot.

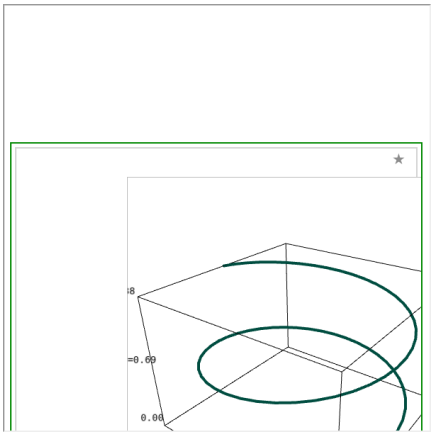


Standalone
In Context
Embed

Figure 13.6.19 A three-dimensional plot of the copium Density plotted for the edges of the mining area

Figure 13.6.19 shows the copium mine plot (in gray) and the paths that are the boundary of the plot in magenta. The curve in green shows the copium Density at each point on the boundary of the mine plot. The area in yellow would be the scalar line integral for the path that is the boundary of the mine plot. In particular, the area in yellow would give the total copium mined from driving our mining machine around the boundary of the mine plot.

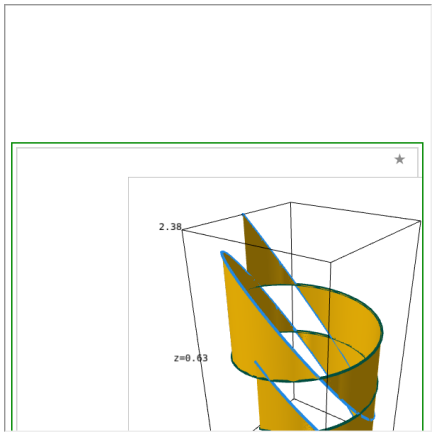
Because the curve we are looking at in Figure 13.6.19 involves straight lines and simple heights, there is no confusion when looking at this plot and using an area under the curve analogy. However, what if we looked at the scalar line integral of a function like $f(x, y, z) = x - y$ along the helix given below?



Standalone
In Context
Embed

Figure 13.6.20 A three-dimensional plot of a helix

Suppose now that above the points of our green helix we try to plot a second blue curve where the position of the point on the blue curve is $f(x, y, z) = x - y$ units above or below the position of the point on the helix. This would be the analogous idea to what we did in Figure 13.6.19, but for a three-dimensional curve. Here we run into issues, however, as our area might intersect other parts of the curve. The plot below shows the confusing plot we would have if we looked at as the height above our curve in blue.



Standalone
In Context
Embed

Figure 13.6.21 A three-dimensional plot of a helix with height given by $f(x, y, z) = x - y$

Summary

- The scalar line integral, denoted by $\int_C f\, ds$, measures the accumulation of the output of f over the points on the curve C.
- Parameterizing the curve used in a scalar line integral allows you to compute the scalar line integral as a definite integral of one variable.
- Scalar line integrals can be split into pieces of the curve or along linear combinations of the scalar valued function being integrated.

13.6.6 Exercises

1. Find the line integral with respect to arc length $\int_C (2x+8y)ds$, where C is the line segment in the xy-plane with endpoints $P = (7, 0)$ and $Q = (0, 8)$.

(a) Find a vector parametric equation $\vec{r}(t)$ for the line segment C so that points P and Q correspond to $t = 0$ and $t = 1$, respectively.

$\vec{r}(t) =$ _____

(b) Using the parametrization in part (a), the line integral with respect to arc length is

$$\int_C (2x + 8y)ds = \int_a^b _____\, dt$$

with limits of integration $a =$ _____ and $b =$ _____

(c) Evaluate the line integral with respect to arc length in part (b).

$\int_C (2x + 8y)ds =$ _____

2. If C is the part of the circle $\left(\frac{x}{6}\right)^2 + \left(\frac{y}{6}\right)^2 = 1$ in the first quadrant, find the following line integral with respect to arc length.

$\int_C (5x - 8y)ds =$ _____

3. Evaluate the line integral $\int_C x^5 z\ ds$, where C is the line segment from $(0, 1, 2)$ to $(8, 3, 7)$.

4. Evaluate the line integral $\int_C 7xy^4\ ds$, where C is the right half of the

circle $x^2 + y^2 = 4$.

5. Consider a wire in the shape of a helix $\mathbf{r}(t) = 3\cos t\mathbf{i} + 3\sin t\mathbf{j} + 4t\mathbf{k}, 0 \leq t \leq 2\pi$ with constant density function $\rho(x, y, z) = 1$.

A. Determine the mass of the wire: _____

B. Determine the coordinates of the center of mass: (_____, _____, _____)

C. Determine the moment of inertia about the z-axis: _____

6. Compute the total mass of a wire bent in a quarter circle with parametric equations: $x = 4\cos t$, $y = 4\sin t$, $0 \leq t \leq \frac{\pi}{2}$ and density function $\rho(x, y) = x^2 + y^2$.

7. Let $f(x, y, z) = x + yz$ and let $\mathcal{C}$ be the line segment from $P = (0, 0, 0)$ to $(6, 7, 6)$.

(a) Calculate $f(\mathbf{c}(t))$ and $ds = \|\mathbf{c}'(t)\|\, dt$ for the parametrization $\mathbf{c}(t) = (6t, 7t, 6t)$ for $0 \leq t \leq 1$.

(b) Evaluate $\int_{\mathcal{C}} f(x, y, z)\, ds$.

(a) $f(\mathbf{c}(t)) =$ _____ , $ds =$ _____ dt

(b) $\int_{\mathcal{C}} f(x, y, z)\, ds =$ _____

8. Calculate the integral of $f(x, y, z) = 6x^2 + 6y^2 + z^5$ over the curve $\mathbf{c}(t) = (\cos t, \sin t, t)$ for $0 \leq t \leq \pi$.

$\int_{\mathcal{C}}(6x^2 + 6y^2 + z^5)\, ds =$ _____

9. Compute the line integral of the scalar function $f(x, y) = \sqrt{1 + 9xy}$ over the curve $y = x^3$ for $0 \leq x \leq 7$

$\int_{\mathcal{C}} f(x, y)\, ds =$ _____

10. Compute the line integral of the scalar function $f(x, y, z) = 2x^2 + 8z$ over the curve $\mathbf{c}(t) = (e^t, t^2, t), \quad 0 \leq t \leq 8$

$\int_{\mathcal{C}} f(x, y, z)\, ds =$ _____

11. Calculate the total mass of a circular piece of wire of radius 9 cm centered at the origin whose mass density is $\rho(x, y) = x^2$ g/cm.

Answer : _____g

12. Consider a wire in the shape of a helix $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k}, 0 \leq t \leq 2\pi$ with constant density function $\rho(x, y, z) = 1$.

A. Determine the mass of the wire: _____

B. Determine the coordinates of the center of mass: (_____, _____, _____)

C. Determine the moment of inertia about the z-axis: _____

Note: If a wire with linear density $\rho(x, y, z)$ lies along a space curve C, its ***moment of inertia*** about the z-axis is defined by $I_z = \int_C (x^2 + y^2)\rho(x, y, z)ds$.

13. Compute

$$\int_C f ds$$

where $f(x, y) = \frac{y^3}{x^7}$ and C is the curve given by $y = \frac{1}{4}x^4$ for $1 \leq x \leq 2$.

14. Let C be the path given below from P to Q with pieces C_1, C_2, and C_3 as labeled. Let f be a scalar-valued function such that $\int_C f\, ds = 13$, $\int_{C_1} f\, ds = 5$,and $\int_{C_3} f\, ds = 9$.

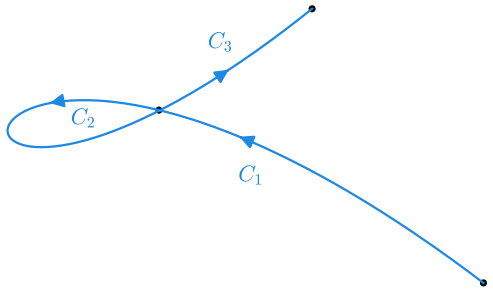


Figure 13.6.22 An oriented path broken into three parts

Find the following:

(a) $\displaystyle\int_{-C_3} f\,ds$

(b) $\displaystyle\int_{C_2} f\,ds$

(c) $\displaystyle\int_{-C_1-C_3} f\,ds$

15. Calculate the following line integral where C is the path on $x = y$ with $-1 \geq y \geq 2$:

$$\int_C (y^2 - 2x)\,ds$$

13.7 The Divergence of a Vector Field

Motivating Questions

- How can you measure where a vector field's strength is increasing or decreasing?
- What does the divergence of a vector field measure and how can you visually estimate whether the divergence of a vector field is positive or negative?

13.7.1 Introduction

As we saw in Section 13.2, there are many physical and theoretical representations for vector fields. A natural question is "Where exactly is the vector field created?" With the vector field in Figure 13.7.1, imagine sketching a curve that follows the direction of the vector field by treating the vectors in the vector field as tangent vectors to your curve. No matter where you start, you should observe that the vector field decreases in strength as you move along the flow. We wish to understand (as a function of position), how much of the vector field is created (or destroyed) at a given location.

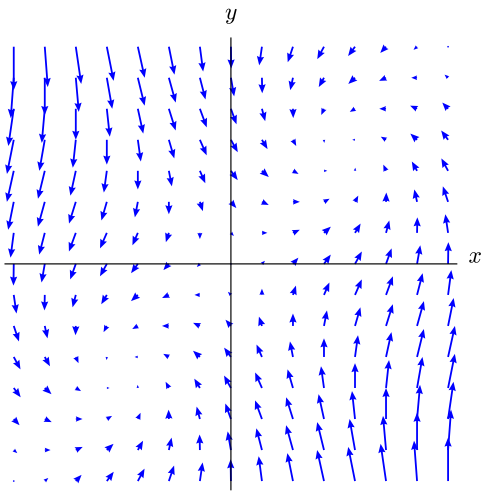


Figure 13.7.1 A vector field with changing strength

Preview Activity 13.7.1 In this preview activity, we will look at several two-dimensional vector fields and try to assess when the vector field has increased or decreased in strength over a given region. We begin with graphs of the three vector fields, $\vec{F}$, $\vec{G}$, and $\vec{H}$. Parts a, b, and c ask you to answer the same three questions about the vector field and square illustrated in each of the figures. Part d asks you to think further about the third vector field.

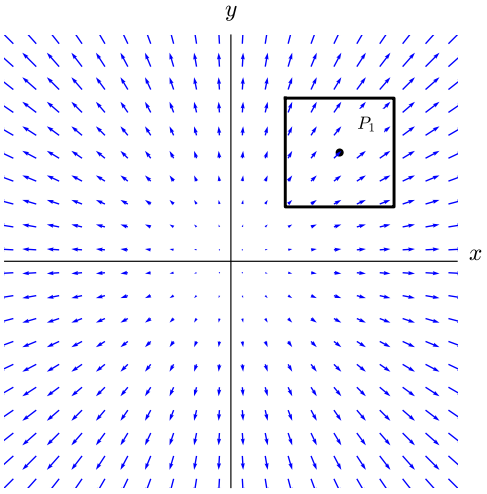


Figure 13.7.2 Vector Field $\vec{F}$

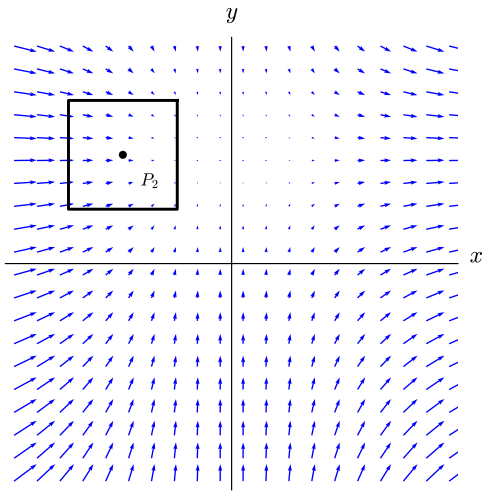


Figure 13.7.3 Vector Field $\vec{G}$

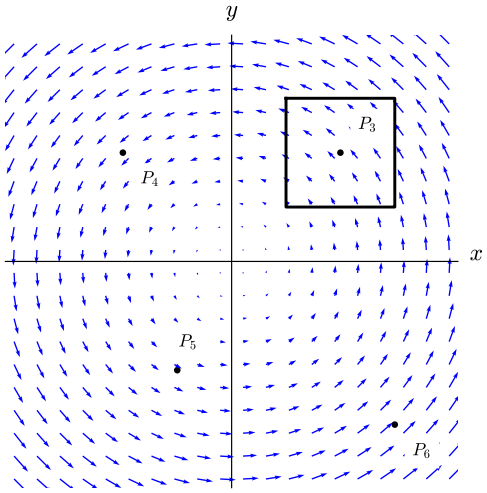


Figure 13.7.4 Vector Field $\vec{H}$

(a) For each of the vector fields $\vec{F}$, $\vec{G}$, and $\vec{H}$ and the square centered on P_1, P_2, and P_3 (respectively), which statement do you think best applies?

i More of the vector field is going into the square than going out.

ii Less of the vector field is going into the square than going out.

iii The same amount of the vector field is going into the square as is going out.

(b) For each of the vector fields $\vec{F}$, $\vec{G}$, and $\vec{H}$ (and corresponding square), does your answer to part a suggest that the vector field is being created, destroyed, or is unchanging in strength inside the square? Write a sentence to explain your thinking for each vector field.

(c) Would the answer to parts a or b change if you used a smaller square centered on P_1, P_2, and P_3 for the corresponding vector fields? Write a sentence to explain your thinking for each vector field.

(d) Thinking now only about the vector field $\vec{H}$, would your answers to parts a, b, or c change if you considered squares around points P_4, P_5, or P_6? Write a couple of sentences to explain your thinking.

13.7.2 Definition of the Divergence of a Vector Field

We begin this subsection by stating a definition that captures analytically the ideas you reasoned about geometrically in Preview Activity 13.7.1. After the statement of the definition, we discuss what it means.

Definition 13.7.5 The **divergence of a vector field**

$$\vec{F}(x,y) = \langle F_1(x,y,z), F_2(x,y,z), F_3(x,y,z)\rangle$$

is given by

$$\text{div}(\vec{F}) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}.$$

In two dimensions, the definition is analogous by omitting the third term. ◆

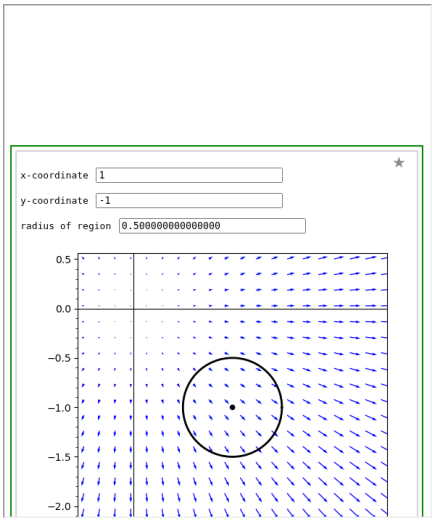
Alternative Notation for Divergence. In other sources you may see the divergence written using a dot product as $\text{div}(\vec{F}) = \nabla \cdot \vec{F}$. This notation is very compact and works well with the understanding that the del operator $\nabla = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\rangle$ is a function that operates on other functions. However, this notation can also be confusing because of its emphasis on computation rather than conceptual understanding. In this text, we will not generally write the divergence using the del operator.

The divergence of a vector field is a scalar measurement at a point that measures how the strength of the vector field is changing as we look in small neighborhoods around our point. While Subsection 13.7.3 will show the details for how the divergence is defined, we can make the following qualitative argument: If we are looking at how the strength of the vector field is changing in a small neighborhood of the point (a, b), then we only need to look how fast the horizontal component is changing horizontally and how the vertical component is changing vertically. We make this measurement of changing strength of the vector field by measuring how much of a vector field flows into versus out of a small neighborhood of our point (as was done in Preview Activity 13.7.1). When looking at a small neighborhood of the point (as shown by the square in Figure 13.7.15), only the change of the horizontal component of the vector field contributes to flow in or out on the sides and only the change in the vertical component contributes to flow in or out on the top and bottom.

Example 13.7.6 Let $\vec{F} = \langle x, y\rangle$. Then

$$\text{div}(\vec{F}) = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) = 1 + 1 = 2.$$

From our conceptual description above, $\text{div}(\vec{F})$ being positive means that our vector field is increasing in strength. Since the divergence of $\vec{F}$ does not have a dependence on the input point, that means the vector field is increasing in strength, regardless of which point we consider. In the interactive element below, you can change the point you would like plotted and the size of the region around the point. You should see that regardless of what point you select or how small you make the region around the point, there will be more of the vector field flowing out of the region than in.



Standalone
In Context
Embed

Figure 13.7.7 An interactive plot of $\vec{F} = \langle x, y \rangle$ with point (a, b) and region plotted

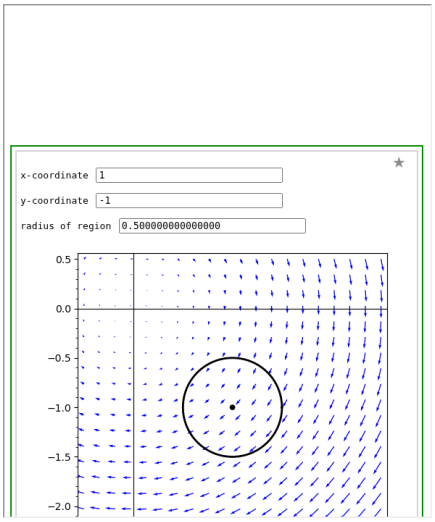
▲

Example 13.7.8 Let $\vec{G} = \langle y, -x \rangle$. We can compute that

$$\text{div}(\vec{G}) = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(-x) = 0.$$

From our conceptual description of divergence, $\text{div}(\vec{G})$ being zero means that our vector field is not changing in strength. Since the divergence of $\vec{G}$ does not have a dependence on the input point, that means the vector field is not changing in strength, regardless of which point we consider.

In the interactive element below, you can change the point you would like plotted and the size of the region around the point. You should see that regardless of what point you select or how small you make the region around the point, there will be exactly as much of the vector field flowing into the region as is flowing out (on the other side).



Standalone
In Context
Embed

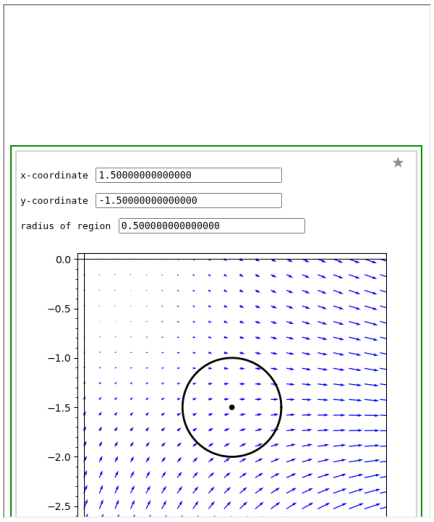
Figure 13.7.9 An interactive plot of $\vec{G} = \langle y, -x \rangle$ with point (a, b) and region plotted

▲

Example 13.7.10 Let $\vec{H} = \langle x^2 - y, y^2 - x \rangle$, which means that

$$\text{div}(\vec{H}) = \frac{\partial}{\partial x}(x^2 - y) + \frac{\partial}{\partial y}(y^2 - x) = 2x + 2y.$$

In contrast to the previous examples, you can see that the value of divergence of $\vec{H}$ will depend on the input point chosen. For instance, $\text{div}(\vec{H})$ will be 4 at the point $(1, 1)$. This means that for a small region around $(1, 1)$ there should be more of the vector field flowing out of the region than into it. Use the interactive element below to verify this.



Standalone
In Context
Embed

Figure 13.7.11 An interactive plot of $\vec{H} = \langle x^2 - y, y^2 - y \rangle$ with point (a, b) and region plotted

The divergence of $\vec{H}$ at the point $(-1, 1)$ will be zero. If you look at this point in Figure 13.7.11, you will see that there is exactly as much of the vector field flowing into the region around $(-1, 1)$ as is flowing out.

The divergence of $\vec{H}$ at the point $(-2, 0)$ is -4. This should mean that there is more of the vector field flowing into the region around $(-2, 0)$ than is flowing out. You should be able to see how difference regions of the xy-plane will have different values for $\text{div}(\vec{H})$. ▲

Another way to "see" divergence on a vector field plot is to look at what happens to the magnitude of vectors as you move along the flow of the vector field. If the vector field is increasing in magnitude as you move along the flow of a vector field, then the divergence is positive. If the vector field is decreasing in magnitude as you move along the flow of a vector field, then the divergence is negative. If the vector field does not change in magnitude as you move along the flow of the vector field, then the divergence is zero. Also, remember that the divergence of a vector field is often a variable quantity and will change depending on location. The next activity asks you to graphically examine the divergence of three vector fields.

Activity 13.7.2 Graphical Representations of Divergence.

(a) For this part of the activity, consider the vector field $\vec{F}$ shown in Figure 13.7.12.

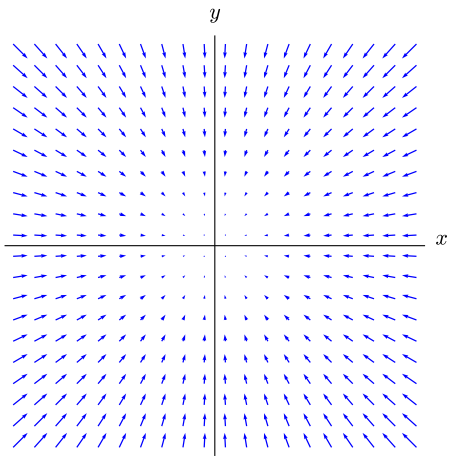


Figure 13.7.12 Vector field $\vec{F}$

(i) Draw a circle in the first quadrant of the vector field $\vec{F}$ depicted in Figure 13.7.12. Based on the flow of the vector field into or out of the circle, do you think the vector field is increasing in strength, decreasing in strength, or not changing in overall strength in the first quadrant?

(ii) As you move along the flow of the vector field in the first quadrant of Figure 13.7.12, does your vector field increase in magnitude, decrease in magnitude, or have constant magnitude?

(iii) Draw a circle in each of quadrants II, III, and IV. Based on the flow of the vector field into or out of your circles, do you think the vector field is increasing in strength, decreasing in strength, or not changing in overall strength in quadrants III, II, and IV?

(iv) As you move along the flow of the vector field in the third quadrant of Figure 13.7.12, does your vector field increase in magnitude, decrease in magnitude, or have constant magnitude.

(v) Based on your arguments above, describe why the divergence of $\vec{F}$ is negative for all points in the xy-plane.

(b) Look at the plot of the vector field $\vec{G}$ in Figure 13.7.13 and state whether you think the vector field is increasing in strength, decreasing in strength, or not changing in overall strength in each of the four quadrants. You can make your argument in terms of the change in magnitude along the flow of the vector field or in terms of the net flow into or out of a small region on the plane. You may need to make separate arguments for each of the four quadrants.

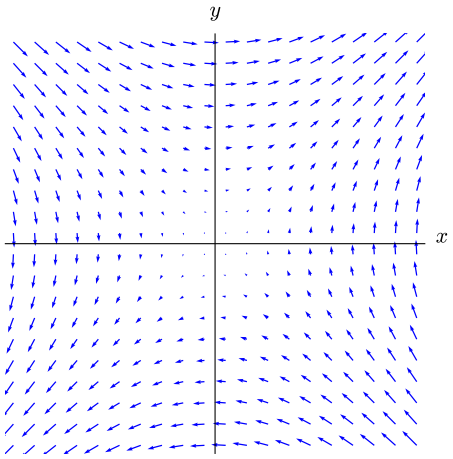


Figure 13.7.13 Vector field $\vec{G}$

(c) Look at the plot of the vector field $\vec{H}$ in Figure 13.7.14 below and state whether you think the vector field is increasing in strength, decreasing in strength, or not changing in overall strength in each of the four quadrants. You can make your argument in terms of the change in magnitude along the flow of the vector field or in terms of the net flow into or out of a small region on the plane. You may need to make separate arguments for each of the four quadrants.

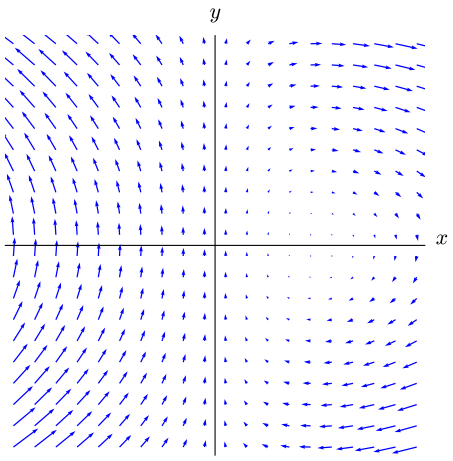


Figure 13.7.14 Vector field $\vec{H}$

The next activity of this section asks you to do some algebraic calculations of divergence using Definition 13.7.5.

Activity 13.7.3

(a) Calculate the divergence of the vector fields given below.

- $\vec{F}(x, y) = \langle -x, -y \rangle$
- $\vec{G}(x, y) = \langle y, x \rangle$
- $\vec{H}(x, y) = \langle xy, 1 - x \rangle$

(b) Explain how your answers to the questions in Activity 13.7.2 can be explained by using your results from part a of this activity.

13.7.3 Measuring the Change in Strength of a Vector Field

In this subsection, we examine the details of how to measure the density of the "creation" or "destruction" of the vector field using a classic calculus approach. Specifically, we will measure how the strength of the vector field changes in a region around a point. Next, using a limit, we examine what happens to our measurement as we shrink the region. Because vector fields change in a continuous fashion, the vector fields don't actually change *at a single point*. Rather, we will measure the density for the change in strength of the vector field using a limit.

We will develop all of our measurements in a two dimensional setting for now. However, our arguments can be applied to three (or more) dimensions. We start in the same fashion as in Preview Activity 13.7.1. Namely, we will look at how much of the vector field is going into or out of a square centered at a point (a, b). For this development, we will consider a two-dimensional vector field given by $\vec{F}(x, y) = \langle F_1(x, y), F_2(x, y)\rangle$.

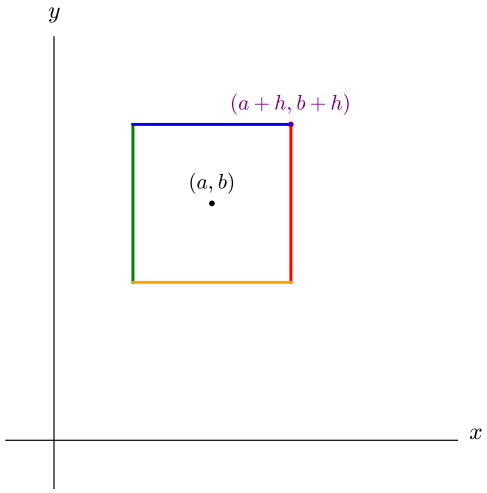


Figure 13.7.15 A square around the point (a, b)

We can parametrize the top edge of the box by $\vec{r}_{\text{top}}(t) = \langle a + t, b + h\rangle$ with $-h \le t \le h$. Similarly, the bottom, right, and left can be parametrized by

$$\begin{aligned} \vec{r}_{\text{bottom}}(t) &= \langle a + t, b - h\rangle \\ \vec{r}_{\text{right}}(t) &= \langle a + h, b + t\rangle \\ \vec{r}_{\text{left}}(t) &= \langle a - h, b + t\rangle \end{aligned}$$

all of which use parameter values in $-h \le t \le h$.

The amount of the vector field $\vec{F}$ that is created inside the square around the point (a, b) can be measured by the net amount of the vector field coming into or going out of the square. The amount of vector flow that goes through each of the boundary segments can be measured by looking at just the orthogonal component of the vector field on each particular segment. For instance, on the top segment, the vertical component F_2 determines how much of the vector field goes in or out of the square. Integrating just the vertical component F_2 of the vector field $\vec{F}$ over the points on the top segment of our square will therefore measure how much of the vector field goes through the top of the square.

The same argument applies to the bottom edge of our square. Similarly, if we want to measure how much of $\vec{F}$ goes through either the left or right side

of the square, we need to integrate the horizontal component F_1. Hence, the net flow N of the vector field into or out of the square will be given by

$$N = \int_{-h}^{h} F_2(a+t, b+h)dt - \int_{-h}^{h} F_2(a+t, b-h)dt + \int_{-h}^{h} F_1(a+h, b+t)dt - \int_{-h}^{h} F_1(a-h, b+t)dt.$$

Notice that the integrals corresponding to the left and bottom segments are subtracted because we need to pay attention to the orientation of the vector field relative to the square. A positive vertical component of the vector field (F_2) will correspond to flow out on the top of the square but will correspond to the vector field flowing into the square on the bottom. In the integrals above, we are counting the flow out of the square as positive and the flow in as negative.

We are measuring the net flow through the square as a scalar quantity. By decreasing h, we can look at what happens to our amount of flow out of the square as we shrink to the point (a, b). In order for this to make sense across different size of squares, we will change what we are measuring to be a density argument by calculating flow in (or out) per unit area. This will allow us to compare our net flow calculations across squares with different areas. In other words, we want to consider what happens to

$$\frac{1}{(2h)^2}\left(\int_{-h}^{h} F_2(a+t, b+h)dt - \int_{-h}^{h} F_2(a+t, b-h)dt + \int_{-h}^{h} F_1(a+h, b+t)dt - \int_{-h}^{h} F_1(a-h, b+t)dt\right)$$

as h goes to zero.

Before we compute our limit, we will take a moment to simplify our integrals in order to make the limit easier to evaluate. Recall that in single-variable calculus, we defined the average value of a function f on an interval $[a, b]$ to be

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

It turns out that we can say something stronger: there will be a value t^* in the interval $[a, b]$ so that

$$f(t^*)(b-a) = \int_a^b f(t)\, dt$$

when f is continuous on $[a, b]$. (You can think of the left-hand side as being the area of a rectangle with the interval $[a, b]$ as its base.) We use this fact, sometimes called the Mean Value Theorem for Integrals, to simplify here. Applying the Mean Value Theorem for Integrals to the first integral gives

$$\int_{-h}^{h} F_2(a+t, b+h)dt = (2h)\ \ F_2(t_1^*, b+h),$$

where t_1^* is some value in the interval $(a-h, a+h)$. Applying the Mean Value Theorem for Integrals to each of the other integrals allows us to simplify the expression above for the net flow N to be

$$(2h)\left(F_2(t_1^*, b+h) - F_2(t_2^*, b-h) + F_1(a+h, t_3^*) - F_1(a-h, t_1^*)\right),$$

where t_1^* and t_2^* are values in $(a-h, a+h)$ and t_3^* and t_4^* are values in $(b-h, b+h)$. Thus our flow density can be measured by looking at the limit as $h \to 0$ of the net flow (in or out) over the square divided by the area of the square.

$$\begin{aligned}\text{Flow Density}(a,b) &= \lim_{h\to 0} \frac{\text{net flow}}{\text{area}} \\ &= \lim_{h\to 0} \frac{(2h)\left(F_2(t_1^*, b+h) - F_2(t_2^*, b-h) + F_1(a+h, t_3^*) - F_1(a-h, t_1^*)\right)}{4h^2}.\end{aligned}$$

To simplify the limit further, we will reorganize our limit expression. Specifically, we collect the F_2 and F_1 terms separately. This gives

$$\begin{aligned}&\text{Flow Density}(a,b) = \\ &\lim_{h\to 0}\left[\frac{F_1(a+h, t_3^*) - F_1(a-h, t_1^*)}{2h} + \frac{F_2(t_1^*, b+h) - F_2(t_2^*, b-h)}{2h}\right].\end{aligned}$$

Recall the central difference method of computing derivatives from Section 1.5.2 and as $h \to 0$, the numbers t_1^*, t_2^* must go to a and t_3^*, t_4^* must go to b. Therefore, after evaluating our limit, the flow density is

$$\text{Flow Density}(a,b) = \frac{\partial F_1}{\partial x}(a,b) + \frac{\partial F_2}{\partial y}(a,b).$$

While this simplification may seem a bit amazing and magical, our conceptual steps should help us make sense of the result. If we are looking at how the strength of the vector field is changing in a small neighborhood of the point (a, b) then we only need to look how fast the horizontal component is changing horizontally and how the vertical component is changing vertically. How the horizontal component changes over small steps in the vertical direction will give us information about how the direction of the vector field changes but not how the strength of the vector field is changing.

The arguments we made about measuring how much of the vector field flows into or out of a square has straightforward generalization to three (or more) dimensions. However, doing so requires a method for measuring how much of a vector field flows through a surface. This will be the subject of Section 13.11.

Summary

- The divergence of a vector field $\vec{F}(x,y) = \langle F_1(x,y), F_2(x,y)\rangle$ is computed as
$$\text{div}(\vec{F}) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y}$$
In three dimensions, the divergence of the vector field $\vec{G}(x,y) = \langle G_1(x,y,z), G_2(x,y,z), G_3(x,y,z)\rangle$ is computed as
$$\text{div}(\vec{G}) = \frac{\partial G_1}{\partial x} + \frac{\partial G_2}{\partial y} + \frac{\partial G_3}{\partial z}$$

- The divergence of a vector field measures the density of change in the strength of the vector field. In other words, the divergence measures the instantaneous rate of change in the strength of the vector field along the direction of flow.

- The accumulation of the divergence over a region of space will measure the net amount of the vector field that exits (versus enters) the region.

- The key ideas when interpreting divergence are:
  - A positive divergence means that the vector field is growing in strength.
  - A negative divergence means that the vector field is decreasing in strength.
  - A zero divergence means that the vector field is not changing in strength.

13.7.4 Exercises

There are several additional exercises involving the calculation and interpretation of the divergence of a vector field in Section 13.8. In those problems, the output of divergence and curl are compared and contrasted.

1. Find the divergence of each of the following vector fields at all points where they are defined.

(a) $\operatorname{div}\left(\left(2x^2-\sin(xz)\right)\boldsymbol{i}+4\boldsymbol{j}-\sin(xz)\,\boldsymbol{k}\right)$ = _____

(b) $\operatorname{div}\left(2e^{xy}\boldsymbol{i}+4\cos(xy)\,\boldsymbol{j}+4e^{\ln\left(x^2+y^2+5\right)}\boldsymbol{k}\right)$ = _____

(c) $\operatorname{div}\left(\dfrac{x}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{i}+\dfrac{y}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{j}+\dfrac{z}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{k}\right)$ = _____

2. Consider $\operatorname{div}\left(\dfrac{-y\vec{i}+x\vec{j}}{(x^2+y^2)^2}\right)$.

(a) Is this a vector or a scalar? (☐ vector ☐ scalar)

(b) Calculate it:

$\operatorname{div}\left(\frac{-y\vec{i}+x\vec{j}}{(x^2+y^2)^2}\right)$ = _____ + _____ = _____

3. Let $f(x,y)=axy+ax^2y+y^3$.

(a) $\operatorname{div}(\operatorname{grad}(f))$ = _____

(b) Find a so that $\operatorname{div}(\operatorname{grad}(f))=0$ for all x,y.

a = _____

4. Let $\vec{F}(x,y,z)=4x^2\vec{i}-\cos(xy)(\vec{i}+\vec{j})$. Calculate the divergence:

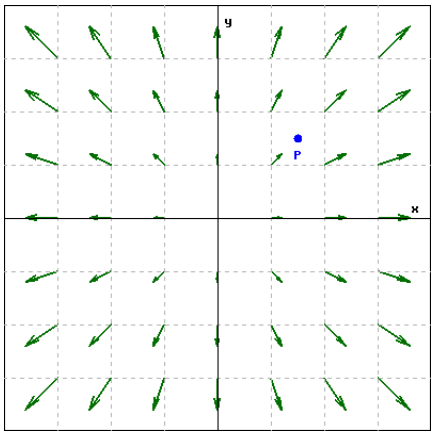
$\operatorname{div}\vec{F}(x,y,z)$ = _____

5. Find the divergence of the vector field $\vec{F}=\vec{a}\times\vec{r}$, if $\vec{r}=x\,\vec{\imath}+y\,\vec{\jmath}+z\,\vec{k}$ and $\vec{a}=6\,\vec{\imath}+3\,\vec{\jmath}+7\,\vec{k}$, by first calculating the cross-product.

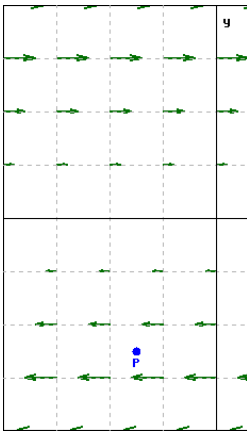
$\vec{F}$ = _____

so that $\operatorname{div}\vec{F}$ = _____

6. Determine whether the divergence of each vector field (in green) at the indicated point P (in blue) is positive, negative, or zero.



(☐ Zero ☐ Negative ☐ Positive)



(☐ Zero ☐ Negative ☐ Po

(Click on a graph to enlarge it)

7. Suppose $\text{div}(\vec{F}) = 4x^2 + y^2 - z^2$. Estimate the flux out of a small sphere of radius 0.3 centered at $(4, 4, 2)$.

Flux out of the small sphere $\approx$ _____

8.

(a) Let $\vec{F} = \langle F_1, F_2, F_3 \rangle$ and let

$$\vec{G} = (\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z})\hat{\imath} - (\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z})\hat{\jmath} + (\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y})\hat{k}$$

Show that $\text{div}(\vec{G}) = \vec{0}$.

(b) Vector fields with a zero divergence everywhere in their domain are called **divergence-free** or **incompressible** vector fields. Which of the following vector fields are divergence-free?

i. $\vec{F} = \langle -y, z, x \rangle$

ii. $\vec{F} = \langle \cos(yz), 3xe^{z-x}, 6(x + y + z)^3 \rangle$

iii. $\vec{F} = \langle 4xyz, y^2z, yz^2 \rangle$

iv. $\vec{F} = \nabla f$ where f is a scalar function of x, y, and z

(c) Let $\vec{F}_1 = \langle 3(x - z)^2, 2\cos(x) + 3yz + y, -(z - 1)^2 + e^{xy} \rangle$. Calculate the divergence of $\vec{F}_1$ and give a point where $\text{div}(\vec{F}_1) = 0$.

(d) Is $\vec{F}_1$ a divergence free vector field?

9. In this exercise, we will look at how different forms of component function's dependence on our variables will change the divergence of the vector field. For each of the vector fields given below, decide if the divergence of the vector field must be zero and write a sentence or two to justify your decision.

(a)

$$\vec{F} = f(x)\hat{\imath} + g(y)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

(b)

$$\vec{G} = f(x)\hat{\imath} + f(y)\hat{k} + \hat{\jmath} + f(z)\hat{\jmath}$$

(c)

$$\vec{H} = f(y,z)\hat{\imath} + g(x,z)\hat{k} + \hat{\jmath} + h(x,y)\hat{\jmath}$$

(d)

$$\vec{A} = g(x,y)\hat{\imath} - g(x,y)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

(e)

$$\vec{B} = g(x,y)\hat{\imath} - g(y,x)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

13.8 The Curl of a Vector Field

Motivating Questions

- What is meant by rotation of a vector field in a plane?
- How can a two-dimensional measurement of rotation be generalized to work in three dimensions?
- How can the rotational strength of a vector field be measured?

13.8.1 Introduction

In Section 13.7, we examined how the strength of a vector field in two (or more) dimensions changed in different regions. In particular, we developed the divergence of a vector field as a local (or density) measurement for how the strength of the vector field changes. The key ideas when interpreting divergence are:

- A positive divergence means that the vector field is growing in strength.
- A negative divergence means that the vector field is decreasing in strength.
- A zero divergence means that the vector field is not changing in strength.

In many physical settings, it is also useful to measure the rotational strength of a vector field at a local scale. For instance, the vector field on the left of Figure 13.8.1 shows a vector field which flows in a counterclockwise fashion around the origin. The vector field on the right of Figure 13.8.1 shows a vector field which does not have a rotational aspect to its flow.

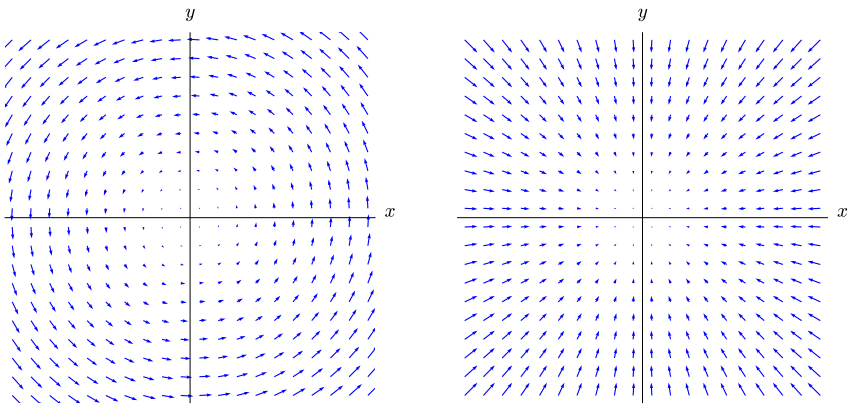


Figure 13.8.1 Examples of vector fields with global rotation properties (left) and without rotation in either the global or local sense (right)

These global ideas of rotation are nice, but are not always visually apparent or may only appear in some regions. For instance, the vector field on the left of Figure 13.8.2 seems to have rotation around at least a couple of different points. In each quadrant, the vector field on the right of Figure 13.8.2 has different rotational patterns.

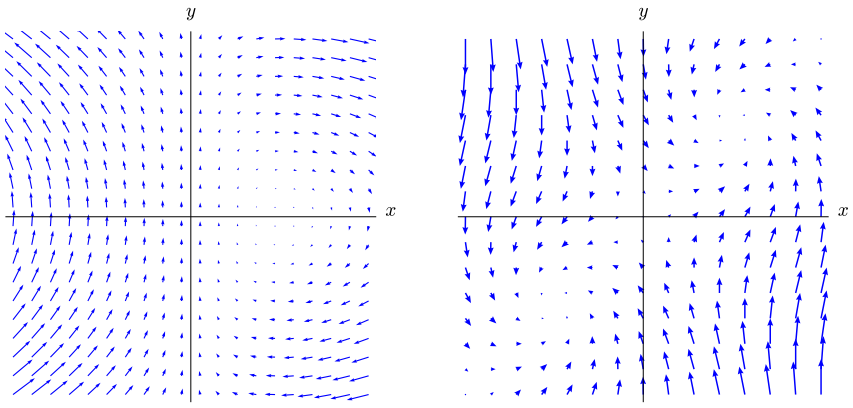
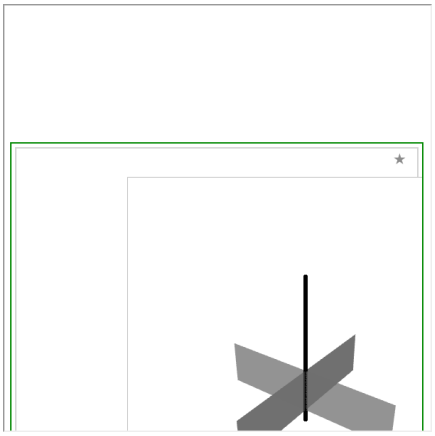


Figure 13.8.2 Examples of vector fields with varying rotation properties

Preview Activity 13.8.1 We would like to understand and measure rotation of a vector field near a particular point. In order to investigate this concept, we will look at some two-dimensional vector fields and think about whether the vector field shown will rotate a small pinwheel or spinner placed at a particular location. The sort of spinner we imagine is illustrated in Figure 13.8.3. It consists of a central axis with a four-bladed paddle placed at one end of the axis. We imagine that the spinner is anchored at a point and the vector field, perhaps thought of as a fluid flow or wind velocity vector field, pushes against the blades of the spinner's paddle. In this activity, we will be trying to assess how the the spinner will rotate around the black axel (as an axis of rotation).



Standalone
In Context
Embed

Figure 13.8.3 A paddle-bladed spinner

To begin our investigation of the rotation of a spinner in a vector field, we will look at the vector field $\vec{F}$ in Figure 13.8.4. As you think about these questions, draw an "X" at each of the points about which you are asked and consider the vector field as being the pattern of a wind blowing across the plane.

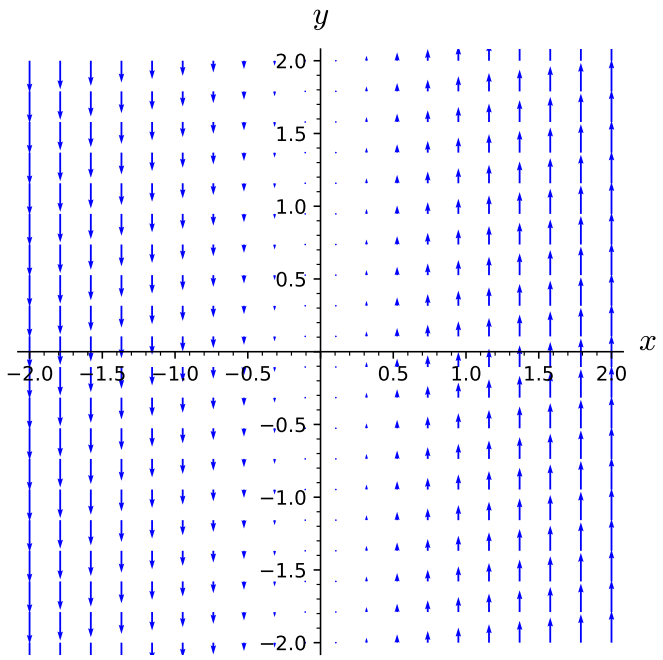


Figure 13.8.4 The vector field $\vec{F}$

(a) Draw an X at the origin to act as your spinner. Draw a vector on the top right blade of your spinner that represents how the wind will push on that blade. Next, draw a vector on each of the other blades of your spinner that represents how the wind will push on that blade.

(b) Use your vector representations from the previous part to describe if a small spinner placed at the origin would spin clockwise, counterclockwise, or not spin at all.

(c) Now we would like to look at the rotational strength of $\vec{F}$ at the point $(1, 1)$. As before, draw a spinner at this point and draw vectors on each of the blades to represent how the wind will push on that blade. It is important at this stage to draw the relative lengths of the vectors on each blade to scale so you can see which blades have a larger force due to the wind.

(d) Use your vector representations from the previous part to describe if a small spinner placed at at $(1, 1)$ would spin clockwise, counterclockwise, or not spin at all. You should pay attention to which blades will have a stronger force (in comparison to the blade on the other side).

(e) Would a small spinner placed at $(-2, 1)$ spin clockwise, counterclockwise, or not spin at all? Draw a representation of a spinner if necessary to demonstrate your ideas.

(f) Are there any points on the xy-plane where the spinner would not turn counterclockwise?

13.8.2 Measuring the Circulation Density of Vector Field in $\mathbb{R}^2$

In this subsection, we will develop the measurement of the circulation density for a two-dimensional vector field. We will use this measurement to generalize to a notion of rotational strength in higher dimensional cases in the next subsection.

We will start by measuring the circulation of a vector field on a path around the point (a, b) and use this measurement to define circulation density. Specifically, we will measure the circulation of a vector field as we move around a square centered at (a, b). Using this measurement, we will calculate the circulation density by dividing our measurement by the area enclosed. This will allow us to compare our measurement across regions of different sizes. By taking the limit of this circulation density as the square's side length goes to zero, we will have the circulation density at the point (a, b). Just as in our discussion of Subsection 13.7.3, we will look at a two dimensional setting first, then examine how our argument can be generalized to higher dimensions.

Let's start by measuring the circulation around a square with side length $2h$ centered at a point (a, b). Namely, we will look at the line integral of our vector field as we move along the square curve shown in Figure 13.8.5.

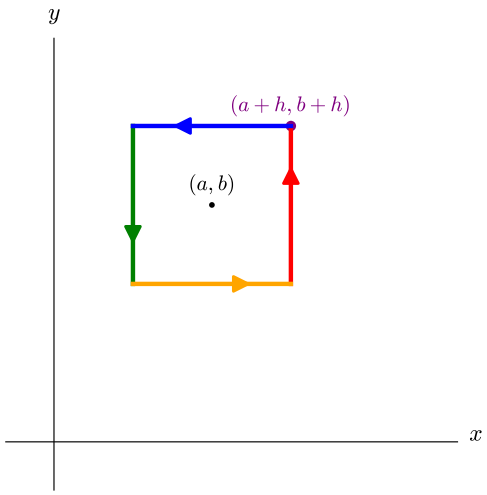


Figure 13.8.5 A square centered at the point (a, b) with side lengths $2h$ traveled counterclockwise

We can parametrize the top edge of the square using the parametrization

$$\vec{r}_{\text{top}}(t) = \langle a - t, b + h \rangle$$

for $-h \leq t \leq h$. Similarly, the bottom, right, and left can be parametrized (over the same range of values for t) as follows:

$$\begin{aligned} \vec{r}_{\text{bottom}}(t) &= \langle a + t, b - h \rangle \\ \vec{r}_{\text{right}}(t) &= \langle a + h, b + t \rangle \\ \vec{r}_{\text{left}}(t) &= \langle a - h, b - t \rangle \end{aligned}$$

In order to measure the circulation around the square, we want to know how much of the vector field $\vec{F} = \langle F_1, F_2 \rangle$ is parallel to the direction of travel. Thus, along each of the sides of the box, we only need to look at one of the components of the vector field. The only contribution to the circulation on the top and bottom comes from the horizontal component F_1. Similarly, the

vertical component, F_2, is all that contributes to the circulation on the right and left sides. This means that we can simplify our line integrals considerably once we apply Theorem 13.4.6. For instance, the circulation on the top can be written as

$$\int_{\text{top}} \vec{F} \cdot d\vec{r} = \int_{-h}^{h} \langle F_1, F_2 \rangle \cdot \langle -1, 0 \rangle \, dt = \int_{-h}^{h} -F_1(a-t, b+h) \, dt$$

Similarly, we can simplify the appropriate line integrals on the other sides (ordered below as top, left, bottom, right) to get the total circulation around the square to be

$$\begin{aligned} \text{Total Circulation} &= \int_{\text{Square}} \vec{F} \cdot d\vec{r} \\ &= \int_{-h}^{h} -F_1(a-t, b+h) \, dt + \int_{-h}^{h} -F_2(a-h, b-t) \, dt \\ &\quad + \int_{-h}^{h} F_1(a+t, b-h) \, dt + \int_{-h}^{h} F_2(a+h, b+t) \, dt \end{aligned}$$

Note that on the top and the left sides, we must switch the sign because the direction of travel on these sides is in the negative coordinate direction. It is important to note that by convention, we talk about "positive" rotation being counterclockwise and "negative" rotation being clockwise when looking at these planar graphs. This will be discussed more fully in Subsection 13.8.3.

As with our development of divergence, we apply the Mean Value Theorem for Integrals to the four integrals above to simplify future calculations. Specifically, the Mean Value Theorem for Integrals applied to the first integral tells us that there is a value, t^*_{top}, in the interval $(a-h, a+h)$ such that

$$\int_{-h}^{h} -F_1(a-t, b+h) \, dt = -2h \left(F_1(t^*_{\text{top}}, b+h) \right)$$

Applying the same arguments to the other three integrals in our Total Circulation calculation gives

$$\begin{aligned} \text{Total Circulation} &= \int_{\text{Square}} \vec{F} \cdot d\vec{r} \\ &= \int_{-h}^{h} -F_1(a-t, b+h) \, dt + \int_{-h}^{h} -F_2(a-h, b-t) \, dt \\ &\quad + \int_{-h}^{h} F_1(a+t, b-h) \, dt + \int_{-h}^{h} F_2(a+h, b+t) \, dt \\ &= -2hF_1(t^*_{\text{top}}, b+h) - 2hF_2(a-h, t^*_{\text{left}}) \\ &\quad + 2hF_1(t^*_{\text{bottom}}, b-h) + 2hF_2(a+h, t^*_{\text{right}}) \end{aligned}$$

Here we have $t^*_{\text{top}}, t^*_{\text{bottom}} \in (a-h, a+h)$ and $t^*_{\text{left}}, t^*_{\text{right}} \in (b-h, b+h)$. We can regroup these terms to get the following expression for the total circulation around our square:

$$\begin{aligned} &2h \left(F_2(a+h, b^*_{\text{right}}) - F_2(a-h, b^*_{\text{left}}) \right) \\ &-2h \left(F_1(a^*_{\text{top}}, b+h) - F_1(a^*_{\text{bottom}}, b-h) \right) \end{aligned}$$

A larger square is likely to have a larger total circulation since there is more distance to accumulate how much the vector field is moving in the same

direction as our path. This is why the side length $2h$ of the square appears as a factor in the formula for our total circulation. In order to compare our rotational or circulation ideas over different sizes of squares, we will now look at the **circulation density** (strength of rotation per unit area), which can be computed directly from our total circulation measurement. We can take the limit of the circulation density as we shrink the square to the central point (a, b) and get a measurement rotational strength of our vector field at a point.

We will compute our circulation density as

$$\text{Circulation Density} = \lim_{h \to 0} \frac{\text{Total Circulation}}{\text{Area of Square}}.$$

Since the area of the square is $(2h)^2 = 4h^2$, this gives us that the circulation density is

$$\lim_{h \to 0} \frac{F_2(a+h, t^*_{\text{right}}) - F_2(a-h, t^*_{\text{left}})}{2h} - \frac{F_1(t^*_{\text{top}}, b+h) - F_1(t^*_{\text{bottom}}, b-h)}{2h}.$$

Recall the central difference method of estimating derivatives from Section 1.5.2 and notice that as $h \to 0$, $t^*_{\text{top}}, t^*_{\text{bottom}}$ must go to a and $t^*_{\text{left}}, t^*_{\text{right}}$ must go to b. Therefore, our circulation density measurement at a point is

$$\text{Circulation Density at the point } (a, b) = \frac{\partial F_2}{\partial x}(a, b) - \frac{\partial F_1}{\partial y}(a, b)$$

You may have noticed that this argument had a similar flow to the development of divergence in Section 13.7. However, the result here is different because the property of the vector field we were trying to measure was different. Because divergence was developed to measure the change in the *strength* of the vector field (without regard to the direction), divergence is computed using the partial derivatives of the horizontal component of the vector field with respect to the horizontal variable and the vertical component of the vector field with respect to the vertical variable. The circulation density we just developed measures how the *direction* of the vector field is changing, and thus uses partial derivatives of the components with respect to the transverse variable. In other words, the rotational aspects of the vector field depend on how the horizontal component of the vector field changes when we move vertically (and vice versa).

We could have used a shape other than a square in the development of the circulation density, but a square allows for calculations that are easier to understand as the area of the shape decreases.

Example 13.8.6 In this example, we will look at the circulation density of the vector field given by $\vec{F} = \langle \sin(x+y), x-y \rangle$ which is plotted in Figure 13.8.7.

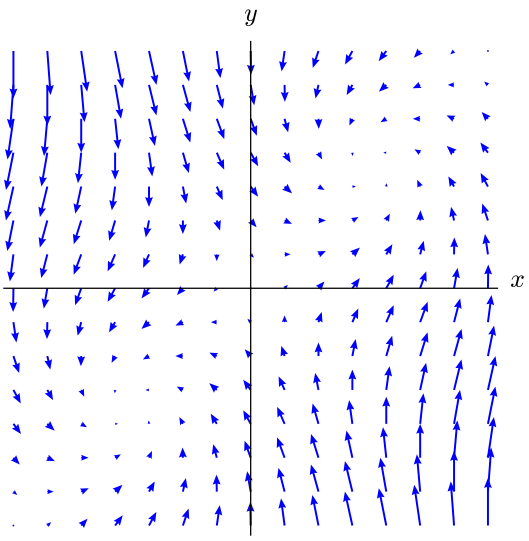


Figure 13.8.7 The vector field $\vec{F} = \langle \sin(x+y), x-y \rangle$

We can compute the partial derivatives of our component functions to see that $\vec{F}$ will have a circulation density of

$$\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 1 - \cos(x+y).$$

Thus, we will have a positive circulation density at all points except where $\cos(x+y) = 1$. You can look at points in the first or third quadrant of Figure 13.8.7 and see that a spinner will rotate in the counterclockwise direction. In particular, you can see that the strongest rotation will occur near points on the line $y = x$ since those will have a circulation density of 2.

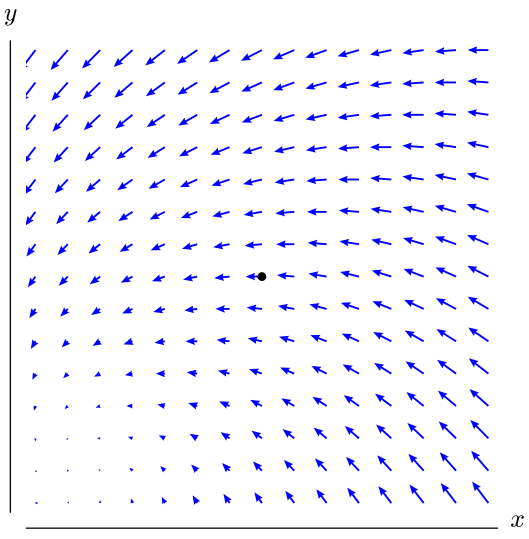


Figure 13.8.8 The vector field $\vec{F} = \langle \sin(x+y), x-y \rangle$ shown around the point (2,2).

If we look at a point where the circulation density is zero, like at $(-2, 2)$, then we can see what appears to be a close to a constant vector field, as shown in Figure 13.8.9. This means that the amount that any blade would be pushed

by our vector field, the blade opposite would be pushed in exactly the same way, yeilding no rotation.

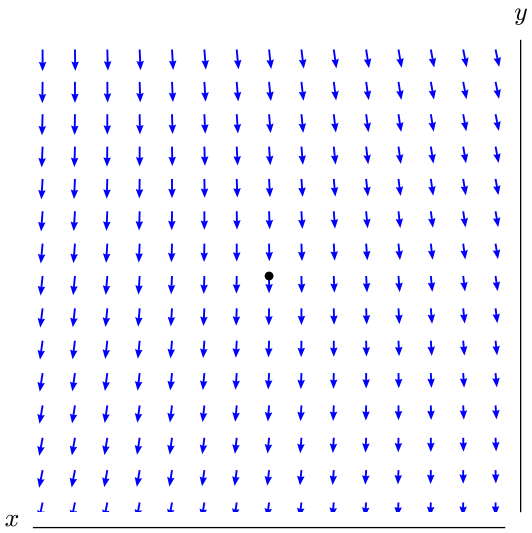


Figure 13.8.9 The vector field $\vec{F} = \langle \sin(x+y), x-y \rangle$ shown around the point (-2,2).

▲

Activity 13.8.2 Matching Visual Measurements to Algebraic Calculations. Remember that the circulation density of a vector field $\vec{F} = \langle F_1, F_2 \rangle$ at a point (a, b) is calculated as

$$\text{Circulation Density at the point } (a,b) = \frac{\partial F_2}{\partial x}(a,b) - \frac{\partial F_1}{\partial y}(a,b).$$

The circulation density will be positive when a small spinner placed at (a, b) will spin counterclockwise, negative when the spinner will move clockwise, and zero when the spinner does not rotate.

For each of the vector fields given below, answer the following questions:

i. What is the formula for the circulation density of the given vector field?

ii. For what points will you have a positive circulation density?

iii. For what points will you have a negative circulation density?

iv. For what points will you have a zero circulation density?

(a) $\langle -y, x \rangle$

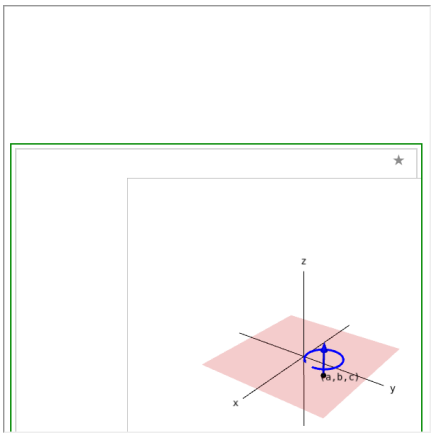
(b) $\langle x, y \rangle$

(c) $\langle 1 - x, xy \rangle$

You can check your work from the previous activity with the ideas presented in previous plots and work. In particular, the vector field from part (a) can be seen as the left plot in Figure 13.8.1, the vector field from part (b) can be seen as exactly opposite of the field shown in the right plot of Figure 13.8.1, and the vector field in part (c) can be seen as the left plot of Figure 13.8.2.

13.8.3 Measuring Rotation in Three Dimensions

The previous subsection showed how we can measure the circulation density, or strength of rotation, at a point for a two-dimensional vector field. In this subsection, we will look at how this two-dimensional measurement can be used to define circulation density for three dimensional vector fields.



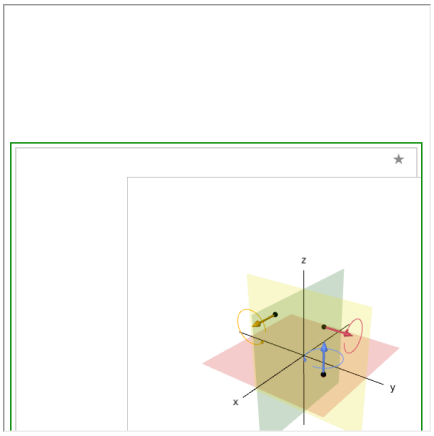
Standalone
In Context
Embed

Figure 13.8.10 Rotation as a vector on the xy-plane

We first need to consider how we want to represent and measure rotation in three dimensions. When we developed the circulation density on the xy-plane, we measured the rotational strength at a point (a, b) with an axis of rotation coming out of the xy-plane. This corresponds to the axis of rotation being given by the blue vector in Figure 13.8.10. Remember that "positive" rotation corresponds to counterclockwise rotation in the xy-plane. We can generalize this idea to think of rotation in three dimensions as being represented by a vector where

- the direction of the vector represents the axis of rotation and
- the magnitude of the vector represents the strength of the rotation.

By convention, we consider positive "rotation" to correspond to counterclockwise rotation when the vector field is viewed looking from the terminal point of the vector to its initial point. In Figure 13.8.11, you can see that each vector corresponds to the rotation displayed and is consistent with the conventions described above.



Standalone
In Context
Embed

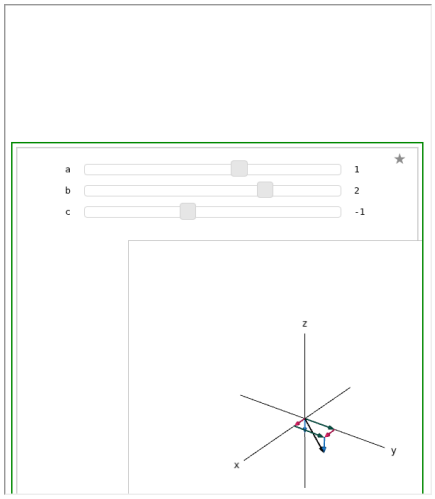
Figure 13.8.11 Rotation as a vector on each coordinate plane

If you look carefully at Figure 13.8.11, you can see that the vector shown in blue will be parallel to $\hat{k}$ and will represent rotations on planes of the form $z = c$. Similarly, the vector shown in yellow will be parallel to $\hat{\imath}$ and will represent rotations on planes of the form $x = a$. When looking down the blue and yellow vectors (from the terminal to the initial point), you can see the positive coordinate axes as pointing to the right and up[1] (as we would expect on a two-dimensional plot). In contrast, when you look down the magenta vector (from the terminal to the initial point), the positive coordinate axes (the x and z axes) point to the left and up.[2] In fact, when we look down the magenta vector the xz-plane is flipped. A positive rotation on a plane of the form $y = b$ will correspond to a rotation vector that is in the direction of $-\hat{\jmath}$. This is a consequence of the right-handed coordinate system and our right-handed idea of rotation, as reflected in the relationships amongst the vectors $\hat{\imath}, \hat{\jmath}\,\hat{k}$ recalled below:

$$\hat{\imath} \times \hat{\jmath} = \hat{k} \qquad \hat{\imath} \times \hat{k} = -\hat{\jmath} \qquad \hat{\jmath} \times \hat{k} = \hat{\imath}$$

By using a vector to represent the rotation of some element, we have enabled all of the tools that vectors allow, especially the ideas of projection and linear combinations. Remember that the projection of a vector $\vec{v}$ onto a vector $\vec{u}$ will give the vector component of $\vec{v}$ that is parallel to $\vec{u}$. In other words, the projection of $\vec{v}$ onto $\vec{u}$ is how much of $\vec{v}$ is parallel to $\vec{u}$. The projection formula will allow us to take a vector representation of rotation and determine how much rotation happens around a given axis of rotation.

Recall that we have used a set of coordinate vectors (namely $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$) to write other vectors as a linear combination of our base vectors. In other words, a vector $\langle a, b, c\rangle$ in $\mathbb{R}^3$ can be written in the form $a\hat{\imath} + b\hat{\jmath} + c\hat{k}$.



Standalone
In Context
Embed

Figure 13.8.12 The black vector, $\langle a, b, c\rangle$, is decomposed into parts parallel to $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

The linear combination of vectors will allow us to build the total rotation vector by measuring the rotational strength in the direction of three coordinate direction vectors. Just as in Figure 13.8.12, we will build our total rotation vector as a sum of $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$, but we will think of the initial point of the black vector as being at our point of interest (u, v, w)A rotation vector in the

[1]When placing the x-axis horizontally for the blue vector and placing the y-axis horizontally for the yellow vector

[2]When placing the x-axis horizontally.

direction of $\hat{k}$ at a point (u, v, w) will measure the circulation density at our point restricted to the trace plane $z = w$. Similarly the rotation vectors in the directions $\hat{\imath}$ and $\hat{\jmath}$ will measure the circulation density on the trace planes $x = u$ and $y = v$, respectively.

13.8.4 Circulation Density in Three Dimensions

The previous subsection discussed how we can view the amount of rotation of a three-dimensional vector field in a plane parallel to the xy-, xz-, or yz-plane and recalled that vectors in $\mathbb{R}^3$ can be written as a linear combination of the vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$. We use our definition of curl for a two-dimensional vector field to measure the amount of rotation in the appropriate planes, which leads to the following definition.

Definition 13.8.13 Let $\vec{F} = \langle F_1, F_2, F_3 \rangle$ be a three-dimensional vector field. The **curl** of $\vec{F}$ is given by

$$\text{curl}(\vec{F}) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\imath} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z}\right)\hat{\jmath} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{k}.$$

◆

Alternate Notation for Curl. As with divergence, there is an alternate notation for curl that uses the del operator $\nabla = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \rangle$. Specifically, for a three-dimensional vector field $\vec{F}$, $\text{curl}(\vec{F}) = \nabla \times \vec{F}$.

The curl is exactly the rotational description described at the end of Subsection 13.8.3. When evaluated at a point (u, v, w), the first component of the curl will measure the circulation density of the vector field restricted to the plane $x = u$. Similarly, the second and third components of the curl will measure the circulation density of the vector field restricted to the planes $y = v$ and $z = w$, respectively.

Example 13.8.14 In this example, we will look at $\vec{F} = \langle 0, x, 0 \rangle$. Applying Definition 13.8.13,

$$\text{curl}(\vec{F}) = \left\langle \frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(x), -\left(\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}(0)\right), \frac{\partial}{\partial x}(x) - \frac{\partial}{\partial y}(0) \right\rangle$$

Thus, $\text{curl}(\vec{F}) = \langle 0, 0, 1 \rangle$. This means that $\vec{F}$ will have a rotational aspect only with an axis parallel to the z-axis. This should not be surprising since the plot of $\vec{F}$ will be the same as Figure 13.8.4 copied on each plane $z = c$. As you saw in Preview Activity 13.8.1, this vector field will have constant rotational strength at every point with axis of rotation in the $\hat{k}$ direction. ▲

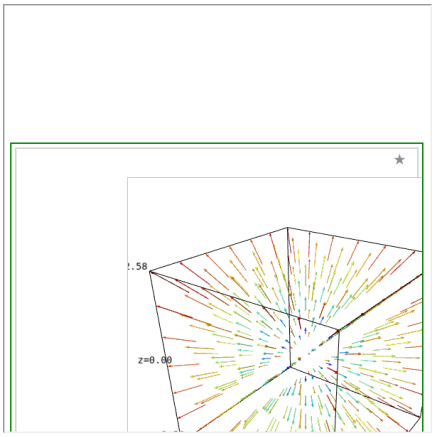
We next consider an example where the trace of the vector field in planes parallel to coordinate planes is not as easy to visualize as in the previous example.

Example 13.8.15 In this example, we will look at $\vec{G} = \langle x, y, z \rangle$. If we apply Definition 13.8.13, we obtain

$$\text{curl}(\vec{G}) = \left\langle \frac{\partial}{\partial y}(z) - \frac{\partial}{\partial z}(y), -\left(\frac{\partial}{\partial x}(z) - \frac{\partial}{\partial z}(x)\right), \frac{\partial}{\partial x}(y) - \frac{\partial}{\partial y}(x) \right\rangle$$

Therefore, we have $\text{curl}(\vec{G}) = \langle 0, 0, 0 \rangle$. Thus, $\vec{G}$ will have no rotational strength anywhere. In Figure 13.8.16 we have a plot of $\vec{G}$, which illustrates that dropping a spinner at any point in space will not have the spinner rotate. No matter

what orientation the spinner will have, there will be equal force on each side of the spinner, and thus the spinner will not rotate.



Standalone
In Context
Embed

Figure 13.8.16 The black vector, $\langle a, b, c\rangle$, is decomposed into parts parallel to $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

▲

Example 13.8.17 In this example, we will look at $\vec{H} = \langle x - y, y + 2z, x^2\rangle$. By Definition 13.8.13,

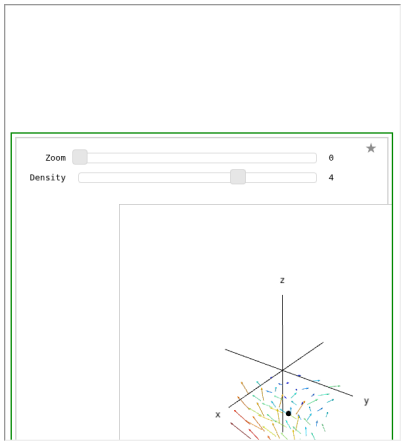
$$\text{curl}(\vec{H}) = \left\langle \frac{\partial}{\partial y}(x^2) - \frac{\partial}{\partial z}(y+2z), -\left(\frac{\partial}{\partial x}(x^2) - \frac{\partial}{\partial z}(x-y)\right), \right.$$
$$\left. \frac{\partial}{\partial x}(y+2z) - \frac{\partial}{\partial y}(x-y)\right\rangle$$

Therefore, $\text{curl}(\vec{H}) = \langle 0, -2x, 1\rangle$. This tells us that the vector field $\vec{H}$ will have a non-zero curl for every point. However, it will be quite difficult to "see" the curl at a particular point. The next activity will take you through the process of trying to visualize and understand the output of curl in three dimensions.

▲

Activity 13.8.3 Estimating Curl in Three Dimensions.

(a) Consider the vector field $\vec{F}$ plotted in Figure 13.8.18. You can adjust the size of the region around $(1, 1, -2)$ over which the vector field is plotted using the "Zoom" slider. The "Density" slider allows you to adjust the number of vectors plotted. Try to identify any rotation in the three dimensional vector field plot at the point $(1, 1, -2)$. Write a sentence describing how a spinner placed at $(1, 1, -2)$ would rotate, including along which axis it would rotate. Try to state your answer as a vector representing the rotational strength of the vector field at $(1, 1, -2)$.



Standalone
In Context
Embed

Figure 13.8.18 A vector field plotted in a region around the point $(1, 1, -2)$

(b) You likely found it difficult to decide how you thought a spinner might rotate in this new, three-dimensional setting. Let's look at the vector field in the plane $z = -2$, as displayed in Figure 13.8.19. Do you think a spinner placed on the red point would rotate clockwise, counterclockwise, or not rotate? If the spinner will rotate, you should think about what the axis of rotation would be and whether the rotation should be positive or negative. Summarize your result as a vector representing the rotational strength of $\vec{F}$ in the plane $z = -2$.

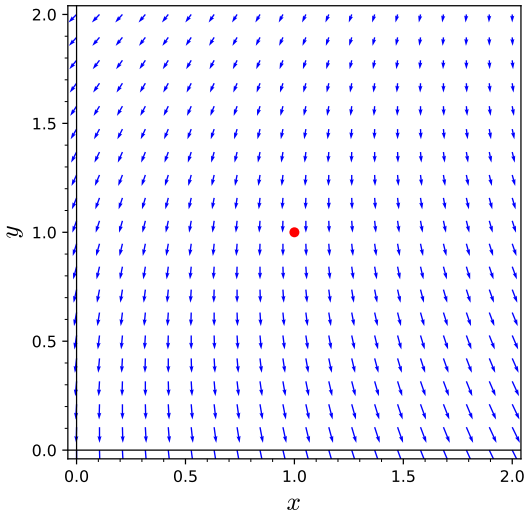


Figure 13.8.19 The trace of $\vec{F}$ in the plane $z = -2$

(c) Next we will look at the vector field in the plane $y = 1$, as displayed in Figure 13.8.20. Do you think a spinner placed on the red point would rotate clockwise, counterclockwise, or not rotate? If the spinner will rotate, you should think about what the axis of rotation would be and whether the rotation should be positive or negative. Summarize your result as a vector representing the rotational strength of your vector field in the plane $y = 1$. Do you think the rotation in this figure is stronger or weaker than in Figure 13.8.19?

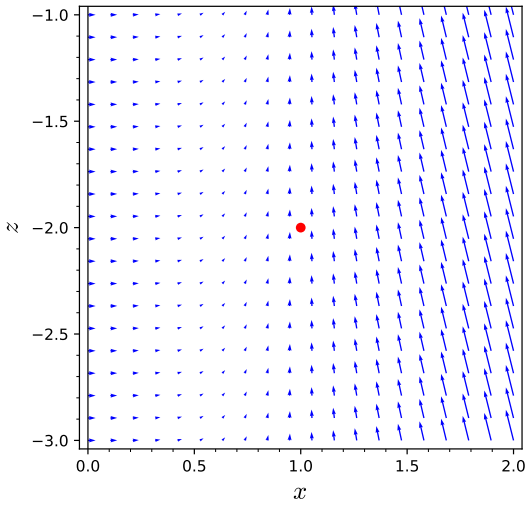


Figure 13.8.20 The trace of $\vec{F}$ in the plane $y = 1$

(d) Finally, consider the trace of $\vec{F}$ in the plane $x = 1$, as displayed in Figure 13.8.21. Do you think a spinner placed on the red point would rotate clockwise, counterclockwise, or not rotate? If the spinner will rotate, you should think about what the axis of rotation would be and whether the rotation should be positive or negative. Summarize your result as a vector representing the rotational strength of your vector field in the plane $x = 1$. How do you think the rotation in this figure compares (i.e., stronger or weaker) to that in Figure 13.8.19 and Figure 13.8.19?

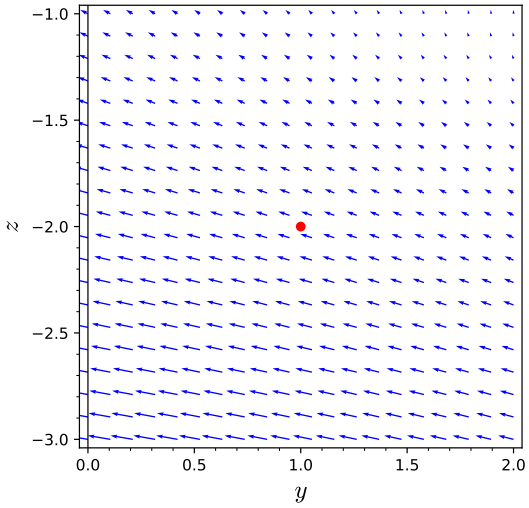


Figure 13.8.21 The trace of $\vec{F}$ in the plane $x = 1$

(e) Summarize your prediction to what you think the three-dimensional rotational strength of the vector field will be at the point $(1, 1, -2)$ in the form of three-dimensional vector.

Hint. Add your results from the previous three steps.

(f) Compute the curl of $\vec{F} = \langle x - y, y + 2z, x^2 \rangle$. Specifically, what is curl$(\vec{F})$ at the point $(1, 1, -2)$?

(g) Compare the result of the curl calculation in part f to your prediction from part e. You likely found it difficult to estimate the magnitude, so your answer there may be incorrect. Hopefully, you did get the signs of the components and their relative strengths (i.e., which is biggest) correct. If you did not, go back and review the previous parts and explain why the calculated components match with the rotational strength for each of the three figures.

13.8.5 Interpretation and Usage of Curl

It is worth making explicit a fact that we have used implicitly throughout this section: the curl of a vector field is itself a vector field! That is, evaluating $\text{curl}(\vec{F})$ at a point gives a vector. As we saw earlier in this section, the vector output of $\text{curl}(\vec{F})$ represents the rotational strength of the vector field $\vec{F}$ as a linear combination of rotational strengths (or circulation densities) from two-dimensional planar descriptions. From our description of vectors as a representation of rotations, We can think of rotation in three dimensions as being represented by a vector where

- the direction of the vector represents the axis of rotation and
- the magnitude of the vector represents the strength of the rotation.

By convention, we consider positive "rotation" to correspond to counterclockwise rotation when the vector field is viewed looking from the end of the vector to the base.

With these conventions, the output vector of the curl evaluated at a point P, written as $\text{curl}(\vec{F})(P)$, will have the following properties:

1. The direction of $\text{curl}(\vec{F})(P)$ will be the axis of rotation that has the strongest positive rotational strength.
2. The magnitude of $\text{curl}(\vec{F})(P)$ will be the circulation density of the vector field in the plane through P with normal vector $\text{curl}(\vec{F})(P)$.

The following theorem allows us to use $\text{curl}(\vec{F})$ to measure the rotational strength of $\vec{F}$ around an arbitrary axis. This will be particularly useful to us in later sections.

Theorem 13.8.22 *The rotational strength of a vector field $\vec{F}$ around an axis given by a unit vector $\vec{v}$ at a point $P = (a, b, c)$ can be computed by $\Big(\text{curl}(\vec{F})(a, b, c)\Big) \cdot \vec{v}$.*

Proof. This follows from the formula for the component of $\text{curl}(\vec{F})(a, b, c)$ along $\vec{v}$ and the fact that $\vec{v}$ was selected to be a unit vector. ■

Activity 13.8.4 In this activity, we will work on calculating curl algebraically and interpreting it.

(a) Consider the vector field $\vec{F} = \langle x, y, z \rangle$. If we plot this vector field in any plane through the origin, we will see the vector field shown in Figure 13.8.23. This two-dimensional vector field has no rotation. The projection of $\text{curl}(\vec{F})(0, 0, 0)$ onto any direction therefore must give the zero vector. The only vector that has a zero projection in every direction is the zero vector. Verify this geometric argument for the curl of $\vec{F}$ by doing the calculations necessary to show that $\text{curl}(\langle x, y, z \rangle) = \vec{0}$.

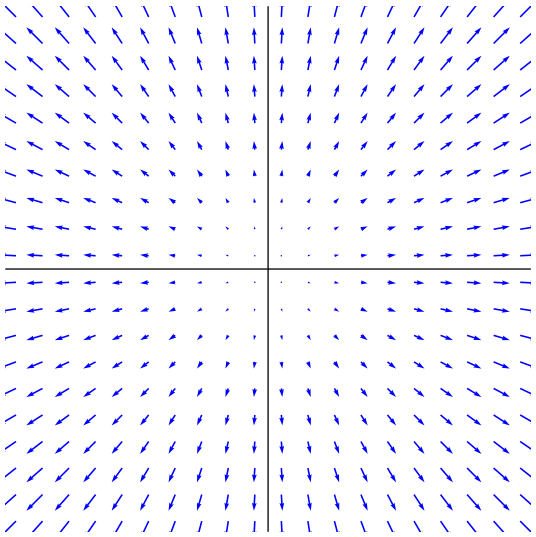
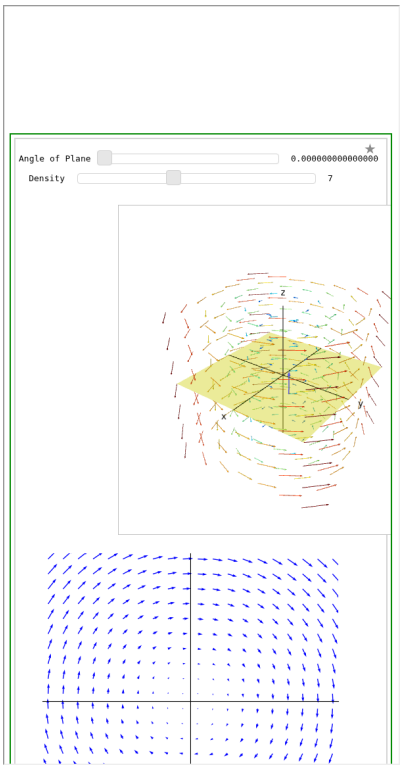


Figure 13.8.23 The vector field $\langle x, y, z \rangle$ on every plane through the origin

(b) Now consider the vector field $\vec{F} = \langle -y, x, 0 \rangle$, which is shown in Figure 13.8.24. The figure includes both an interactive three-dimensional plot of $\vec{F}$ as well as a two-dimensional plot of $\vec{F}$ in the yellow plane, which can be adjusted using the "Angle of Plane" slider.



Standalone
In Context
Embed

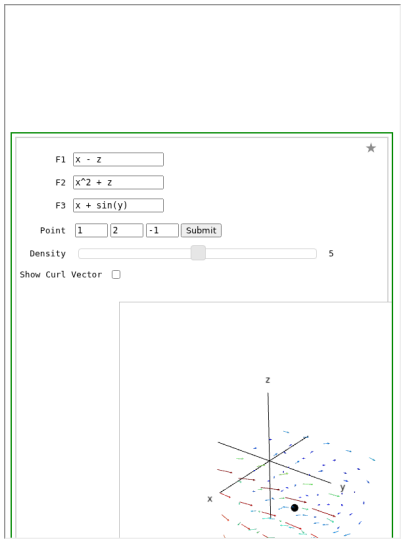
Figure 13.8.24 Two views of the vector field $\langle -y, x, 0 \rangle$

(i) Based on the figure, would you estimate the components of curl$(\vec{F})$ to be positive, negative, or zero at the origin? Does your answer change if you pick a different point?

(ii) Calculate curl($\vec{F}$) algebraically. Does the curl of this vector field vary depending on the point at which the curl is measured?

(iii) Explain why the rotational strength of the vector field on a titled plane through the origin will be given by $2\cos(\theta)$, where θ is the angle between $\hat{k}$ and the normal vector of the tilted plane.

(c) In Figure 13.8.25 you can plot a vector field in a region around a point of your choosing in order to look at the rotational properties of the vector field. The check box in Figure 13.8.25 will show the curl vector at the base point specified so you can make sense of your vector field and its curl.



Standalone
In Context
Embed

Figure 13.8.25 A plot of the vector field $\langle F_1, F_2, F_3\rangle$

Use the figure to estimate the direction of curl($\langle x - z, x^2 + z, x + \sin(y)\rangle$) at the point $(1, 2, -1)$. Confirm your estimate by calculating the curl at this point algebraically. Is there any point at which the direction of greatest rotational strength of this vector field has negative x-component? If there is, find such a point. If not, explain why not.

Summary

- The circulation of a vector field on a closed path measures how strong the vector field moves in the direction of travel for the path. The circulation density of a two-dimensional vector field $\vec{F} = \langle F_1, F_2\rangle$ is given by

$$\text{Circulation Density at the point } (a,b) = \frac{\partial F_2}{\partial x}(a,b) - \frac{\partial F_1}{\partial y}(a,b).$$

  The circulation density can be visualized in terms of how fast a very small spinner anchored to (a, b) will rotate as a result of the force of $\vec{F}$.

- By considering the circulation density of a three-dimensional vector field $\vec{F} = \langle F_1, F_2, F_2\rangle$ in planes parallel to the coordinate planes, we can calculate the curl of $\vec{F}$ as

$$\text{curl}(\vec{F}) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\imath} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z}\right)\hat{\jmath} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{k}.$$

- For a three-dimensional vector field $\vec{F}$, the vector curl$(\vec{F})(a, b, c)$ points along the axis of rotation with the greatest rotational strength at the

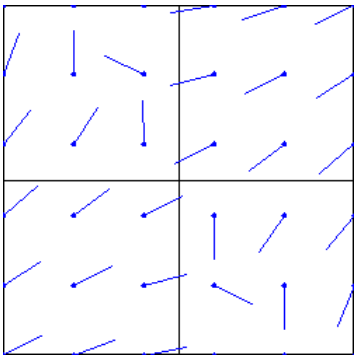
point (a, b, c). The magnitude of $\text{curl}(\vec{F})(a, b, c)$ measures the strength of the rotation in this maximal direction. The strength of rotation around an axis determined by a unit vector $\vec{v}$ is found by calculating $\Big(\text{curl}(\vec{F})(a, b, c)\Big) \cdot \vec{v}$

13.8.6 Exercises

1. Compute the curl of the vector field $\vec{F} = 8z\,\vec{\imath} - 5y\,\vec{\jmath} - 8x\,\vec{k}$.
curl = _____

2. A)
Consider the vector field $F(x, y, z) = (-8yz, 7xz, -xy)$.
Find the divergence and curl of F.
$\text{div}(F) = \nabla \cdot F =$ _____.
$\text{curl}(F) = \nabla \times F =$ (_____, _____, _____).
B)
Consider the vector field $F(x, y, z) = (-3x^2, 2(x + y)^2, 7(x + y + z)^2)$.
Find the divergence and curl of F.
$\text{div}(F) = \nabla \cdot F =$ _____.
$\text{curl}(F) = \nabla \times F =$ (_____, _____, _____).

3. The figure below gives a sketch of a velocity vector field $\vec{F} = (-x - y)\vec{i} + (-x)\vec{j}$ in the xy-plane. Note that vectors are given with a dot at the *tail* of each arrow.



(a) What is the direction of rotation of a thin twig placed at the origin along the x-axis?
(☐ clockwise ☐ counterclockwise ☐ it won't rotate)
(b) What is the direction of rotation of a thin twig placed at the origin along the y-axis?
(☐ clockwise ☐ counterclockwise ☐ it won't rotate)
(c) If we put a paddle wheel at the origin, instead of a twig, in what direction would you expect that to rotate?
(☐ clockwise ☐ counterclockwise ☐ it won't rotate)
(d) Compute curl $\vec{F}$.
curl$\vec{F}$ =_____

4. For each of the following vector fields, find its curl and determine if it is a gradient field.
(a) $\vec{F} = 2\big(xy + z^2\big)\,\vec{\imath} + 4\big(x^2 + yz\big)\,\vec{\jmath} + 4\big(xz + y^2\big)\,\vec{k}$.

curl $\vec{F}$ = _____

$\vec{F}$ (☐ is a gradient field ☐ is not a gradient field)

(b) $\vec{G} = \left(4xz + y^2\right)\,\vec{\imath} + 2xy\,\vec{\jmath} + 2x^2\,\vec{k}$:

curl $\vec{G}$ = _____

$\vec{G}$ (☐ is a gradient field ☐ is not a gradient field)

(c) $\vec{H} = (2xy + yz)\,\vec{\imath} + \left(x^2 + z^2\right)\,\vec{\jmath} + 2xz\,\vec{k}$:

curl $\vec{H}$ = _____

$\vec{H}$ (☐ is a gradient field ☐ is not a gradient field)

5. Let $\mathbf{F} = (5xy, 9y, 8z)$.

The curl of $\mathbf{F}$ = (_____,_____,_____).

Is there a function f such that $\mathbf{F} = \nabla f$? _____ (y/n)

6. Three small squares, S_1, S_2, and S_3, each with side 0.05 and centered at the point $(4, 1, 2)$, lie parallel to the xy-, yz- and xz-planes, respectively. The squares are oriented counterclockwise when viewed from the positive z-, x-, and y-axes, respectively. A vector field $\vec{G}$ has circulation around S_1 of 1, around S_2 of -0.0075, and around S_3 of -1. Estimate curl $\vec{G}$ at the point $(4, 1, 2)$

curl $\vec{G} \approx$ _____

7. Compute the curl of the vector field $\vec{F} = (x + y)\,\vec{\imath} + (z - y)\,\vec{\jmath} + (z + x)\,\vec{k}$

curl = _____

8. Compute the curl of the vector field $\vec{F} = \cos(y)\,\vec{\imath} + \sin(z)\,\vec{\jmath} + 6x^3\,\vec{k}$

curl = _____

9. A vortex that rotates at constant angular velocity w about the z-axis has velocity vector field $\vec{v} = w(-y\vec{i} + x\vec{j})$.

(a) Sketch, on a separate sheet of paper, the vector field with $w = 1$ and the vector field with $w = -1$. Then determine the speed $\|\vec{v}\|$ of the vortex as a function of the distance from its center, r.

speed = _____

(b) Compute div $\vec{v}$ and curl $\vec{v}$.

div $\vec{v}$ = _____

curl $\vec{v}$ = _____

(c) Compute the circulation of $\vec{v}$ counterclockwise about the circle of radius R in the xy-plane, centered at the origin.

circulation = _____

10. Let $\mathbf{F} = 10x\mathbf{i} + 1y\mathbf{j} + 8z\mathbf{k}$. Compute the divergence and curl of F.

A. div $\mathbf{F}$ = _____

B. curl $\mathbf{F}$ = _____

11. Let $\mathbf{F} = (1yz)\,\mathbf{i} + (8xz)\,\mathbf{j} + (2xy)\,\mathbf{k}$. Compute the following:

A. div $\mathbf{F}$ = _____

B. curl $\mathbf{F}$ = _____ **i**+ _____ **j**+ _____ **k**

C. div curl $\mathbf{F}$ = _____

Note: Your answers should be expressions of x, y and/or z; e.g. "3xy" or "z" or "5"

12. In this exercise, we will look at how different forms of component function's dependence on our variables will change the curl of the vector field. For each of the vector fields given below, decide if the curl of the vector field must be zero (as a vector) and write a sentence or two to justify your decision.

(a)
$$\vec{F} = f(x)\hat{\imath} + g(y)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

(b)
$$\vec{G} = f(x)\hat{\imath} + f(y)\hat{k} + \hat{\jmath} + f(z)\hat{\jmath}$$

(c)
$$\vec{H} = f(y,z)\hat{\imath} + g(x,z)\hat{k} + \hat{\jmath} + h(x,y)\hat{\jmath}$$

(d)
$$\vec{A} = g(x,y)\hat{\imath} - g(x,y)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

(e)
$$\vec{B} = g(x,y)\hat{\imath} - g(y,x)\hat{k} + \hat{\jmath} + h(z)\hat{\jmath}$$

13. Let $h(x,y,z)$ be a smooth function of three variables and let $\vec{F} = grad(h) = \nabla h$.

(a) Write out the component functions of $\vec{F}$ in terms of partial derivatives of h.

(b) Write the divergence of $\vec{F}$ in terms of derivatives of h.

(c) Is the divergence of $\vec{F}$ necessarily zero? Explain why or why not.

(d) Calculate $curl(\vec{F})$ in terms of derivatives of h.

(e) Show that $curl(\vec{F}) = \vec{0}$.

13.9 Green's Theorem

Motivating Questions

- How can we calculate the circulation of a two-dimensional vector field $\vec{F}$ around a closed curve when $\vec{F}$ is not path-independent?
- What is the meaning of the double integral of the circulation density of a smooth two-dimensional vector field on a region R bounded by a closed curve that does not intersect itself?

13.9.1 Introduction

We know from Section 13.5 that a vector field is path-independent if and only if the circulation around every closed curve in its domain is 0. It is probably not surprising that, given the multitude of names we have for path-independent vector fields, they are important vector fields that arise frequently. However, not every vector field is path-independent, and many times we will want to calculate the circulation around a closed curve in a vector field that is not path-independent. This section explores a connection between line integrals and double integrals that you may find surprising. It will be the first of three major theorems that connect types of integrals that seem very different on the surface.

Preview Activity 13.9.1 We will consider the vector field $\vec{F} = \langle 2y, 3x^2y \rangle$, which is defined on the entire xy-plane. Suppose that we want to calculate the circulation of $\vec{F}$ around the circle C of radius 2, centered at $(0,0)$, and oriented

counterclockwise.

(a) Verify that $\vec{F}$ is not path-independent by calculating the circulation of $\vec{F}$ around the circle C (Use Theorem 13.5.3). The SageMath cell below is set up to assist you with this, but you will need to supply a parametrization of C on line 4.

```
var('r,t,x,y')
F = vector([2*y,3*x^2*y])
#Supply a parametrization of C on the next line inside
    the [ ]
C = vector([ , ])
integral(F(x=C[0],y=C[1]).dot_product(derivative(C,t)),t,0,2*pi)
```

(b) Recall from Section 13.8 that if $\vec{F} = \langle F_1(x,y), F_2(x,y)\rangle$ is a vector field, then the circulation density (in 2D) is given by

$$\text{Circulation Density at the point } (a,b) = \frac{\partial F_2}{\partial x}(a,b) - \frac{\partial F_1}{\partial y}(a,b)$$

What is the circulation density of $\vec{F} = \langle 2y, 3x^2y\rangle$?

(c) Sketch the curve C and shade the region it bounds. Describe the region bounded by C in both rectangular and polar coordinates.

(d) Calculate the double integral of $\dfrac{\partial F_2}{\partial x} - \dfrac{\partial F_1}{\partial y}$ over the region inside the circle C. This integral is not the most fun to do by hand, so a SageMath cell has been provided to assist you.

```
integral(integral( FUNCTION HERE , var1, start, stop),
    var2, start, stop)
```

Hint. If you choose polar coordinates, use θ for the first variable and r for the second variable.

(e) What do you notice about your results to parts (a) and (d)? Do you think this will happen in general? Write a sentence or two about what you think is happening here.

A natural question after completing the Preview Activity is if there is something special about the vector field $\vec{F}$ or the curve C that led to the results you obtained, and investigating this will be our principal task in this section.

13.9.2 Circulation

In Preview Activity 13.9.1, you integrated the circulation density of a smooth vector field over a disk and found the result was equal to the circulation of the vector field along the region's circular boundary. This relationship is the theme of this section. To see why this would make sense, consider the region R bounded by the curve C shown in Figure 13.9.1. We have placed a square grid inside R to suggest the idea of breaking R up into many smaller regions, most of which are square. This idea should make you think of the methods we have already seen of breaking up a region into smaller and smaller regions for Riemann sums.

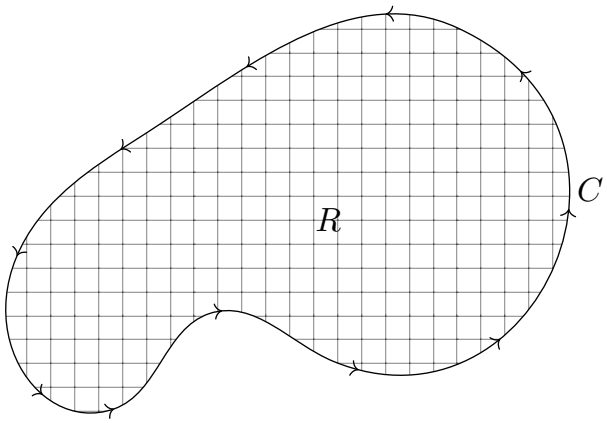


Figure 13.9.1 A region bounded by a curve

The critical idea here is that if we integrate the circulation *density* over each of the small regions and add those up, this is the same as integrating the circulation density over the entirety of the region R because of the fundamental properties of integrals. Also, integrating circulation *per unit area* over a two-dimensional region should be related to the total circulation on that region. Now look at Figure 13.9.2 and think of these two square regions as being two of the square regions inside R in Figure 13.9.1.

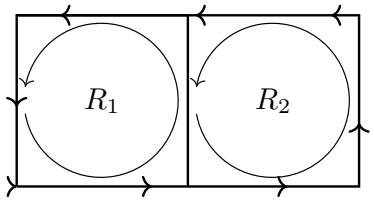


Figure 13.9.2 A square region divided into two smaller squares

We orient the boundary C_i of square region R_i in the manner suggested by the circular arrows. This means that the vertical boundary in common between R_1 and R_2 is oriented *up* when we calculate the line integral $\oint_{C_1} \vec{F} \cdot d\vec{r}$ and oriented *down* when we evaluate $\oint_{C_2} \vec{F} \cdot d\vec{r}$. Thus, this line segment does not contribute to the sum $\oint_{C_1} \vec{F} \cdot d\vec{r} + \oint_{C_2} \vec{F} \cdot d\vec{r}$. Therefore,

$$\oint_{C_1} \vec{F} \cdot d\vec{r} + \oint_{C_2} \vec{F} \cdot d\vec{r}$$

is equal to the line integral of $\vec{F}$ along the boundary of the large rectangle.

Returning to Figure 13.9.1, if we find the circulation along the boundary of each of the smaller regions, the line integrals along the boundaries that lie inside the region R will all offset. Thus the value should equal $\oint_C \vec{F} \cdot d\vec{r}$. If we make our grid fine enough, all of the smaller regions into which R is divided will be very close to rectangular.

The calculation of circulation around a nice closed curve will be the sum of the circulation over pieces that enclose the same area, regardless of the size or number of the pieces used. In other words,

$$\oint_C \vec{F} \cdot d\vec{r} = \sum_{R_i} \oint_{C_i} \vec{F} \cdot d\vec{r}$$

where C_i is a bounding curve for subregion R_i. We defined our circulation density as the ratio of circulation to area, so we can also substitute the circulation around a region for the circulation density times the area of that region. Additionally, we can take the limit as our regions get smaller, since size and the number of regions does not affect the sum. Letting ΔR_i represent the area of region R_i with boundary curve C_i, we have

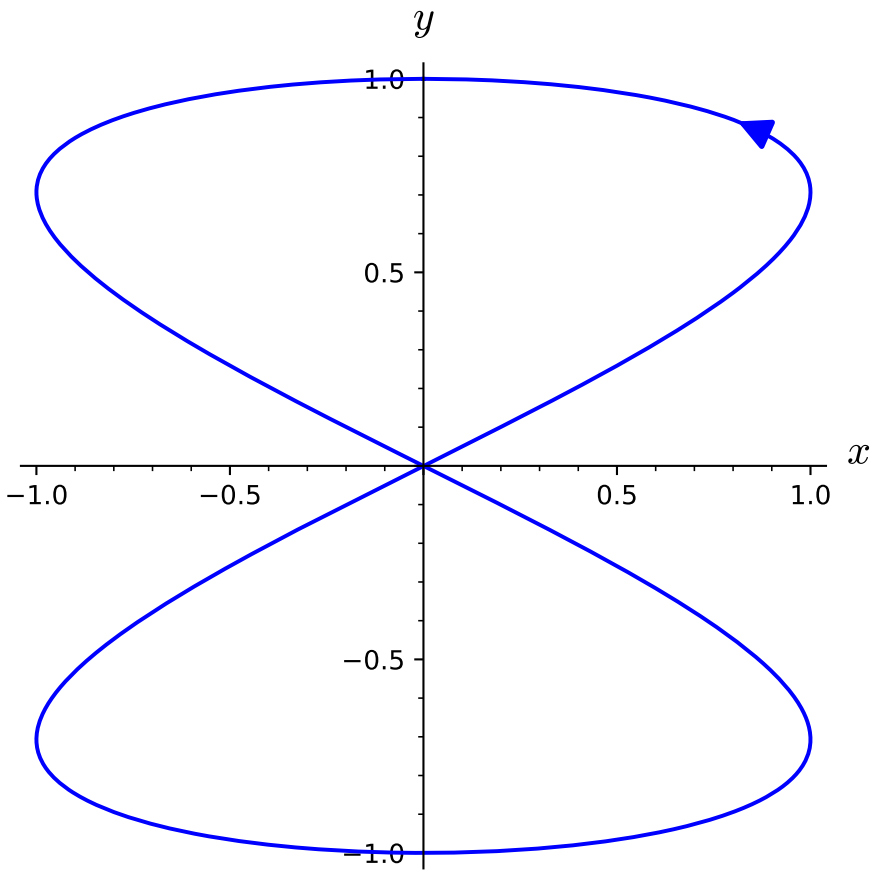
$$\begin{aligned}\oint_C \vec{F} \cdot d\vec{r} &= \lim_{\Delta R_i \to 0} \sum_{R_i} \oint_{C_i} \vec{F} \cdot d\vec{r} \\ &= \lim_{\Delta R_i \to 0} \sum_{R_i} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \Delta R_i.\end{aligned}$$

This expression should look familiar as the Riemann sum definition for the double integral of $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ over the region R. For "nice" closed curves we can equate the circulation around the boundary (computed as a line integral) with a double integral of the circulation density over the region enclosed (computed with a double integral).

13.9.3 Green's Theorem

So far in this section, we have restricted ourselves to relatively nice closed curves when thinking about circulation. While the main theorem of this section will not allow us to consider arbitrary closed curves, it does cover more varied curves than we have discussed so far. A **simple closed curve** is a closed curve that does not cross itself, and these are the curves to which our next theorem applies.

A closed curve that is not simple. The restriction that the curve in Green's Theorem prohibits curves such as the one below, which crosses itself.



Theorem 13.9.3 Green's Theorem. *Let C be a simple closed curve in the plane that bounds a region R with C oriented in such a way that when walking along C in the direction of its orientation, the region R is on our left. Suppose that $\vec{F} = \langle F_1, F_2 \rangle$ is vector field with continuous partial derivatives on the region R and its boundary C. The circulation of $\vec{F}$ along C may be calculated as*

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA.$$

At first glance, it may seem that Green's Theorem is of purely intellectual interest. However, we have already encountered situations where parameterizing a curve can be complicated. This is particularly true when the curve has "corners" that require us to give separate parameterizations for several pieces of the curve, such as with the rectangular curve pictured in Figure 13.9.4. However, as with this rectangle, is often the case that a curve that is difficult to parametrize bounds a region that is not too complex to describe using rectangular or polar coordinates. For instance, the rectangular region in Figure 13.9.4 can be described as $1 \leq x \leq 4$ and $2 \leq y \leq 4$, which is a simpler description than needing to parameterize each of the four sides of the rectangle separately. Additionally, the integrand of the double integral in Green's Theorem involves partial derivatives that can sometimes result in an integrand that is easy to work with. The purpose of Green's Theorem is, at its core, to allow you to exchange one type of integration problem (a line integral) for another type of integration problem (a double integral). This will be a recurring theme as this chapter continues.

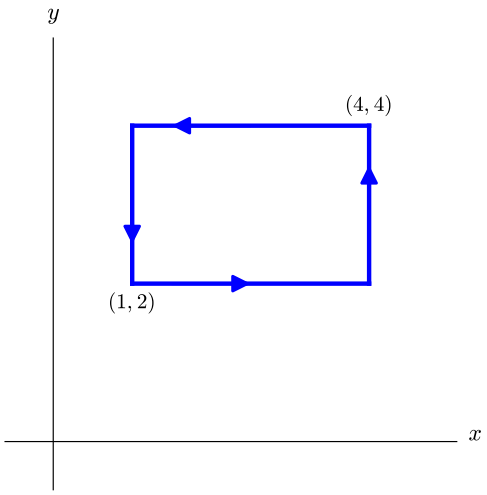


Figure 13.9.4 A rectangular curve

Example 13.9.5 We will verify Green's Theorem by computing both sides of the equation separately when $\vec{F} = \langle x^2 - xy, y^2 - 2x \rangle$ and C is the bounding curve given in Figure 13.9.4.

First, we calculate the double integral (the right hand side of the equation in Green's Theorem). The circulation density of $\vec{F} = \langle x^2 - xy, y^2 - 2x \rangle$ will be $-2 + x$, which is continuous everywhere in the region R. Our iterated integral becomes

$$\begin{aligned} \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA &= \int_2^4 \int_1^4 (-2 + x)\, dx\, dy \\ &= \int_2^4 \left[-2x + x^2/2 \right]_{x=1}^{x=4} dy \\ &= \int_2^4 \frac{3}{2}\, dy \\ &= 3 \end{aligned}$$

Note how easy it is to set up and evaluate this double integral.

We will need to compute the line integral of $\vec{F}$ over the four segments of the bounding rectangle. We will use the following parameterizations:

$$\begin{aligned} \vec{r}_{\text{bottom}}(t) &= \langle 1 + 3t, 2 \rangle \\ \vec{r}_{\text{right}}(t) &= \langle 4, 2 + 2t \rangle \\ \vec{r}_{\text{top}}(t) &= \langle 4 - 3t, 4 \rangle \\ \vec{r}_{\text{left}}(t) &= \langle 1, 4 - 2t \rangle \end{aligned}$$

where each parameterization uses $0 \le t \le 1$. These parameterizations will yield the following line integrals:

$$\begin{aligned} \int_{\text{bottom}} \vec{F} \cdot d\vec{r} &= \int_0^1 ((1 + 3t)^2 - (1 + 3t)(2))3\, dt = 6 \\ \int_{\text{right}} \vec{F} \cdot d\vec{r} &= \int_0^1 ((2 + 2t)^2 - (2)(4))2\, dt = 8/3 \\ \int_{\text{top}} \vec{F} \cdot d\vec{r} &= \int_0^1 ((4 - 3t)^2 - (4 - 3t)(4))(-3)\, dt = 9 \end{aligned}$$

$$\int_{\text{left}} \vec{F} \cdot d\vec{r} = \int_0^1 ((4-2t)^2 - (2)(1))(-2)\, dt = -44/3$$

Note that the sum of these line integrals is 3, and thus Green's Theorem is verified for our case since

$$\oint_C \vec{F} \cdot d\vec{r} = 3 = \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA$$

▲

Activity 13.9.2 For each of the curves described below, find the circulation of the given vector field around the curve. Do this both by calculating the line integral directly as well as by calculating the double integral from Green's Theorem.

(a) The curve C_1 is the circle of radius 3 centered at the point $(2, 1)$ (oriented counterclockwise) and the vector field is $\vec{F} = \langle y^2, 5x + 2xy \rangle$.

(b) The curve C_2 is the triangle with vertices $(0, 0)$, $(3, 0)$, and $(3, 3)$ (oriented counterclockwise) and the vector field is $\vec{G} = \langle y^2, 3xy \rangle$.

Generally, Green's Theorem is useful for allowing us to calculate line integrals by instead calculating a double integral. Going the other direction is harder, since finding a vector field $\langle F_1, F_2 \rangle$ so that $\partial F_2/\partial x - \partial F_1/\partial y$ is equal to the integrand in the double integral can be difficult. However, the exercises will explore some situations where calculating a suitable line integral is a viable alternative to the double integral.

13.9.4 What happens when vector fields are not smooth?

Notice that the assumptions in Green's Theorem require that the region R be bounded by a simple closed curve C and that the vector field have continuous partial derivatives on R and C. In Exercise 13, we will explore what happens when the region R cannot be bounded by a simple closed curve, as sometimes multiple applications of Green's Theorem can be used in those circumstances. Now, however, we will take a look at what happens in some cases where the vector field $\vec{F}$ is not smooth.

Activity 13.9.3 Consider the vector field $\vec{F} = \dfrac{-y}{x^2+y^2}\hat{\imath} + \dfrac{x}{x^2+y^2}\hat{\jmath}$. Notice that $\vec{F}$ is smooth everywhere in the plane other than at the point $(0, 0)$. This vector field is plotted in Figure 13.9.6, but we have not plotted the vectors close to the origin as their magnitudes get so large that they make it hard to interpret the figure. Here Green's Theorem applies to any simple closed curve C that neither passes through $(0, 0)$ nor bounds a region containing $(0, 0)$.

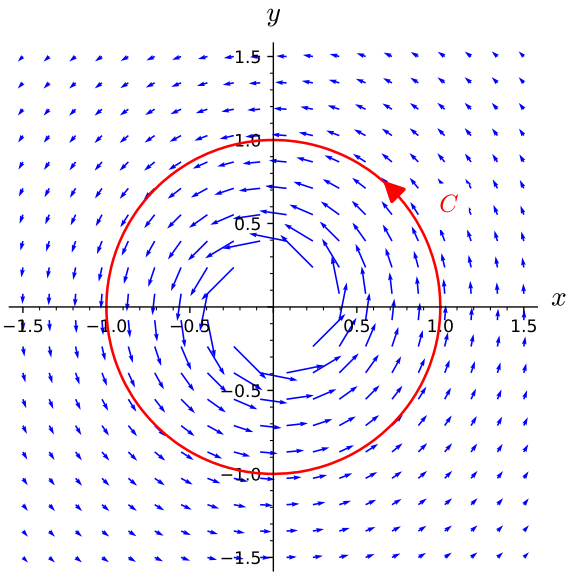


Figure 13.9.6 The vector field $\vec{F}$

(a) Find the circulation density of $\vec{F}$ (i.e., the integrand of the double integral in Green's Theorem).

Hint. Make sure to note any exceptional points at which your formula is not valid.

(b) Suppose that C is the unit circle centered at the origin. Without doing any calculations, what can you say about $\oint_C \vec{F} \cdot d\vec{r}$? What does this tell you about if $\vec{F}$ is path-independent?

(c) What would you get if you integrated the circulation density of $\vec{F}$ over the region bounded by C?

(d) Do the previous two parts contradict Green's Theorem? Explain your reasoning.

(e) Is the vector field $\vec{G} = \dfrac{x}{x^2+y^2}\hat{\imath} + \dfrac{y}{x^2+y^2}\hat{\jmath}$, which is shown in Figure 13.9.7, path-independent? Why or why not?

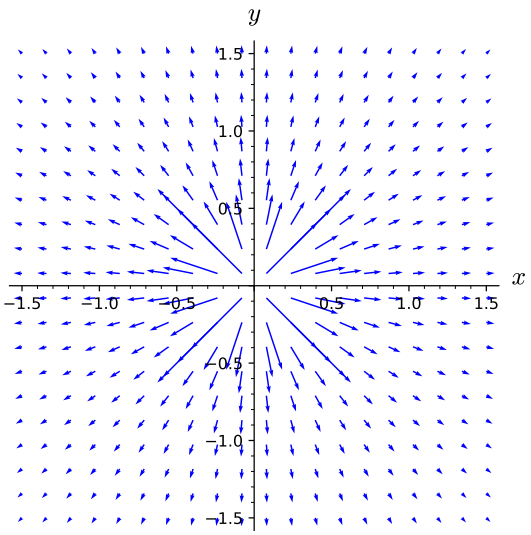


Figure 13.9.7 The vector field $\vec{G}$

(f) Suppose that C is the unit circle centered at the origin. Find $\oint_C \vec{G} \cdot d\vec{r}$. Can you do this using Green's Theorem?

We can now see that Green's Theorem is a powerful tool. However it cannot be used for all line integrals. In particular, the restriction that the vector field be smooth on the entire region bounded by a simple closed curve C creates limitations. This can occur even in cases where holes in the vector field's domain or points where the vector field is not continuously differentiable lie away from C.

Summary

- Green's Theorem tells us that we can calculate the circulation of a smooth vector field along a simple closed curve that bounds a region in the plane on which the vector field is also smooth by calculating the double integral of the circulation density instead of the line integral.
- Integrating the circulation density of a smooth vector field on a region bounded by a simple closed curve gives the same value as calculating the circulation of the vector field around the boundary of the region with suitable orientation.

13.9.5 Exercises

1. Calculate $\int_C \left(3(x^2 - y)\vec{i} + 4(y^2 + x)\vec{j}\right) \cdot d\vec{r}$ if:

(a) C is the circle $(x-5)^2 + (y-5)^2 = 9$ oriented counterclockwise.

$\int_C \left(3(x^2 - y)\vec{i} + 4(y^2 + x)\vec{j}\right) \cdot d\vec{r} =$ _____

(b) C is the circle $(x-a)^2 + (y-b)^2 = R^2$ in the xy-plane oriented counterclockwise.

$\int_C \left(3(x^2 - y)\vec{i} + 4(y^2 + x)\vec{j}\right) \cdot d\vec{r} =$ _____

2. Let $\vec{F} = 6xe^y\,\vec{i} + 3x^2e^y\,\vec{j}$ and $\vec{G} = 6(x-y)\,\vec{i} + 3(x+y)\,\vec{j}$. Let C be the path consisting of lines from $(0,0)$ to $(4,0)$ to $(4,2)$ to $(0,0)$. Find each of the following integrals exactly:

(a) $\int_C \vec{F} \cdot d\vec{r} =$ _____

(b) $\int_C \vec{G} \cdot d\vec{r} =$ _____

3. Let $\vec{F} = (9x^2y + 3y^3 + 5e^x)\,\vec{i} + (2e^{y^2} + 144x)\,\vec{j}$. Consider the line integral of $\vec{F}$ around the circle of radius a, centered at the origin and traversed counterclockwise.

(a) Find the line integral for $a = 1$.

line integral = _____

(b) For which value of a is the line integral a maximum?

$a =$ _____

(Be sure you can explain why your answer gives the correct maximum.)

4. ***(a)*** Show that each of the vector fields $\vec{F} = 3y\,\vec{i} + 3x\,\vec{j}$, $\vec{G} = \frac{2y}{x^2+y^2}\,\vec{i} + \frac{-2x}{x^2+y^2}\,\vec{j}$, and $\vec{H} = \frac{5x}{\sqrt{x^2+y^2}}\,\vec{i} + \frac{5y}{\sqrt{x^2+y^2}}\,\vec{j}$ are gradient vector fields on some domain (not necessarily the whole plane) by finding a potential function for each.

For $\vec{F}$, a potential function is $f(x, y) =$ _____

For $\vec{G}$, a potential function is $g(x, y) =$ _____

For $\vec{H}$, a potential function is $h(x, y) =$ _____

(b) Find the line integrals of $\vec{F}, \vec{G}, \vec{H}$ around the curve C given to be the unit circle in the xy-plane, centered at the origin, and traversed counterclockwise.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

$\int_C \vec{G} \cdot d\vec{r} =$ _____

$\int_C \vec{H} \cdot d\vec{r} =$ _____

(c) For which of the three vector fields can Green's Theorem be used to calculate the line integral in part (b)?

It may be used for (☐ only F ☐ only G ☐ only H ☐ F and G ☐ F and H ☐ G and H ☐ F, G and H)

(Be sure that you are able to explain why or why not.)

5. Let C be the positively oriented square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$, $(0, 1)$. Use Green's Theorem to evaluate the line integral $\int_C 7y^2x\,dx + 1x^2y\,dy$.

6. Let $\mathbf{F} = -4y\mathbf{i} + 4x\mathbf{j}$. Use the tangential vector form of Green's Theorem to compute the circulation integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ where C is the positively oriented circle $x^2 + y^2 = 16$.

7. Use Green's Theorem to calculate the circulation of $\vec{F} = xy\,\vec{i}$ around the rectangle $0 \le x \le 7$, $0 \le y \le 3$, oriented counterclockwise.

circulation = _____

8. Use Green's Theorem to calculate the circulation of $\vec{F} = 8y\,\vec{i} + 8xy\,\vec{j}$ around the unit circle, oriented counterclockwise.

circulation = _____

9. Calculate $\int_C ((8x + 2y)\,\vec{i} + (4x + 6y)\,\vec{j}) \cdot d\vec{r}$ where C is the circular path with center (a, b) and radius m, oriented counterclockwise. Use Green's Theorem.

$\int_C ((8x+2y)\,\vec{i} + (4x+6y)\,\vec{j}) \cdot d\vec{r} =$ _____

10. Let $\vec{F} = 3(x + y)\,\vec{i} + 3\sin(y)\,\vec{j}$. Find the line integral of $\vec{F}$ around the perimeter of the rectangle with corners $(3, 0)$, $(3, 4)$, $(-3, 4)$, $(-3, 0)$, traversed in that order.

line integral = _____

11. Use Green's Theorem to evaluate the line integral $\oint_{\mathcal{C}} y^8\,dx + x^4\,dy$ where $\mathcal{C}$ is the boundary of the square $-1 \le x \le 1$, $-1 \le y \le 1$ oriented counterclockwise.

$\oint_{\mathcal{C}} y^8\,dx + x^4\,dy =$ _____

12. Sometimes, the value in Green's Theorem is in converting a double integral into a line integral. Recall that the area of a region R in the plane can be found by calculating $\iint_R 1\,dA$.

(a) Find a smooth vector field $\vec{F}$ such that the circulation density of $\vec{F}$ is 1 everywhere.

Hint. There are several ways to choose $\vec{F} = \langle F_1, F_2\rangle$ so that $\partial F_2/\partial x - \partial F_1/\partial y = 1$.

(b) Write out both integrals in Green's Theorem for the vector field you selected in part a.

(c) For positive real numbers a, b, the ellipse $x^2/a^2 + y^2/b^2 = 1$ can be parametrized as $\vec{r}(t) = \langle a\cos(t), b\sin(t)\rangle$ with $0 \le t \le 2\pi$. Find the area of this ellipse by calculating a line integral.

(d) Find another vector field $\vec{G}$ (different from the one you found in part a) that has circulation density 1 everywhere, and calculate the area of the ellipse in part c by calculating the circulation of $\vec{G}$ along the ellipse.

(e) How many possible vector fields are there with circulation density 1 everywhere? Why?

13. The types of regions in $\mathbb{R}^2$ to which Green's Theorem applies are formally called **simply connected regions**. To be precise, a simply connected region D in $\mathbb{R}^2$ is a set so that there is a path between every pair of points in D that stays inside D and any simple closed curve in D can be shrunk to a point while remaining inside D. We can think of a simply connected region as being a region that does not have any "holes". There are many instances where we can find a way to apply Green's Theorem multiple times to work with regions that are not simply connected, however.

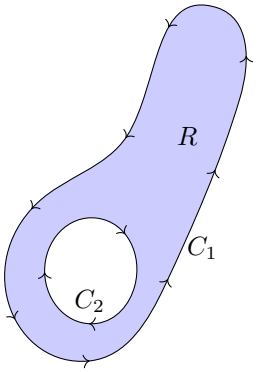


Figure 13.9.8 A region that is not simply connected

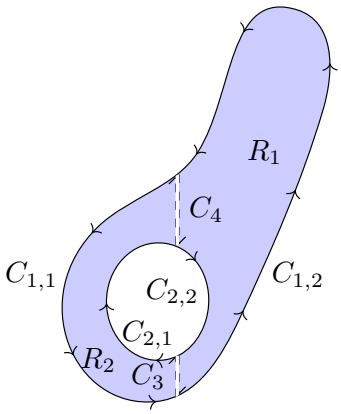


Figure 13.9.9 Breaking up a region

(a) Explain why the region R in Figure 13.9.8 is not simply connected.

(b) If we let $C = C_1 + C_2$, then the orientation of C is as required by

Green's Theorem—when walking along the pieces of C, the region R is always on the left-hand side. Assume that $\vec{F}$ is a vector field that is smooth on R and C. Using Figure 13.9.9 as a guide, write $\oint_C \vec{F} \cdot d\vec{r}$ as a sum of double integrals, one over R_1 and the other over R_2.

(c) In Activity 13.9.3, we considered the vector field $\vec{F} = \frac{-y}{x^2+y^2}\hat{\imath} + \frac{x}{x^2+y^2}\hat{\jmath}$ and found that the circulation of $\vec{F}$ around the unit circle centered at the origin (oriented counterclockwise) was 2π. Show that for every simple closed curve C in the plane that bounds a region containing the origin, $\oint_C \vec{F} \cdot d\vec{r} = 2\pi$.

Hint. For part c, create a region R as in Figure 13.9.8 using the curve C in place of C_1.

14. This exercise presents another occasion where Green's Theorem can be used to convert a double integral into a line integral.

(a) Recall from Section 12.5 that the centroid of a lamina D of area A is given by

$$\overline{x} = \frac{1}{A}\iint_D x\,dA \qquad \text{and} \qquad \overline{y} = \frac{1}{A}\iint_D y\,dA.$$

Find vector fields $\vec{F}$ and $\vec{G}$ so that

$$\overline{x} = \oint_C \vec{F} \cdot d\vec{r} \qquad \text{and} \qquad \overline{y} = \oint_C \vec{G} \cdot d\vec{r},$$

where C is the boundary of the lamina D.

(b) Find the centroid of the triangle with vertices $(0,0)$, $(a,0)$, and $(0,b)$ for real numbers $a, b > 0$.

13.10 Parameterizations of Surfaces and Surface Area

Motivating Questions

- What does it mean to parameterize a surface?
- How can parameterizations of curved surfaces be used to compute surface area?

13.10.1 Introduction

In Section 9.6 and Chapter 10, we used parameterizations of curves as a means to use the many tools of single-variable calculus to analyze the curve. One of the central goals of this section is to extend the idea of parameterizations to surfaces in an attempt to apply the many *multivariable* tools of Chapter 11 and Chapter 12. In the preview activity, we look at how geometric knowledge of right circular cylinders and cones allows us to create a vector-valued function that describes the points on the surface of a cone.

Preview Activity 13.10.1 Recall the standard parameterization of the unit circle given by

$$x(t) = \cos(t) \qquad \text{and} \qquad y(t) = \sin(t),$$

where $0 \leq t \leq 2\pi$.

(a) Find a parameterization of a circle of radius of 2 in $\mathbb{R}^3$ centered at $(0, 0, 1)$ and lying in the plane $z = 1$. Be sure to give the bounds on the parameter as well.

Hint. The z-coordinate is constant, so only x and y will change in the parameterization.

(b) Find a parameterization of a circle of radius of 2 in $\mathbb{R}^3$ centered at $(0, 0, -5)$ and lying in the plane $z = -5$. Be sure to give the bounds on the parameter as well.

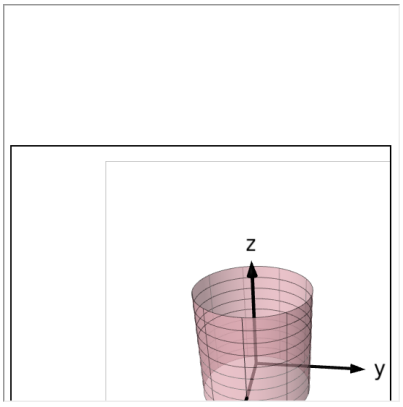
(c) Find a parameterization of a circle of radius of 2 in $\mathbb{R}^3$ centered at $(0, 0, a)$ and lying in the plane $z = a$. Be sure to give the bounds on the parameter as well.

(d) A right circular cylinder surface centered on the z-axis, such as the one shown in Figure 13.10.1, can be thought of as being built from a stack of circles of fixed radius centered at points along the z-axis. Find a parameterization of a right circular cylinder surface of radius 2, centered on the z-axis with heights from b to c. In other words, express the x-, y-, and z-coordinates of points on the right circular cylinder in terms of two variables. Use the parameter s to measure the height of the circle above or below the xy-plane and use the parameter t to measure the location around the circle centered on the z-axis. Use the previous part to write out each of the following:

$$x(s,t) = \qquad\qquad y(s,t) = \qquad\qquad z(s,t) =$$

with the bounds on s and t given by

$$\leq s \leq \quad \text{and} \quad \leq t \leq$$



Standalone
In Context
Embed

Figure 13.10.1 A right circular cylinder centered on the z-axis

(e) Just as a right circular cylinder can be viewed as a "stack" of circles of fixed radius, a cone can be viewed as a stack of circles with *varying* radius. In particular, the cone $z^2 = x^2 + y^2$ has contours that are circles centered on the z-axis with radius changing according to the rule $r = z$. Find a parameterization of the cone $z^2 = x^2 + y^2$ with heights from b to c. In

other words, we want to express the x, y, and z coordinates of points on the cone in terms of two variables. We will let the parameter s measure the height of the circle (in our stack) above or below the xy-plane and let the parameter t measure the location around the circle centered on the z-axis. Write out each of the following:

$$x(s,t) = \qquad y(s,t) = \qquad z(s,t) =$$

with the bounds on s and t given by

$$\leq s \leq \qquad \text{and} \qquad \leq t \leq$$

13.10.2 Parameterizations of Surfaces

Preview Activity 13.10.1 concluded with a description of a cone in terms of measurements corresponding to the angle around the z-axis and the position relative to the xy-plane. Geometric intuition described the coordinate relationships of the two surfaces in Preview Activity 13.10.1. This subsection discusses how to use algebraic and geometric tools to find parameterizations of curved surfaces. The next subsection describes some important geometric measurements of a surface that can be done using a parameterization. The section concludes by examining how parameterizations can be used to compute the surface area of a surface in three dimensions.

In a single-variable setting, any function may have its graph expressed parametrically. For instance, the graph of $y = g(x)$ can be parameterized as $\langle t, g(t)\rangle$, where t belongs to the domain of g. Certain curves that are not expressible with y as a function of x can also be represented parametrically. For instance, the unit circle cannot be represented with either the x- or y-coordinate written as a function of the other. However, it can be parameterized as $\langle \cos(t), \sin(t)\rangle$, where $0 \leq t < 2\pi$.

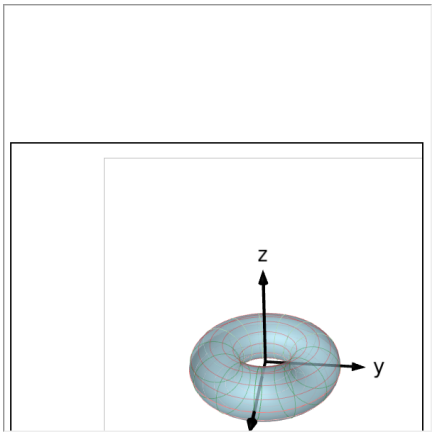
A surface of the form $z = f(x,y)$ can be expressed parametrically by

$$\langle x(s,t), y(s,t), z(s,t)\rangle = \langle s, t, f(s,t)\rangle,$$

where (s,t) varies over the entire domain of f. In the parameterization above, we can think of s as acting like x and t as acting like y. Therefore, any familiar surface expressed as $z = f(x,y)$ can be written as a parametric surface. The greater power of parameterizations is realized when dealing with surfaces such as a sphere that *cannot* be expressed by a single function $z = f(x,y)$ but *can* be represented parametrically.

For surfaces such as ellipsoids or hyperboloids where one coordinate cannot be expressed as a function of the other two, a good strategy is to use geometric knowledge of these surfaces to express each of the x-, y-, and z-coordinates in terms of two measurements/parameters. In Preview Activity 13.10.1, the parameters measured the rotational coordinate around the z-axis and the height above or below the xy-plane (along with the radial coordinate in the cone). In the next example, we will look at how to parameterize a torus using this same strategy.

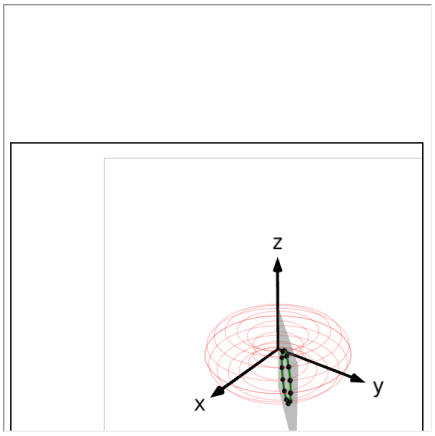
Example 13.10.2 In this example, we want to parameterize the torus shown in Figure 13.10.3.



Standalone
In Context
Embed

Figure 13.10.3 The surface of a torus shown centered at the origin with hole oriented around the z-axis

To describe each point on the torus in terms of two parameters, we can think of the torus as being generated by rotating a circle or radius a around the z-axis. Looking at a radial slice from the z-axis out, as shown by the gray plane in Figure 13.10.4, we generate the torus by revolving the points on the green circle around the z-axis, as shown by the red circles. This means that we can describe each point on the torus in terms of its location along the green circle and its location along a red circle, representing rotation around the z-axis.



Standalone
In Context
Embed

Figure 13.10.4 A slice of a torus shown as the revolution of a circle (shown in green) around the z-axis.

The green slice of the torus corresponds to a constant value of the cylindrical coordinate θ. We will use the t parameter to describe position in terms of rotation around the z-axis; in other words, the parameter t acts as the θ-coordinate from cylindrical coordinates. We will use the other parameter s to describe the location of a point on the green circle.

Looking at the slice of the torus shown by the grey plane in Figure 13.10.4 reveals a two-dimensional plot as shown in Figure 13.10.5. Points on the circle are described in r and z in terms of the parameter s. The green circle in the rz-plane is described by $r(s) = b + a\cos(s)$ and $z(s) = \sin(s)$ with $0 \leq s \leq 2\pi$. This is a typical parameterization of a circle with a horizontal translation by a.

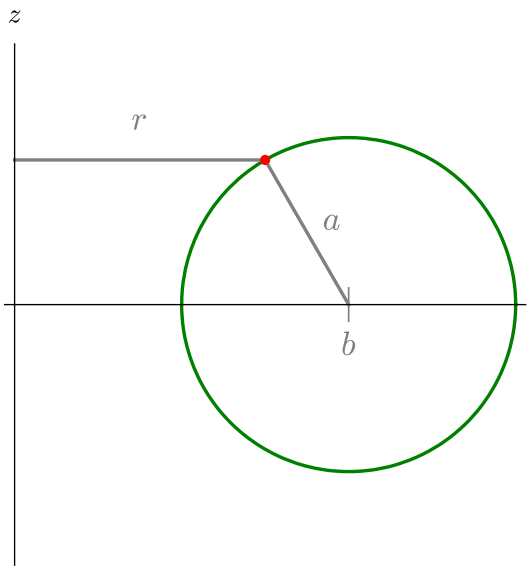


Figure 13.10.5 A circle of radius a centered at the point $(b, 0)$

These cylindrical coordinate transformations lead to the x, y, and z-coordinates in terms of parameters s and t. The cylindrical coordinates of the points of the torus can be written in terms of s and t as

$$\theta = t, \quad r = b + a\cos(s), \quad z = a\sin(s)$$

Therefore, the rectangular coordinates of the points on the torus can be written in terms of s and t as

$$\begin{aligned} x &= r\cos(\theta) & &= (b + a\cos(s))\cos(t) \\ y &= r\sin(\theta) & &= (b + a\cos(s))\sin(t) \\ z &= z & &= a\sin(s) \end{aligned}$$

Just as in the case of parameterizing a curve in space using a vector-valued function of one variable, a parametric description of a surface requires bounds on the parameters used to describe the surface. Remember that for a point on the torus, the t-parameter describes the location in terms of rotation around the z-axis and the s-parameter describes the point's location around the circular slice (in the rz-plane). Both the s and t parameters are bounded below by 0 and above by 2π. Therefore, a parametric description for the torus using a vector-valued function of two variables can be written as

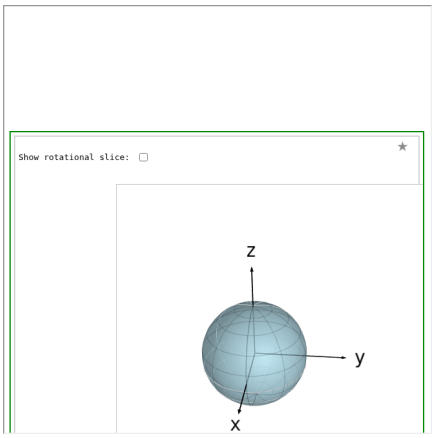
$$\begin{aligned} \vec{r}(s,t) =& \langle x(s,t), y(s,t), z(s,t)\rangle \\ =& \langle (b + a\cos(s))\cos(t), (b + a\cos(s))\sin(t), a\sin(s)\rangle \end{aligned}$$

for $s, t \in [0, 2\pi)$. ▲

The next activity finds two ways to parameterize a sphere centered at the origin. The first views the sphere as a surface of revolution around the z-axis, similar to the preceding example. The second uses spherical coordinates.

Activity 13.10.2 The goal of this activity is to find a parametrization of the sphere of radius R centered at the origin, as shown in Figure 13.10.6. Check the box at the top of the plot to highlight what a semicircular slice of the sphere looks like along a slice with a constant θ coordinate. Notice that this

slice shows how the sphere can be made by rotating this half circle around the z-axis.



Standalone
In Context
Embed

Figure 13.10.6 A plot of a sphere of radius R centered at the origin

(a) One approach to parameterizing the sphere of radius R is similar to the approach to parameterizing the torus as a surface of revolution around the z-axis. In other words, let the parameter t represent the location as a rotation around the z-axis. This allows for parameterization the sphere as

$$x(s,t) = r(s)\cos(t), \quad y(s,t) = r(s)\sin(t), \quad z(s,t) = z(s), \quad (13.10.1)$$

where $r(s)$ and $z(s)$ describe the other cylindrical coordinates in terms of another parameter.

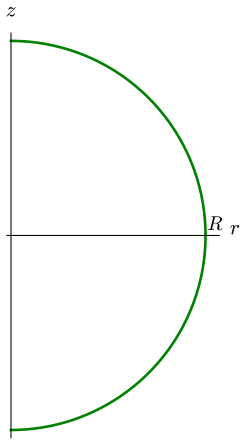


Figure 13.10.7 A plot of a half circle in the rz-plane representing the slice of the sphere rotated around the z-axis

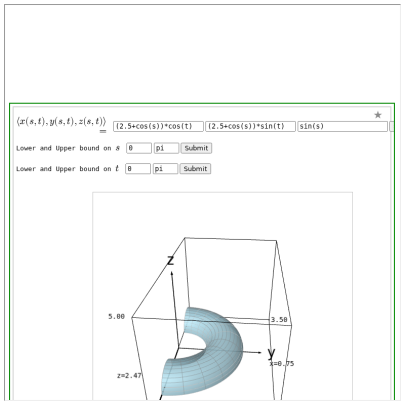
We want to describe the points on the green half circle shown in Figure 13.10.7 in terms of the parameter s. Remember that there is not a unique way to parameterize a curve. Give a parameterization of the green half circle in Figure 13.10.7 as a function of the parameter s. Be sure to state the bounds on s.

Hint. Think of the sphere as a surface of rotation and write r as a function of z. In other words, how large is the radius of the revolution around the z-axis in terms of the z coordinate?

(b) Combine your answer to the previous part with (13.10.1) to create a parameterization of the sphere of radius R. Substitute your answer for the previous part into (13.10.1), writing your answer as parametric functions for x, y, and z. Be sure to state the bounds on the parameters s and t.

$$x(s,t) =$$
$$y(s,t) =$$
$$z(s,t) =$$
$$\leq s \leq$$
$$\leq t \leq$$

(c) Use Figure 13.10.8 to verify that your parameterization for the previous part plots the sphere of radius 3 centered at the origin. Remember to adjust the upper and lower bounds of s and t to match your parameterization.



Standalone
In Context
Embed

Figure 13.10.8 An interactive plot where readers can input their parameterization and bounds

(d) Next we use spherical coordinates to create a parameterization of the sphere of radius R. We want to express the rectangular coordinate equation $x^2+y^2+z^2 = R^2$ in spherical coordinates. State an equation in spherical coordinates that describes the same set of points as $x^2+y^2+z^2 = R^2$.

Hint. This equation is of the form $\rho =$ constant

(e) Surfaces of the form $z = f(x, y)$ were easily parameterized because each of the coordinates can be expressed in terms of two variables: x and y. We can apply a similar idea here because in spherical coordinates, the sphere has one spherical coordinate fixed and the other two spherical coordinates can be used as parameters. Write a couple of sentences to describe which spherical coordinate s will act as and which coordinate t will act as. Include bounds on these parameters as part of your descriptions.

(f) To create a spherical coordinate parameterization of $x^2 + y^2 + z^2 = R^2$, substitute the constant value for the appropriate spherical coordinate and s or t, as determined by your choices in the previous part, for the other two spherical coordinates into the following transformation equations:

$$x = \rho\cos(\theta)\sin(\phi) \qquad y = \rho\sin(\theta)\sin(\phi) \qquad z = \rho\cos(\phi)$$

This gives equations for x, y, and z in terms of the parameters s and t, as well as bounds on both s and t:

$x(s,t) =$ $\quad y(s,t) =$ $\quad z(s,t) =$

with the bounds on s and t given by

$$\quad \leq s \leq \quad \text{and} \quad \leq t \leq \quad .$$

(g) Use Figure 13.10.8 to verify that your parameterization does produce the sphere of radius 3.

The previous activity demonstrates that it is sometimes possible to parameterize a surface in multiple ways. The parameterizations for the sphere in the activity depended on whether you viewed the sphere as a surface of revolution or whether you used spherical coordinates. There are many other ways to parameterize a sphere, but these are likely the two most familiar and common ideas.

This section's exercises ask you to generate and understand the parameterization of the quadric surfaces and cylinder surfaces described in Section 9.8. Additionally, the exercises ask you to understand how shifts and stretches can be applied to each rectangular coordinate separately.

13.10.3 The Geometry of a Parametric Surface

The parameterization of a surface allows us to access important geometric information about the surface. In this subsection, we consider a surface S_1, given by a parameterization $\vec{r}(s,t) = \langle x(s,t), y(s,t), z(s,t)\rangle$ with $a \leq s \leq b$ and $c \leq t \leq d$.

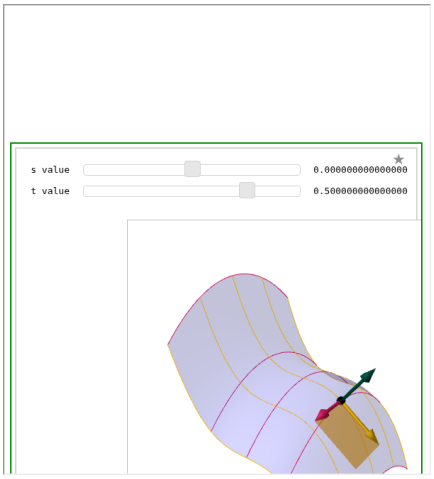
In Subsection 11.2.4, we considered a function $f(x,y)$ and saw how holding one of the input coordinates constant allowed us to restrict the graph of $z = f(x,y)$ to a single-variable slice, which we called the trace. For example, we looked at the trace along $x = a$, which is given by $z = f(a,y)$. Refer to Figure 11.2.11 for an illustration of this. Focusing on a single slice allowed us to define partial derivatives and many other tools in Chapter 11.

What happens when we hold a parameter for a parametric surface constant and allow the other to vary? This gives a slice of the parametric surface, but the slice is not necessarily in the direction of one of the coordinates (x, y, or z). Instead, the slice describes the surface in a constant direction of one parameters, s or t. The trace given by holding s constant yields $\vec{r}(a,t)$, a curve parameterized as a vector-valued function of one variable. This allows the use of the tools from Chapter 10 to understand this curve. Specifically, the derivative,

$$\frac{d\vec{r}}{dt}(a,t) = \langle \frac{dx}{dt}(a,t), \frac{dy}{dt}(a,t), \frac{dz}{dt}(a,t)\rangle$$

gives a vector that is tangent to the curve given by $\vec{r}(a,t)$. Similarly, we can look at the trace generated in the direction of a constant value of t. This gives a trace parameterized as $\vec{r}(s,b)$ with a tangent vector $\frac{d\vec{r}}{ds}(s,b)$.

Figure 13.10.9 shows plot of a parametric surface given by $\vec{r}(s,t)$ for $s_1 \leq s \leq s_2$ and $t_1 \leq t \leq t_2$ plotted as a light blue surface. There are five constant values of t plotted along the surface in shown by the magenta curves and five constant values of s plotted along the surface as shown by the yellow curves. The tangent vectors given by $\frac{d\vec{r}}{ds}(s,b)$ and $\frac{d\vec{r}}{dt}(a,t)$ are shown at a particular point on the surface. Use the sliders at the top of Figure 13.10.9 to change the location at which the tangent vectors are shown.



Standalone
In Context
Embed

Figure 13.10.9 A three-dimensional plot of a parametric surface with traces corresponding to constant s values highlighted in yellow and constant t values highlighted in magenta. The tangent vector corresponding to $\frac{d\vec{r}}{dt}(a,t)$ is shown with the yellow arrow and the tangent vector corresponding to $\frac{d\vec{r}}{ds}(s,b)$ is shown with the red arrow.

The red and yellow tangent vectors at the highlighted point allow us to describe the tangent plane at that point. A piece of the tangent plane is shown as an orange parallelogram. Note that the orange parallelogram moves and tilts to approximate the curved surface near the highlighted point. Using the properties of the cross product allows us both to compute a vector that is orthogonal to the curved surface at the highlighted point *and* to find the area of the orange parallelogram. (Refer to Key Idea 9.5.5 and Key Idea 9.5.6 for the original introduction of these ideas.)

We now summarize the ideas above. Notice that because $\vec{r}(s,t)$ is a function of two variables, we write its derivatives using partial derivative notation with respect to s or t. These partial derivatives correspond, respectively, to holding the parameter t or s constant.

- Curves in the direction of change of s and t are given by vector-valued functions $\vec{r}(s,b)$ and $\vec{r}(a,t)$, respectively.
- The vectors

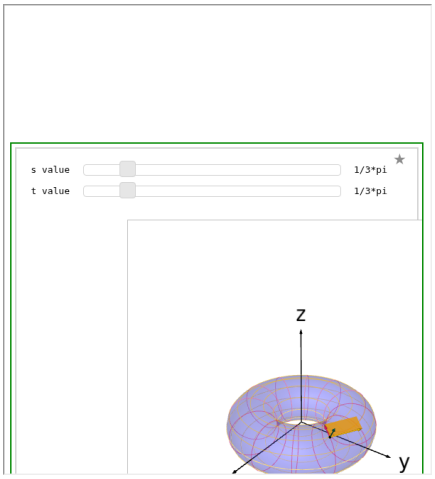
$$\frac{\partial \vec{r}}{\partial s}(a,b) = \vec{r}_s(a,b) \quad \text{and} \quad \frac{\partial \vec{r}}{\partial t}(a,b) = \vec{r}_t(a,b)$$

  are tangent to the parametric surface given by $\vec{r}(s,t)$ through the point $\vec{r}(a,b)$.
- The vector $(\vec{r}_s \times \vec{r}_t)(a,b)$ is orthogonal to surface at the point $\vec{r}(a,b)$. The area of the the parallelogram along the tangent plane to the surface with edges $\vec{r}_s(a,b)$ and $\vec{r}_t(a,b)$ is $\|(\vec{r}_s \times \vec{r}_t)(a,b)\|$.

Example 13.10.10 Recall the parameterization of the torus from Example 13.10.2 given by:

$$\begin{aligned}\vec{r}(s,t) =& \langle x(s,t), y(s,t), z(s,t)\rangle \\ =& \langle (b + a\cos(s))\cos(t), (b + a\cos(s))\sin(t), a\sin(s)\rangle\end{aligned}$$

for $s,t \in [0, 2\pi)$.



Standalone
In Context
Embed

Figure 13.10.11 A torus along with tangent vectors to the traces in the s- and t-directions at the point determined by the sliders. Also shown at that point are are a vector orthogonal to the torus and a parallelogram-shaped patch of the tangent plane.

- Fixing $t = t_0$ gives the parameterized curve $\vec{r}(s, t_0) = \langle (b + a\cos(s))\cos(t_0), (b + a\cos(s))\sin(t_0), a\sin(s)\rangle$. These are the traces with constant values of t and are shown in magenta in Figure 13.10.11.
- Simmilarly, the traces with constant values of $s = s_0$ are shown in yellow in Figure 13.10.11. They are parameterized by $\vec{r}(s_0, t) = \langle (b + a\cos(s_0))\cos(t), (b + a\cos(s_0))\sin(t), a\sin(s_0)\rangle$.
- The vector tangent to the surface along the trace with constant t-value is
$$\vec{r}_s = \langle (-a\sin(s))\cos(t), (-a\sin(s))\sin(t), 0\rangle.$$
This vector is shown by the magenta arrow in Figure 13.10.11.
- The vector tangent to the surface along the trace with constant s-value is
$$\vec{r}_t = \langle (a + b\cos(s))(-\sin(t)), (a + b\sin(s))(\cos(t)), 0\rangle.$$
This vector is shown by the yellow arrow in Figure 13.10.11.
- The green vector in Figure 13.10.11 is calculated by $\vec{r}_s \times \vec{r}_t$. This vector is orthogonal to the surface at the highlighted point. The length of the green vector corresponds to the area of the orange parallelogram, which lies in the tangent plane to the torus at the point. Notice that as the length of the yellow vector increases, the parallelogram's area and the length of the green vector both increase correspondingly.

▲

It may have seemed unusual for us to draw attention to the area of the parallelogram formed by the tangent vectors to the trace curves in the discussion above. However, the final subsection will demonstrate how this property can be used to understand a property of the larger surface.

13.10.4 The Surface Area of Parametrically Defined Surfaces

We are now ready to use the classic calculus approach to compute the surface area of a surface. For step one of the classic calculus approach, we break the parametrized surface into pieces corresponding to a grid of s- and t-steps as shown in Figure 13.10.12. On each of these pieces of the st-grid, we approximate the surface with a parallelogram tangent to the surface at that point. The sum of the areas of these parallelograms gives an approximation of the area of the surface. In step two of the classic calculus approach, we consider smaller step sizes in s and t as part of the sum used in the approximation. The third step of the classic calculus approach is a limit of the surface area approximation which leads to a double Riemann sum that corresponds to a double integral.

Standalone
In Context
Embed

Figure 13.10.12 Pieces of the tangent planes used to approximate surface area

Key Idea 13.10.13 Surface area. *Let $\vec{r}(s,t) = \langle x(s,t), y(s,t), z(s,t)\rangle$ be a parameterization of a smooth surface over a domain D. The* ***area of the surface*** *defined by $\vec{r}$ on D is given by*

$$S = \iint_D \|\vec{r}_s \times \vec{r}_t\| \; dA. \tag{13.10.2}$$

Proof. For this development, we will use a parametrized surface given by

$$\vec{r}(s,t) = x(s,t)\hat{\imath} + y(s,t)\hat{\jmath} + z(s,t)\hat{k}$$

with $a \le s \le b$ and $c \le t \le d$. Although we illustrate the idea using a rectangular domain in the st-plane, this same approach works for non-rectangular domains. The partial derivatives of this parameterization

$$\vec{r}_s(s,t) = x_s(s,t)\hat{\imath} + y_s(s,t)\hat{\jmath} + z_s(s,t)\hat{k}$$
$$\vec{r}_t(s,t) = x_t(s,t)\hat{\imath} + y_t(s,t)\hat{\jmath} + z_t(s,t)\hat{k}$$

give vectors tangent to the surface in the direction of change in s and t, respectively.

To break the section of the surface into steps in s and t, we partition the interval of s-values $[a,b]$ into m subintervals of length $\Delta s = \frac{b-a}{m}$ and let s_0, s_1, …, s_m be the endpoints of these subintervals, where $a = s_0 < s_1 < s_2 < \cdots <$

$s_m = b$. We also partition the interval of t-values $[c, d]$ into n subintervals of equal length $\Delta t = \frac{d-c}{n}$ and let t_0, t_1, ..., t_n be the endpoints of these subintervals, where $c = t_0 < t_1 < t_2 < \cdots < t_n = d$.

These subintervals partition the rectangle R with $a \le s \le b$ and $c \le t \le d$ in st-coordinates into mn sub-rectangles R_{ij} with opposite vertices (s_{i-1}, t_{j-1}) and (s_i, t_j) for i between 1 and m and j between 1 and n. These rectangles all have equal area $\Delta A = \Delta s \cdot \Delta t$.

Each of these rectangles from the st-domain splits the surface into an $m \times n$ grid of patches, but these patches of the surface do *not* all have the same (surface) area. Figure 13.10.9 shows a surface with traces given for five equally-spaced values of s and t. This breaks the surface into a 4×4 grid of patches. Since these patches of the surface do not have the same surface area, we estimate the surface area for each particular piece of the grid using the tools from the parameterization.

Each of the pieces of the surface corresponds to increasing s by a small amount Δs or increasing t by a small amount Δt from the point (s_{i-1}, t_{j-1}) in the st-parameter plane. We want to find the vectors corresponding to each of the sides of the orange parallelogram in Figure 13.10.9. In particular, the vector $\vec{r}_t \Delta t$ corresponds to the side of the parallelogram along a constant value of s (shown as a yellow vector in Figure 13.10.9) and $\vec{r}_s \Delta s$ corresponds to the side of the parallelogram along a constant value of t (shown as a red vector in Figure 13.10.9).

The area of this parallelogram is $S_{ij} = \|(\vec{r}_s \Delta s) \times (\vec{r}_t \Delta t)\|$ evaluated at the point (s_{i-1}, t_{j-1}). This is an approximation of the surface area for a patch of the surface. The magnitude simplifies to $S_{ij} = \|\vec{r}_s \times \vec{r}_t\| \, \Delta s \Delta t$, which means that the approximation over all of these pieces of the grid sums to

$$SA \approx \sum_{i=1}^{m} \sum_{j=1}^{n} S_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} \|\vec{r}_s(s_{i-1}, t_{j-1}) \times \vec{r}_t(s_{i-1}, t_{j-1})\| \, \Delta s \Delta t. \quad (13.10.3)$$

We have now completed both the first and second steps of the classic calculus approach by approximating the surface area and quantifying how the approximation changes with a smaller scale of steps (in s and t). You can use Figure 13.10.12 to see how increasing the number of pieces gives successively better approximations for the surface area of each piece.

Note that this approximation is a double Riemann sum corresponding to a double integral. In particular, taking the limit as Δs and Δt go to zero, the Riemann sum becomes a double integral of $\|\vec{r}_s \times \vec{r}_t\|$ over the rectangular region of integration with $a \le s \le b$ and $c \le t \le d$, which gives the following formula for calculating the surface area. ■

Our next example shows how to calculate the surface area of sphere using Key Idea 13.10.13 applied to the parameterizations of the sphere from Activity 13.10.2.

Example 13.10.14 Recall that in Activity 13.10.2, we found two parameterizations for the sphere of radius R centered at the origin. We will now explore how those parameterizations and Key Idea 13.10.13 can be used to find the surface area of this sphere. The first approach to parameterizing the sphere was to consider the sphere as a surface of revolution around the z-axis. This gave the parameterization

$$\vec{r}(s, t) = \langle \sqrt{R^2 - s^2} \cos(t), \sqrt{R^2 - s^2} \sin(t), s \rangle$$

with s acting as the z-coordinate, ranging from R to $-R$, and t acting as θ, ranging from 0 to 2π.

We calculate $\vec{r}_s, \vec{r}_t, \|\vec{r}_s \times \vec{r}_t\|$ as follows:

$$\begin{aligned}
\vec{r}_s &= \left\langle \frac{-2s}{\sqrt{R^2-s^2}}\cos(t), \frac{-2s}{\sqrt{R^2-s^2}}\sin(t), 1 \right\rangle \\
\vec{r}_t &= \left\langle -\sqrt{R^2-s^2}\sin(t), \sqrt{R^2-s^2}\cos(t), 0 \right\rangle \\
\vec{r}_s \times \vec{r}_t &= \left\langle -\sqrt{R^2-s^2}\cos(t), -\sqrt{R^2-s}\sin(t), -\frac{2s}{\sqrt{R^2-s^2}} \right\rangle \\
\|\vec{r}_s \times \vec{r}_t\| &= \sqrt{(R^2-s^2)\cos^2(t) + (R^2-s)\sin^2(t) + \frac{4s^2}{R^2-s^2}}
\end{aligned}$$

Unfortunately, this expression for $\|\vec{r}_s \times \vec{r}_t\|$ does not simplify much more. This gives us the surface area double integral

$$\int_{-R}^{R}\int_{0}^{2\pi} \sqrt{(R^2-s^2) + \frac{4s^2}{R^2-s^2}}\,dt\,ds.$$

While this integral *can* be evaluated, it requires a trigonometric substitution. We will omit the calculations that result from this and turn to our other parameterization for an easier calculation of the surface area.

Recall the parameterization of the sphere that came from using spherical coordinates:

$$\vec{r}(s,t) = \langle R\sin(s)\cos(t), R\sin(s)\sin(t), R\cos(s)\rangle$$

with s acting as the ϕ-coordinate, ranging from 0 to 2π, and t acting as the θ, ranging from 0 to 2π.

We calculate $\vec{r}_s, \vec{r}_t, \|\vec{r}_s \times \vec{r}_t\|$ as follows:

$$\begin{aligned}
\vec{r}_s &= \langle R\cos(s)\cos(t), R\cos(s)\sin(t), -R\sin(s)\rangle \\
\vec{r}_t &= \langle -R\sin(s)\sin(t), R\sin(s)\cos(t), 0\rangle \\
\vec{r}_s \times \vec{r}_t &= \langle R^2\sin^2(s)\cos(t), R^2\sin^2(s)\sin(t), R^2\cos^2(t)\cos(s)\sin(s) + R^2\sin^2(t)\cos(s)\sin(s)\rangle \\
&= R^2\sin(s)\langle \sin(s)\cos(t), \sin(s)\sin(t), \cos(s)\rangle \\
\|\vec{r}_s \times \vec{r}_t\| &= R^2\sin(s)
\end{aligned}$$

You may recognize this expression from our work on the volume element in spherical coordinates, Key Idea 12.8.5. In both this case and the volume element of spherical coordinates, we are measuring how warped our space is based on transformations using spherical coordinates.

The surface area double integral is

$$\int_0^{\pi}\int_0^{2\pi} R^2\sin(s)dt\,ds,$$

which evaluates to

$$R^2(2\pi)\left[-\cos(s)\Big|_{s=0}^{s=\pi}\right] = R^2(2\pi)(2) = 4\pi R^2,$$

the familiar formula for the surface area of a sphere for which you likely previously had no explanation. ▲

Activity 13.10.3 This activity computes the surface area of a right circular cylinder. In particular, consider the cylinder with radius a and height h defined

parametrically by

$$\vec{r}(s,t) = a\cos(s)\hat{\imath} + a\sin(s)\hat{\jmath} + t\hat{k}$$

for $0 \le s \le 2\pi$ and $0 \le t \le h$, as shown in Figure 13.10.15.

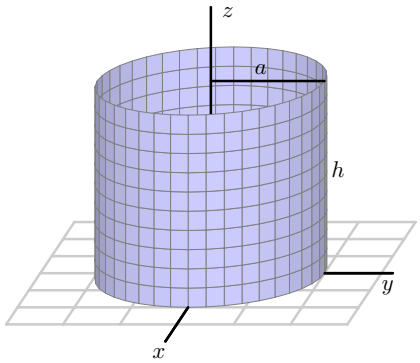


Figure 13.10.15 A part of the cylinder of radius a about the z-axis

(a) Calculate $\vec{r}_s, \vec{r}_t, \|\vec{r}_s \times \vec{r}_t\|$ based on the parameterization given above.

(b) Use the calculations from the previous part to set up an iterated integral to determine the surface area of this cylinder.

(c) Evaluate your iterated integral from the previous part.

(d) One way to think about the surface area of a cylinder is to cut the cylinder horizontally and find the perimeter of the resulting cross sectional circle, then multiply by the height. Calculate the surface area of the given cylinder using this alternate approach, and compare your result the value from the previous part.

As we noted earlier, we can take any surface $z = f(x,y)$ and obtain a parameterization for the surface by writing $\langle s, t, f(s,t)\rangle$. Hence, we can use Key Idea 13.10.13 to find the area of a surface $z = f(x,y)$ over a given domain.

Activity 13.10.4 Let $z = f(x,y)$ define a smooth surface, and consider the parameterization $\vec{r}(s,t) = \langle s, t, f(s,t)\rangle$.

(a) Let D be a region in the domain of f. Using equation (13.10.2), show that the area of the surface defined by the graph of f over D is

$$\iint_D \sqrt{(f_x(x,y))^2 + (f_y(x,y))^2 + 1}\, dA.$$

(b) Use the formula developed in the previous part to calculate the area of the surface defined by $f(x,y) = \sqrt{4-x^2}$ over the rectangle $D = [-2,2] \times [0,3]$.

(c) Observe that the surface of the solid describe the previous part is half of a circular cylinder. Use the standard formula for the surface area of a cylinder to calculate the surface area in a different way, and compare your result from above.

Summary

- A parameterization of a curve describes the coordinates of a point on the curve in terms of a single parameter t, while a parameterization of a

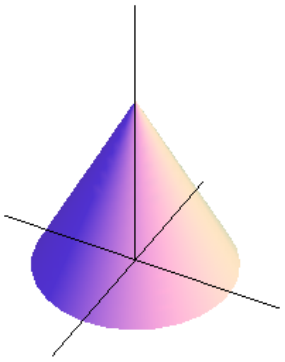
surface describes the coordinates of points on the surface in terms of two independent parameters.

- If $\vec{r}(s,t) = \langle x(s,t), y(s,t), z(s,t)\rangle$ describes a smooth surface in 3-space on a domain D, then the area, S, of that surface is given by

$$S = \iint_D \|\vec{r}_s \times \vec{r}_t\| \; dA.$$

13.10.5 Exercises

1. Consider the cone shown below.



If the height of the cone is 6 and the base radius is 5, write a parameterization of the cone in terms of $r = s$ and $\theta = t$.

$x(s,t) =$ _____,
$y(s,t) =$ _____, and
$z(s,t) =$ _____, with
_____ $\le s \le$ _____ and
_____ $\le t \le$ _____.

2. Parameterize a vase formed by rotating the curve $z = 6\sqrt{x-2}$, $2 \le x \le 4$, around the z-axis. Use s and t for your parameters.

$x(s,t) =$ _____,
$y(s,t) =$ _____, and
$z(s,t) =$ _____, with
_____ $\le s \le$ _____ and
_____ $\le t \le$ _____

3. Find parametric equations for the sphere centered at the origin and with radius 7. Use the parameters s and t in your answer.

$x(s,t) =$ _____,
$y(s,t) =$ _____, and
$z(s,t) =$ _____, where
_____ $\le s \le$ _____ and
_____ $\le t \le$ _____.

4. Match the parametric equations with the verbal descriptions of the surfaces by putting the letter of the verbal description to the left of the letter of the parametric equation.

(a) $\mathbf{r}(u,v) = u\cos v\mathbf{i} + u\sin v\mathbf{j} + u^2\mathbf{k}$

(b) $\mathbf{r}(u,v) = u\mathbf{i} + \cos v\mathbf{j} + \sin v\mathbf{k}$

(c) $\mathbf{r}(u,v) = u\mathbf{i} + u\cos v\mathbf{j} + u\sin v\mathbf{k}$

(d) $\mathbf{r}(u,v) = u\mathbf{i} + v\mathbf{j} + (2u - 3v)\mathbf{k}$

A. circular paraboloid

B. circular cylinder

C. cone

D. plane

5. Parameterize the plane that contains the three points
(4,-3,4), (-2,-6,-6), (5,35,10).

Use s and t for the parameters in your parameterization and enter your formula as a single vector, with angle brackets, for example, $\langle 1+s+t, s-t, 3-t\rangle$. Each component should be a function of the form a+bs+ct where a,b,c are appropriate real numbers.

$\vec{r}(s,t) =$ _____

6. Find the surface area of that part of the plane $10x + 9y + z = 4$ that lies inside the elliptic cylinder $\frac{x^2}{16} + \frac{y^2}{9} = 1$

Surface Area = _____

7. Find the surface area of the part of the circular paraboloid $z = x^2 + y^2$ that lies inside the cylinder $x^2 + y^2 = 1$.

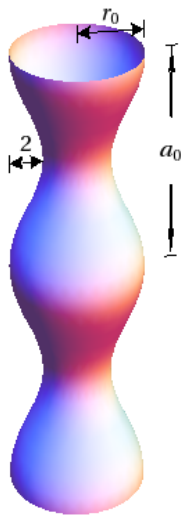
8. Find the surface area of the part of the plane $1x + 1y + z = 1$ that lies inside the cylinder $x^2 + y^2 = 9$.

9. Write down the iterated integral which expresses the surface area of $z = y^3\cos^7 x$ over the triangle with vertices (-1,1), (1,1), (0,2):

$$\int_a^b \int_{f(y)}^{g(y)} \sqrt{h(x,y)}\, dxdy$$

$a =$ _____
$b =$ _____
$f(y) =$ _____
$g(y) =$ _____
$h(x,y) =$ _____

10. A decorative oak post is 48 inches long and is turned on a lathe so that its profile is sinusoidal as shown in the figure below.



In this figure, $r_0 = 4$ inches and $a_0 = 8$ inches.

(a) Describe the surface of the post parametrically using cylindrical coordinates and the parameters s and t.

$x(s,t) =$ _____,
$y(s,t) =$ _____, and
$z(s,t) =$ _____, where
_____ $\leq s \leq$ _____ and
_____ $\leq t \leq$ _____.

(b) Find the volume of the post.

volume = _____
(Include .)

11. Find the surface area of the part of the sphere $x^2 + y^2 + z^2 = 25$ that lies above the cone $z = \sqrt{x^2 + y^2}$

12. If a parametric surface given by $\mathbf{r_1}(u,v) = f(u,v)\mathbf{i} + g(u,v)\mathbf{j} + h(u,v)\mathbf{k}$ and $-4 \leq u \leq 4, -4 \leq v \leq 4$, has surface area equal to 4, what is the surface area of the parametric surface given by $\mathbf{r_2}(u,v) = 3\mathbf{r_1}(u,v)$ with $-4 \leq u \leq 4, -4 \leq v \leq 4$?

13. Write down the iterated integral which expresses the surface area of $z = y^6 \cos^4 x$ over the triangle with vertices (-1,1), (1,1), (0,2):

$$\int_a^b \int_{f(y)}^{g(y)} \sqrt{h(x,y)}\, dxdy$$

$a =$ _____
$b =$ _____
$f(y) =$ _____
$g(y) =$ _____
$h(x,y) =$ _____

14. Find the area of the portion of the sphere of radius 9 (centered at the origin) that is in the cone $z > \sqrt{x^2 + y^2}$.

15. Find the surface area of the part of the plane $5x + 4y + z = 1$ that lies inside the cylinder $x^2 + y^2 = 16$.

16. Find the surface area of the part of the circular paraboloid $z = x^2 + y^2$ that lies inside the cylinder $x^2 + y^2 = 1$.

17. Find the area of the surface obtained by rotating the curve

$$y = 2x^3$$

from $x = 0$ to $x = 1$ about the x-axis.

The area is _____ square units.

18. The vector equation $\mathbf{r}(u, v) = u\cos v\mathbf{i} + u\sin v\mathbf{j} + v\mathbf{k}$, $0 \le v \le 3\pi$, $0 \le u \le 1$, describes a helicoid (spiral ramp). What is the surface area?

19. Find the surface area of the portion S of the cone $z^2 = x^2 + y^2$, where $z \ge 0$,

contained within the cylinder $y^2 + z^2 \le 25$.

Area(S) = _____

20. A torus of radius 6 (and cross-sectional radius 1) can be represented parametrically by the function $\mathbf{r} : D \to \mathbb{R}^3$:

$$\mathbf{r}(\theta, \phi) = ((6 + \cos\phi)\cos\theta, (6 + \cos\phi)\sin\theta, \sin\phi)$$

where D is the rectangle given by $0 \le \theta \le 2\pi,\ 0 \le \phi \le 2\pi$.

The surface area of the torus is _____

21. Find an ***equation*** of the tangent plane (in the variables x, y and z) to the parametric surface

$\mathbf{r}(u, v) = \langle u, u^2 + 5v, -5v^2\rangle$ at the point $(-2, 4, 0)$.

22. Calculate $\mathbf{T}_u$, $\mathbf{T}_v$, and $\mathbf{n}(u, v)$ for the parametrized surface at the given point.

Then find the equation of the tangent plane to the surface at that point.

$\Phi(u, v) = (2u + v, u - 4v, 8u)$; $\quad u = 4, \quad v = 9$

$\mathbf{T}_u$ = _____, $\mathbf{T}_v$ = _____, $\mathbf{n}(u, v)$ = _____

The tangent plane:

_____ $= 9z$

23. Consider the ellipsoid given by the equation

$$\frac{x^2}{16} + \frac{y^2}{25} + \frac{z^2}{9} = 1.$$

In Activity 13.10.2, we found that a parameterization of the sphere S of radius R centered at the origin is

$$x(r, s) = R\cos(s)\cos(t),\ y(s, t) = R\cos(s)\sin(t),\ \text{ and }\ z(s, t) = R\sin(s)$$

for $-\frac{\pi}{2} \le s \le \frac{\pi}{2}$ and $0 \le t \le 2\pi$.

a. Let (x, y, z) be a point on the ellipsoid and let $X = \frac{x}{4}$, $Y = \frac{y}{5}$, and $Z = \frac{z}{3}$. Show that (X, Y, Z) lies on the sphere S. Hence, find a parameterization of S in terms of X, Y, and Z as functions of s and t.

b. Use the result of part (a) to find a parameterization of the ellipse in terms of x, y, and z as functions of s and t. Check your parametrization by substituting x, y, and z into the equation of the ellipsoid. Then check your work by plotting the surface defined by your parameterization.

24. In this exercise, we explore how to use a parametrization and iterated integral to determine the surface area of a sphere.

a. Set up an iterated integral whose value is the portion of the surface area of a sphere of radius R that lies in the first octant (see the parameterization you developed in Activity 13.10.2).

b. Then, evaluate the integral to calculate the surface area of this portion of the sphere.

c. By what constant must you multiply the value determined in (b) in order to find the total surface area of the entire sphere.

d. Finally, compare your result to the standard formula for the surface area of sphere.

25. Consider the plane generated by $z = f(x, y) = 24 - 2x - 3y$ over the region $D = [0, 2] \times [0, 3]$.

a. Sketch a picture of the overall solid generated by the plane over the given domain.

b. Determine a parameterization $\vec{r}(s, t)$ for the plane over the domain D.

c. Use Equation (13.10.2) to determine the surface area generated by f over the domain D.

d. Observe that the vector $\vec{u} = \langle 2, 0, -4 \rangle$ points from $(0, 0, 24)$ to $(2, 0, 20)$ along one side of the surface generated by the plane f over D. Find the vector $\vec{v}$ such that $\vec{u}$ and $\vec{v}$ together span the parallelogram that represents the surface defined by f over D, and hence compute $|\vec{u} \times \vec{v}|$. What do you observe about the value you find?

26. A cone with base radius a and height h can be realized as the surface defined by $z = \frac{h}{a}\sqrt{x^2 + y^2}$, where a and h are positive.

a. Find a parameterization of the cone described by $z = \frac{h}{a}\sqrt{x^2 + y^2}$. (Hint: Compare to the parameterization of a cylinder as seen in Activity 13.10.3.)

b. Set up an iterated integral to determine the surface area of this cone.

c. Evaluate the iterated integral to find a formula for the lateral surface area of a cone of height h and base a.

27. Parameterize the following surfaces and be sure to state bounds for your parameterization.

(a) The first octant portion of $2x + y + 2z = 4$

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

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

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

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

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

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

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

(i) $z = \left(\frac{x}{2}\right)^2 - \left(\frac{y}{3}\right)^2$

13.11 Flux Integrals

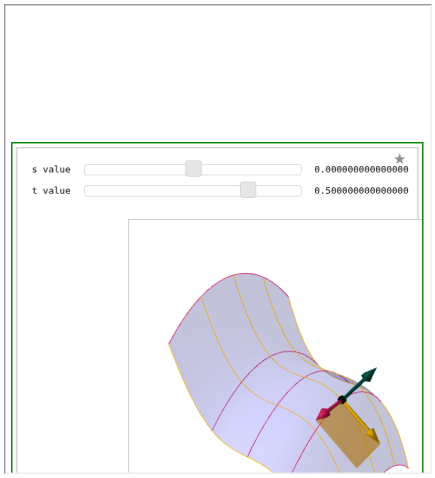
Motivating Questions

- How can we measure how much of a vector field flows through a surface in space?
- How can we calculate the amount of a vector field that flows through common surfaces, such as the graph of a function $z = f(x, y)$?

13.11.1 Introduction

Section 13.10 showed how we can use vector-valued functions of two variables to give a parameterization of a surface in space. For instance, the function $\vec{r}(s,t) = \langle 2\cos(t)\sin(s), 2\sin(t)\sin(s), 2\cos(s)\rangle$ with domain $0 \le t \le 2\pi$ and $0 \le s \le \pi$ parameterizes a sphere of radius 2 centered at the origin. Section 13.10 also gives examples of how to write parameterizations based on other geometric relationships like when one coordinate can be written as a function of the other two. In Subsection 13.10.4, we set up a Riemann sum based on a parameterization that would measure the surface area of our curved surfaces in space.

In Figure 13.11.1, we plot a surface using a parametrization $\vec{r}(s,t) = \langle f(s,t), g(s,t), h(s,t)\rangle$. The magenta curves represent curves where s varies and t is held constant, while the yellow curves represent curves where t varies and s is held constant. The vector in magenta is $\vec{r}_s = \frac{\partial \vec{r}}{\partial s} = \langle f_s, g_s, h_s\rangle$ which measures the direction and magnitude of change in the coordinates of the surface when only s is varied. Similarly, the vector in yellow is $\vec{r}_t = \frac{\partial \vec{r}}{\partial t} = \langle f_t, g_t, h_t\rangle$ which measures the direction and magnitude of change in the coordinates of the surface when only t is varied. We also plot the parallelogram that is formed by $\vec{r}_s$ and $\vec{r}_t$, which is tangent to the surface. The area of this parallelogram offers an approximation for the surface area of a patch of the surface.



Standalone
In Context
Embed

Figure 13.11.1 A three-dimensional plot showing how a parallelogram can be used to approximate a section of a curved surface

From Section 9.5, we also know that $\vec{r}_s \times \vec{r}_t$ (plotted in green) will be orthogonal to both $\vec{r}_s$ and $\vec{r}_t$ and its magnitude will be given by the area of the parallelogram.

Preview Activity 13.11.1 In this preview activity, we will explore the parameterizations of a few familiar surfaces and confirm some of the geometric properties described in the introduction above.

(a) Use the ideas from Section 13.10 to give a parameterization $\vec{r}(s,t)$ of the following surface. Be sure to specify the bounds on each of your parameters.

S_1: the right circular cylinder centered on the x-axis of radius 2 when $0 \leq x \leq 5$

(b) Draw a graph of S_1 from the previous part. Label the points that correspond to (s,t) points of $(0,0)$, $(0,1)$, $(1,0)$, and $(2,3)$.

(c) For the parameterization from part a, calculate $\vec{r}_s$, $\vec{r}_t$, and $\vec{r}_s \times \vec{r}_t$.

(d) For the parameterization from part a, find the value for $\vec{r}_s$,$\vec{r}_t$, and $\vec{r}_s \times \vec{r}_t$ at the (s,t) points of $(0,0)$, $(0,1)$, $(1,0)$, and $(2,3)$.

(e) Draw your vector results from d on your graph and confirm the geometric properties described in the introduction to this section. Namely, $\vec{r}_s$ and $\vec{r}_t$ should be tangent to the surface, while $\vec{r}_s \times \vec{r}_t$ should be orthogonal to the surface (in addition to $\vec{r}_s$ and $\vec{r}_t$).

(f) Repeat steps (a) through (e) for the following two surfaces:

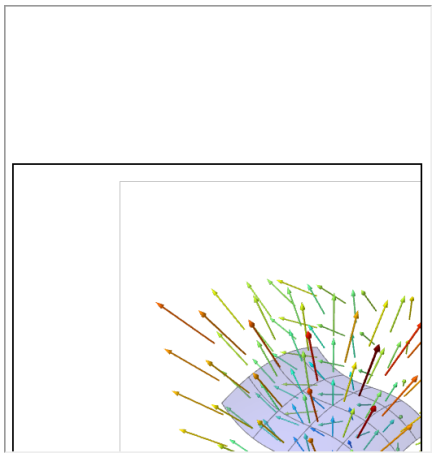
(i) S_2: the sphere centered at the origin of radius 3

(ii) S_3: the first octant portion of the plane $x + 2y + 3z = 6$

As we saw in Section 13.10, we can set up a Riemann sum of the areas for the parallelograms in Figure 13.11.1 to approximate the surface area of the region plotted by our parametrization. Equation (13.10.2) shows that we can compute the exact surface area by taking a limit of a Riemann sum which will correspond to integrating the magnitude of $\vec{r}_s \times \vec{r}_t$ over the appropriate parameter bounds. What if we wanted to measure a quantity other than the surface area? Our focus in this section we will be the exploration of a specific case of this question: How can we measure the amount of a three dimensional vector field that flows through a particular section of a surface? The geometric tools we have reviewed in this section, especially the vector $\vec{r}_s \times \vec{r}_t$, will be valuable.

13.11.2 The Idea of the Flux of a Vector Field through a Surface

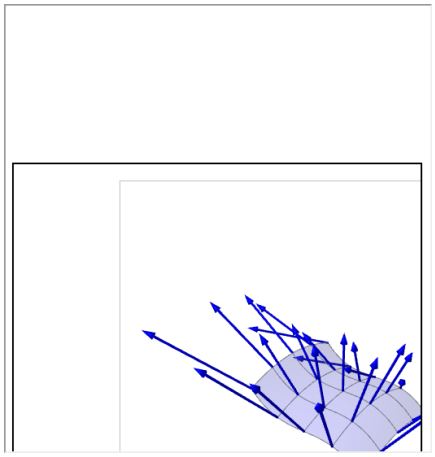
In this section we will look at how to measure the amount of a vector field that flows through a surface in space. Figure 13.11.2 illustrates a plot that demonstrates this idea. Our definition of divergence in Section 13.7 looked at measuring the amount of vector field flowing out of a small region on a 2D plane. In this subsection, we will set up a precise measurement of this same measurement but over a region of a curved surface in 3D.



Standalone
In Context
Embed

Figure 13.11.2 A three-dimensional vector field and a surface

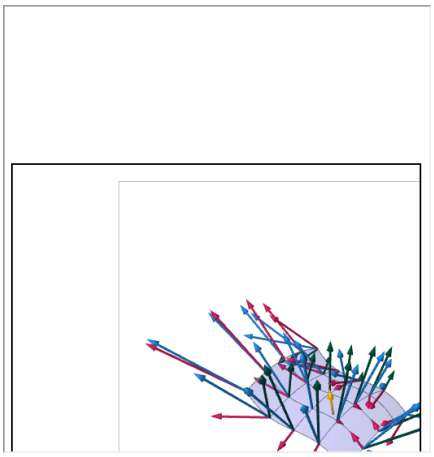
As with understanding line integrals of vector-valued functions in Section 13.3, we don't care about the output of the vector field at points away from the surface. We would really would like to examine the output vectors for the points on our surface. To do this, we will look at Figure 13.11.3, which plots the output of our vector field at an array of points on our surface.



Standalone
In Context
Embed

Figure 13.11.3 A three-dimensional vector field evaluated along a surface

The central question we would like to consider is "How can we measure the amount of a three dimensional vector field that flows through a particular section of a curved surface?", so we only need to consider the amount of the vector field that flows *through* the surface. Any portion of our vector field that flows along (or tangent) to the surface will not contribute to the amount that goes through the surface. In Figure 13.11.4, we have split the vector field for points on our surface into two components. One component, plotted in green, is orthogonal to the surface. The component that is tangent to the surface is plotted in magenta.



Standalone
In Context
Embed

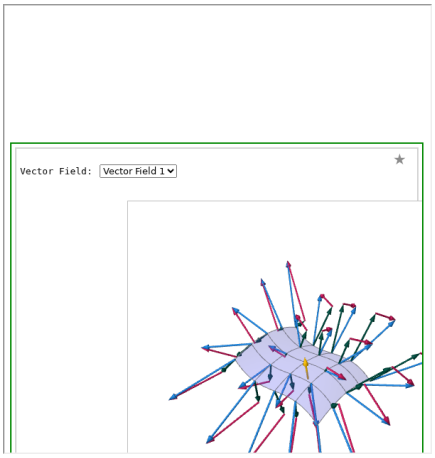
Figure 13.11.4 The decomposition of three-dimensional vector field evaluated along a surface into normal and tangent components

In order to measure the amount of the vector field that moves through the plotted section of the surface, we must find the accumulation of the lengths of the green vectors in Figure 13.11.4. Notice that some of the green vectors are moving through the surface in a direction opposite of others. In other words, we will need to pay attention to the direction in which these vectors move through our surface and not just the magnitude of the green vectors.

If we have a parameterization of the surface, then the vector $\vec{r}_s \times \vec{r}_t$ varies smoothly across our surface and gives a consistent way to describe which direction we choose as "through" the surface. If we define a positive flow through our surface as being consistent with the yellow vector in Figure 13.11.4, then there is more positive flow (in terms of both magnitude and area) than negative flow through the surface. Thus, the *net* flow of the vector field through this surface is positive.

Activity 13.11.2 Visualizing flux through a surface. In this activity, you will compare the net flow of different vector fields through our sample surface. In Figure 13.11.5 you can select between five different vector fields. Once you select a vector field, the vector field for a set of points on the surface will be plotted in blue. Each blue vector will also be split into its normal component (in green) and its tangential component (in magenta). The yellow vector defines the direction for positive flow through the surface.

Look at each vector field and order the vector fields from greatest net flow through the surface to least net flow through the surface. Remember that a negative net flow through the surface should be lower in your rankings than any positive net flow.



Standalone
In Context
Embed

Figure 13.11.5 Five vector fields flowing through a surface, with decomposition into normal and tangential components

13.11.3 Computing the Flux of a Vector Field through a Surface

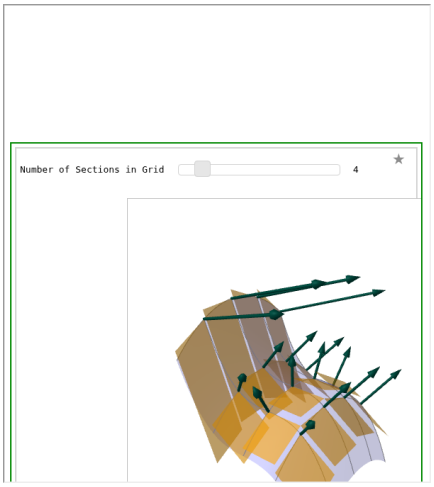
Now that we have developed a conceptual understanding of what we are trying to measure, we can set up a corresponding Riemann sum to measure the flux of a vector field through a section of a surface. In the following theorem, we utilize a parameterization to calculate both the normal vector to the surface at a location as well an measure of how the surface area is chaninging at that point. You can look at the proof that follows this theorem to see precisely how the elements discussed give rise to this method of calculation.

Theorem 13.11.6 *Let a smooth surface Q be parametrized by $\vec{r}(s,t)$ over a domain D. The total flux of a smooth vector field $\vec{F}$ through Q is given by*

$$\iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

Proof. Let Q be the section of our surface and suppose that Q is parameterized by $\vec{r}(s,t)$ with $a \leq s \leq b$ and $c \leq t \leq d$. The domain of $\vec{r}$ is a region of the st-plane, which we call D, and the range of $\vec{r}$ is Q.

As with most problems in integral calculus, we slice our region of interest into smaller pieces. Specifically, we slice $a \leq s \leq b$ into n equally-sized subintervals with endpoints $s_0, s_1, \ldots, s_n$ and $c \leq t \leq d$ into m equally-sized subintervals with endpoints $t_0, t_1, \ldots, t_n$. This divides D into nm rectangles of size $\Delta s = \frac{b-a}{n}$ by $\Delta t = \frac{d-c}{m}$. We index these rectangles as $D_{i,j}$. Every $D_{i,j}$ has area (in the st-plane) $\Delta s \Delta t$. The partition of D into the rectangles $D_{i,j}$ also partitions Q into nm corresponding pieces which we call $Q_{i,j} = \vec{r}(D_{i,j})$. From Section 13.10 (specifically (13.10.3)) the surface area of $Q_{i,j}$ is approximated by $S_{i,j} = \|(\vec{r}_s \times \vec{r}_t)(s_i, t_j)\| \Delta s \Delta t$.



Standalone
In Context
Embed

Figure 13.11.7 Pieces of the tangent planes that approximate a surface

We want to measure the total flow of the vector field $\vec{F}$ through Q, which we will approximate on each $Q_{i,j}$ and then sum to obtain the total flow. In other words, the flux of $\vec{F}$ through Q is

$$\text{Flux} = \sum_{i=1}^{n}\sum_{j=1}^{m} \left\|\vec{F}_{\perp Q_{i,j}}\right\| \cdot S_{i,j},$$

where $\left\|\vec{F}_{\perp Q_{i,j}}\right\|$ is the length of the component of $\vec{F}$ orthogonal to $Q_{i,j}$.

For each $Q_{i,j}$, we approximate the surface Q by the tangent plane to Q at a corner of that partition element. This corresponds to using the planar elements in Figure 13.11.7, which have surface area $S_{i,j}$. The vector $\vec{w}_{i,j} = (\vec{r}_s \times \vec{r}_t)(s_i, t_j)$ can be used to measure the orthogonal direction (and thus define which direction we mean by positive flow through Q) on the i, j partition element. This means that

$$\vec{F}_{\perp Q_{i,j}} = \left\|\text{proj}_{\vec{w}_{i,j}} \vec{F}(s_i, t_j)\right\| = \frac{\vec{F}(s_i, t_j) \cdot \vec{w}_{i,j}}{\|\vec{w}_{i,j}\|}$$

Combining these pieces, we find that the flux through $Q_{i,j}$ is approximated by

$$\begin{aligned}\text{Flux through } Q_{i,j} = \left\|\vec{F}_{\perp Q_{i,j}}\right\| \cdot S_{i,j} &= \left(\frac{\vec{F}_{i,j} \cdot \vec{w}_{i,j}}{\|\vec{w}_{i,j}\|}\right)(\|\vec{w}_{i,j}\|\,\Delta s \Delta t) \\ &= \left(\vec{F}_{i,j} \cdot (\vec{r}_s \times \vec{r}_t)\right)(\Delta s \Delta t)\,,\end{aligned}$$

where $\vec{F}_{i,j} = \vec{F}(s_i, t_j)$. Therefore we may approximate the total flux by

$$\text{Total Flux} = \sum_{i=1}^{n}\sum_{j=1}^{m} \left(\vec{F}_{i,j} \cdot \vec{w}_{i,j}\right)(\Delta s \Delta t)\,.$$

In Figure 13.11.7, you can change the number of sections in the partition and see the geometric result of refining the partition. Taking the limit as $n, m \to \infty$ gives the following double integral.

$$\iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA$$

Conceptually, the above limit shows how the argument in Subsection 13.10.4 generalizes when measuring the flux of a vector field through a surface rather than just the surface area. The key difference between the Riemann sum above and Riemann sum used to set up Surface area is that we are using the dot product of the vector field with the normal vector to the surface (given by $\vec{w}_{i,j} = (\vec{r}_s \times \vec{r}_t)(s_i, t_j)$) as the scalar value in our sum. This realization gives the following theorem, which states that the flux through a curved surface in space can be computed using a parameterization of the surface and a double integral over a flat region in the plane of parameter values. ■

In the next example, we will look at how Theorem 13.11.6 is used on a part of a cone and make sense of the vector $\vec{w}(t, s) = \vec{r}_s \times \vec{r}_t$ and the scalar $\vec{F} \cdot (\vec{r}_s \times \vec{r}_t)$.

Remark 13.11.8 In some Webwork exercises and other sources, you will see flux integrals specified with the following notations:

$$\iint_{S_1} \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA = \iint_{S_1} \vec{F} \cdot d\vec{A} = \iint_{S_1} \vec{F} \cdot d\vec{S}$$

While we will use the notation specified in Theorem 13.11.6 primarily, the other notations come from combining the normal vector (calculated from the parameterization of the surface) and the area element into a single vector-valued area element. In other words, $(\vec{r}_s \times \vec{r}_t)\, dA = d\vec{A} = d\vec{S}$.

Example 13.11.9 In this example we will compute the flux of the vector field $\vec{F} = \langle xz - 2, y + x, 1\rangle$ through the surface of the cone given by $x^2 + y^2 = z^2$ with $1 \leq z \leq 3$. So that we can use Theorem 13.11.6, we first parameterize our surface and calculate $\vec{w}(t, s) = \vec{r}_s \times \vec{r}_t$ based on that parameterization.

In cylindrical coordinates, our cone is described by $r = z$ which suggests that we can use r and θ as s and t, respectively. Consider the parameterization $\vec{r}(s, t) = \langle s\cos(t), s\sin(t), s\rangle$ with bounds $1 \leq s \leq 3$ and $0 \leq t < 2\pi$. Notice that s is acting as both r and z. Calculating the partial derivatives of $\vec{r}$ gives the following:

$$\vec{r}_s = \langle \cos(t), \sin(t), 1\rangle$$
$$\vec{r}_t = \langle -s\sin(t), s\cos(t), 0\rangle$$

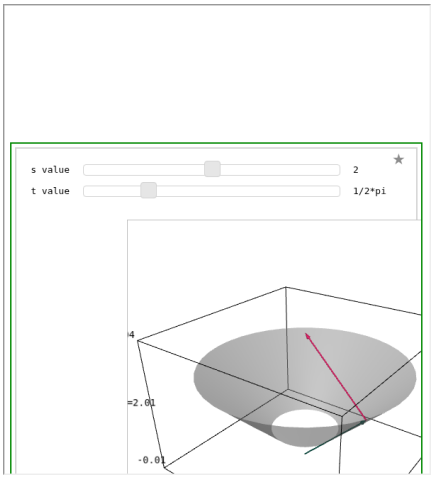
Taking the cross product of $\vec{r}_s$ and $\vec{r}_t$, we obtain the vector-valued function $\vec{w}$:

$$\vec{w} = \vec{r}_s \times \vec{r}_t = \langle -s\cos(t), -s\sin(t), s(\cos(t)^2 + \sin(t)^2)\rangle.$$

We may simplify this to

$$\vec{w} = \langle -s\cos(t), -s\sin(t), s\rangle.$$

Notice that $\vec{w}$ is very close to $\vec{r}$. If we think of $\vec{r}$ as giving $\langle x, y, z\rangle$, then $\vec{w} = \langle -x, -y, z\rangle$. The figure below shows the relationship between vectors $\vec{r}$ and $\vec{w}$ for a point on the surface.



Standalone
In Context
Embed

Figure 13.11.10 A plot of our surface with the parameterization vector in green and the normal vector $\vec{w}$ in magenta

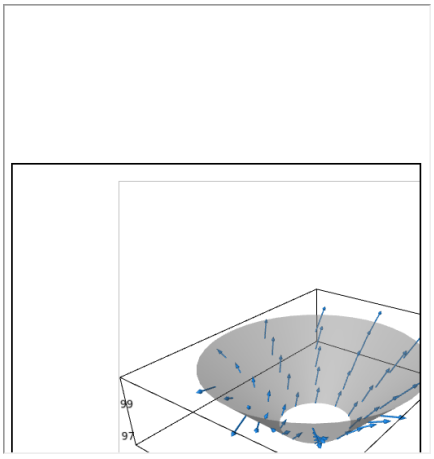
In Figure 13.11.10, the normal vector (as calculated from the parameterization) points inside the cone. This is the direction of positive flow when measuring the flux of the vector field. Our parameterization will also transform our vector field into a function of s and t. Specifically,

$$\vec{F}(s,t) = \langle (s\cos(t))(s) - 2, s\sin(t) + s\cos(t), 1\rangle$$

Applying Theorem 13.11.6 to our parameterization and vector field allows us to compute $I = \iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA$ as an interated integral

$$\begin{aligned} I &= \int_0^{2\pi}\int_1^3 \langle (s\cos(t))(s) - 2, s\sin(t) + s\cos(t), 1\rangle \cdot \langle -s\cos(t), -s\sin(t), s\rangle \, ds\, dt \\ &= \int_0^{2\pi}\int_1^3 ((s\cos(t))(s) - 2)(-s\cos(t)) + (-s\sin(t))(s\sin(t) + s\cos(t)) + s \\ &= -\frac{62}{3}\pi. \end{aligned}$$

Note that the result of this flux integral is negative, which means more of the vector field is going out of the cone than is coming into the cone.



Standalone
In Context
Embed

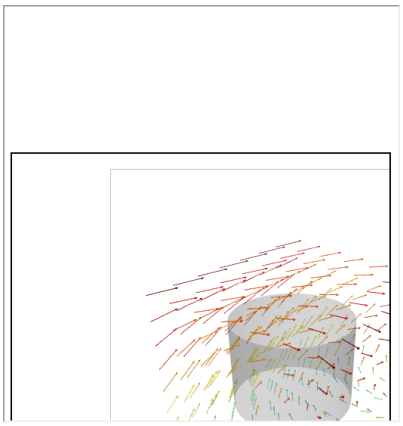
Figure 13.11.11 A plot of the surface with the vector field $\vec{F}$ for an array of points on the surface. The vector field has been scaled down by a factor of 5 in this plot.

In Figure 13.11.11 you can see much more of the vector field is flowing outside of the cone than inside, which matches with our result of calculating the flux as negative. ▲

The next activity asks you to carefully go through the process of calculating the flux of some vector fields through a cylindrical surface.

Activity 13.11.3 Checking the Visualization for Flux.

(a) Figure 13.11.12 shows a plot of the vector field $\vec{F} = \langle y, z, 2 + \sin(x) \rangle$ and a right circular cylinder of radius 2 and height 3 (with open top and bottom). Consider the vector field going into the cylinder (toward the z-axis) as corresponding to positive flux.



Standalone
In Context
Embed

Figure 13.11.12 The vector field $\vec{F} = \langle y, z, 2 + \sin(x) \rangle$ and a right circular cylinder

(i) Reasoning graphically, do you think the flux of $\vec{F}$ throught the cylinder will be positive, negative, or zero? Write a few sentences jutifying your answer.

(ii) Parameterize the right circular cylinder of radius 2, centered on the z-axis, for $0 \leq z \leq 3$. Be sure to give bounds on your parameters.

(iii) Based on your parameterization, compute $\vec{r}_s$, $\vec{r}_t$, and $\vec{r}_s \times \vec{r}_t$. Confirm that these vectors are either orthogonal or tangent to the right circular cylinder. Is your orthogonal vector pointing in the direction of positive flux or negative flux?

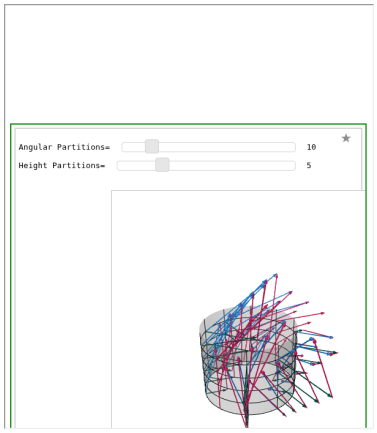
(iv) Use your parametrization to write $\vec{F}$ as a function of s and t.

Hint. The x-coordinate is given by the first component of $\vec{r}$.

(v) Compute the flux of $\vec{F}$ through the parameterized portion of the right circular cylinder.

(vi) Does your computed value for the flux match your prediction from earlier?

(vii) Use Figure 13.11.13 to make an argument about why the flux of $\vec{F} = \langle y, z, 2 + \sin(x) \rangle$ through the right circular cylinder is zero.



Standalone
In Context
Embed

Figure 13.11.13 The vector field $\vec{F} = \langle y, z, 2 + \sin(x) \rangle$ with normal and tangential components plotted on a right circular cylinder

(b) Write a couple of sentences to explain how the results of the flux calculations would be different if we used the vector field $\vec{F} = \left\langle y, z, \cos(xy) + \frac{9}{z^2+6.2} \right\rangle$ and the same right circular cylinder.

(c) Write a couple of sentences to explain how the results of the flux calculations would be different if we used the vector field $\vec{F} = \langle y, -x, 3 \rangle$ and the same right circular cylinder.

In the exercises for this section, we will look at some computational ideas to help us more efficiently compute the value of a flux integral. In many cases, the surface we are looking at the flux through can be written with one coordinate as a function of the others. For simplicity, we consider $z = f(x, y)$. Additionally, there will be exercises that guide you through common surfaces like spheres and cylinder surfaces.

Remark 13.11.14 Note that throughout this section, we have implicitly assumed that we can parametrize the surface S in such a way that $\vec{r}_s \times \vec{r}_t$ gives a well-defined normal vector. Technically, this means that the surface be **orientable**. Most "reasonable" surfaces are orientable. However, there are surfaces that are not orientable. Perhaps the most famous is formed by taking a long, narrow piece of paper, giving one end a half twist, and then gluing the ends together. Try doing this yourself, but before you twist and glue (or tape), poke a tiny hole through the paper on the line halfway between the long edges of your strip of paper and circle your hole. After gluing, place a pencil with its eraser end on your dot and the tip pointing away. Think of this as a potential normal vector. Keep the eraser on the paper, and follow the middle of your surface around until the first time the eraser is again on the dot. Is your pencil still pointing the same direction relative to the surface that it was before?

Summary

- A flux integral of a vector field, $\vec{F}$, on a surface in space, S, measures how much of $\vec{F}$ goes through S_1.
- Let the smooth surface, S, be parametrized by $\vec{r}(s, t)$ over a domain D. The total flux of a smooth vector field $\vec{F}$ through S is given by

$$\iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

- If S_1 is of the form $z = f(x, y)$ over a domain D, then the total flux of a smooth vector field $\vec{F}$ through S_1 is given by

$$\iint_D \vec{F}(x, y, f(x, y)) \cdot \left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle \, dA.$$

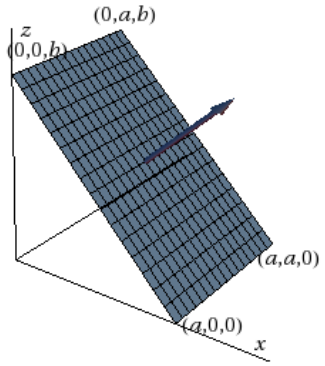
13.11.4 Exercises

1. Calculate the flux of the vector field $\vec{F}(x, y, z) = 8y\vec{j}$ through a square of side length 9 in the plane $y = 4$. The square is centered on the y-axis, has sides parallel to the axes, and is oriented in the positive y-direction.

Flux = _____

Hint. $y = 4$

2. Compute the flux of the vector field $\vec{v} = 5\,\vec{\imath} + \vec{\jmath} + 3\,\vec{k}$ through the rectangular region shown below, assuming it is oriented as shown and that $a = 1$ and $b = 2$.



flux = _____

3. Find the flux of the constant vector field $\vec{v} = -5\,\vec{\imath} - 3\,\vec{\jmath} - 3\,\vec{k}$ through a square plate of area 25 in the xy-plane oriented in the positive z-direction.

flux = _____

4. Calculate the flux of the vector field $\vec{F} = (y+7)\vec{j}$ through a square of side 4 in the xz-plane, oriented in the negative y-direction.

flux = _____

5. Calculate the flux of the vector field $\vec{F} = -3\,\vec{\imath} + 5\,\vec{\jmath} + 5\,\vec{k}$ through a cube of side 2 with sides parallel to the axes.

flux = _____

Suppose that one face of the cube is at $y = 0$ and the opposite face is at $y = 2$. If the face at $y = 0$ is removed, what is the flux through the resulting surface?

flux = _____

6. (a) Set up a double integral for calculating the flux of the vector field $\vec{F}(x, y, z) = z^2\vec{k}$ through the upper hemisphere of the sphere $x^2+y^2+z^2 = 9$, oriented away from the origin. If necessary, enter ρ as *rho,* θ as *theta,* and ϕ as *phi.*

Flux = $\int_A^B \int_C^D$ _____ $d\phi\, d\theta$

A = _____

B = _____
C = _____
D = _____
(b) Evaluate the integral.
Flux = $\iint_S \vec{F} \cdot d\vec{A}$ = _____

Hint. $z = \rho\cos(\phi)$

7. Compute the flux of the vector field $\vec{F} = 5x^2y^2z\vec{k}$ through the surface S which is the cone $\sqrt{x^2+y^2} = z$, with $0 \le z \le R$, oriented downward.

(a) Parameterize the cone using cylindrical coordinates (write θ as *theta*).
$x(r,\theta)$ = _____
$y(r,\theta)$ = _____
$z(r,\theta)$ = _____
with _____ $\le r \le$ _____
and _____ $\le \theta \le$ _____

(b) With this parameterization, what is $d\vec{A}$?
$d\vec{A}$ = _____

(c) Find the flux of $\vec{F}$ through S.
flux = _____

8. Compute the flux of the vector field $\vec{F} = z\vec{i} + 6x\vec{k}$ through the parameterized surface S oriented upward and given, for $0 \le s \le 2, \quad 2 \le t \le 5$, by

$$x = s^2, \quad y = 2s + t^2, \quad z = 4t.$$

flux = _____

9. Are the following statements true or false?

(a) If S is the unit sphere centered at the origin, oriented outward and the flux integral $\iint_S \vec{F} \cdot d\vec{A}$ is zero, then $\vec{F} = \vec{0}$.

(b) If S is the unit sphere centered at the origin, oriented outward and the flux integral $\iint_S \vec{F} \cdot d\vec{A}$ is zero, then $\vec{F}(x,y,z)$ is perpendicular to $\vec{r} = \langle x, y, z\rangle$ at every point of S.

(c) If S is the unit sphere centered at the origin, oriented outward and $\vec{F} = x\vec{i} + y\vec{j} + z\vec{k} = \vec{r}$, then the flux integral $\iint_S \vec{F} \cdot d\vec{A}$ is positive.

(d) If S_1 is a rectangle with area 1 and S_2 is a rectangle with area 2, then $2\iint_{S_1} \vec{F} \cdot d\vec{A} = \iint_{S_2} \vec{F} \cdot d\vec{A}$

(e) If S is an open-ended circular cylinder centered about the z-axis, oriented away from the z-axis, and $\vec{F} = \langle 3, -2, 6\rangle$, then the flux of $\vec{F}$ through S is zero.

(f) The area vector of a flat, oriented surface is parallel to the surface.

(g) If S is the cube bounded by the six planes $x = \pm 3$, $y = \pm 3$, $z = \pm 3$, oriented outward, and $\vec{F} = 2\vec{i} - \vec{k}$, then the flux of $\vec{F}$ through S is zero.

(h) If S is an open-ended circular cylinder centered about the z-axis, oriented away from the z-axis, and $\vec{F} = \langle x, y, 0\rangle$, then the flux of $\vec{F}$ through S is positive.

10. F $= x\boldsymbol{i} + y\boldsymbol{j} + 5z\boldsymbol{k}$ and σ is the portion of the cone $z = \sqrt{x^2 + y^2}$ between the planes $z = 5$ and $z = 12$ oriented with normal vectors pointing upwards. Find the flux of the flow field F across σ: _____

11. Let $\mathbf{F} = 5x\mathbf{i} + 3y\mathbf{j}$ and let $\boldsymbol{n}$ be the outward unit normal vector to the positively oriented circle $x^2+y^2 = 9$. Compute the flux integral $\int_C \mathbf{F} \cdot \mathbf{n}\, ds$.

12. Calculate the flux of the vector field $\vec{F} = -8\vec{i} + x^2\vec{j} - 5\vec{k}$, through the square of side 4 in the plane $y = 8$, centered on the y-axis, with sides parallel to the x and z axes, and oriented in the positive y-direction.

flux = _____

13. Calculate the flux of the vector field $\vec{F} = (x^2 + z^2)\vec{j}$ through the disk of radius 2 in the xz-plane and centered at the origin and oriented in the direction of the positive y axis.

flux = _____

14. Calculate the flux integral $\int_S (3\vec{i}+3\vec{j}+3\vec{k}) \cdot d\vec{A}$ where S is a disk of radius 5 in the plane $x + y + z = 3$, oriented upward.

$\int_S (3\vec{i} + 3\vec{j} + 3\vec{k}) \cdot d\vec{A}$ = _____

15. Let S be the cube with side length 4, faces parallel to the coordinate planes, and centered at the origin.

(a) Calculate the total flux of the constant vector field $\vec{v} = -5\,\boldsymbol{\imath} - 3\,\boldsymbol{\jmath} + 4\,\boldsymbol{\kappa}$ out of S by computing the flux through each face separately.

flux through the face at $x = 2$: _____
flux through the face at $y = 2$: _____
flux through the face at $z = 2$: _____
flux through the face at $x = -2$: _____
flux through the face at $y = -2$: _____
flux through the face at $z = -2$: _____

Thus the total flux is

flux = _____

Calculate the flux out of S for any constant vector field $\vec{v} = a\vec{i}+b\vec{j}+c\vec{k}$.

flux = _____

(Be sure that you can explain why your answers to parts (a) and (b) make sense.)

16. Calculate the flux integral. $\int_S (3\vec{i} + 4z\vec{k}) \cdot d\vec{A}$ where S is $z = x^2 + y^2$ with $0 \le x \le 2, 0 \le y \le 1$, oriented upward.

$\int_S (3\vec{i} + 4z\vec{k}) \cdot d\vec{A}$ = _____

17. Compute the flux of the vector field $\vec{F} = x\vec{i} + y\vec{j} + z\vec{k}$ through the surface S, which is a closed cylinder of radius 3, centered on the x-axis, with $-2 \le x \le 2$, and oriented outward.

flux = _____

18. Compute the flux of the vector field $\vec{F} = 4x\vec{i} + 4y\vec{j}$ through the surface S, which is the part of the surface $z = 4 - (x^2 + y^2)$ above the disk of radius 2 centered at the origin, oriented upward.

flux = _____

19. Compute the flux of the vector field $\vec{F} = 5y\vec{i} + 4\vec{j} - 5xz\vec{k}$ through the surface S, which is the surface $y = x^2 + z^2$, with $x^2 + z^2 \le 9$, oriented in

the positive y-direction.

flux = _____

20. Compute the flux of the vector field $\vec{F} = z\vec{k}$ through the parameterized surface S, which is oriented toward the z-axis and given, for $0 \le s \le 2$, $0 \le t \le 1$, by

$$x = 5s + 2t, \quad y = 5s - 2t, \quad z = s^2 + t^2.$$

flux = _____

21. Compute the flux of $\vec{F} = x\vec{i} + y\vec{j} + z\vec{k}$ through the curved surface of the cylinder $x^2 + y^2 = 1$ bounded below by the plane $x + y + z = 3$, above by the plane $x + y + z = 8$, and oriented away from the z-axis.

flux = _____

22. Compute the surface integral over the given oriented surface:

$\mathbf{F} = y^3\mathbf{i} + 7\mathbf{j} - x\mathbf{k}$, portion of the plane $x + y + z = 1$ in the octant $x, y, z \ge 0$, downward-pointing normal

$\iint_S \mathbf{F} \cdot d\mathbf{S}$ = _____

23. Calculate the flux of the vector field $\vec{F}(x, y, z) = (4x + 4)\vec{i}$ through a disk of radius 4 centered at the origin in the yz-plane, oriented in the negative x-direction.

Flux = _____

Hint.

24. Flux Through Surfaces of the Form $z = f(x, y)$. In this exercise, we will look at how to use a parametrization of a surface that can be described as $z = f(x, y)$ to efficiently calculate flux integrals.

(a) Suppose that S is a surface given by $z = f(x, y)$. Find a parametrization $\vec{r}(s, t)$ of S.

Hint. Use $s = x$ and $t = y$.

(b) Show that the vector orthogonal to the surface S has the form

$$\vec{r}_s \times \vec{r}_t = \left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle.$$

(c) For each of the three surfaces given below, compute $\vec{r}_s \times \vec{r}_t$, graph the surface, and compute $\vec{r}_s \times \vec{r}_t$ for four different points of your choosing. You should make sure your vectors $\vec{r}_s \times \vec{r}_t$ are orthogonal to your surface.

(i) $z = x^2 + y^2$

(ii) $x + 2y + z = -4$

(iii) $z = x^2 - y^2$

(d) For each of the three surfaces in part c, use your calculations and Theorem 13.11.6 to compute the flux of each of the following vector fields through the part of the surface corresponding to the region D in the xy-plane.

(i) $\vec{F} = \langle x, y, z \rangle$ with D given by $0 \le x, y \le 2$

(ii) $\vec{F} = \langle -y, x, 1 \rangle$ with D as the triangular region of the xy-plane with vertices $(0, 0)$, $(1, 0)$, and $(1, 1)$

(iii) $\vec{F} = \langle z, y - x, (y - x)^2 - z^2 \rangle$ with D given by $0 \le x, y \le 2$

25. Calculating the Flux through a Sphere. For this activity, let S_R be the sphere of radius R centered at the origin.

(a) Parametrize S_R using spherical coordinates. Give your parametrization as $\vec{r}(s,t)$, and be sure to state the bounds of your parametrization.

Hint. Use $s = \theta$ and $t = \phi$.

(b) Use your parametrization of S_R to compute $\vec{r}_s \times \vec{r}_t$.

(c) Your result for $\vec{r}_s \times \vec{r}_t$ should be a scalar expression times $\vec{r}(s,t)$. Explain why the outward pointing orthogonal vector on the sphere is a multiple of $\vec{r}(s,t)$ and what that scalar expression means.

(d) Use your parametrization of S_2 and the results of part b to calculate the flux through S_2 for each of the three following vector fields.

(i) $\vec{F}_1 = \langle x, y, z \rangle$

(ii) $\vec{F}_2 = \langle -y, x, -1 \rangle$

(iii) $\vec{F}_3 = \langle x - y, y + x, z - 1 \rangle$

(e) Use computer software to plot each of the vector fields from part d and interpret the results of your flux integral calculations.

(f) If we used the sphere of radius 4 instead of S_2, explain how each of the flux integrals from part d would change. You do not need to calculate these new flux integrals, but rather explain if the result would be different and how the result would be different.

13.12 Surface Integrals of Scalar Valued Functions

Motivating Questions

- How can we measure the accumulation of a scalar-valued function along a surface in space?
- What does that accumulation measure?
- How can we efficiently calculate scalar surface integrals?

13.12.1 Introduction

In Section 13.11 we looked at flux integrals measure how much of a vector field flows through a given surface. In other words we calculated the accumulation of of the component of the vector field that is normal to the surface over all the points on the surface. In this section, we will generalize this ideas to calculate the accumluation of other scalar functions over a region given by a curved surface. The Preview Activity below will use our mining example that motivated several of our ideas related to line integrals to understand a few key ideas.

Preview Activity 13.12.1 In Preview Activity 13.6.1 we looked at how to understand line integrals of scalar functions through the analogy of running a mining machine along a given path. In short, the amount of copium mined

depended on the density of copium at points on the path and the length of the path driven by the mining rig.

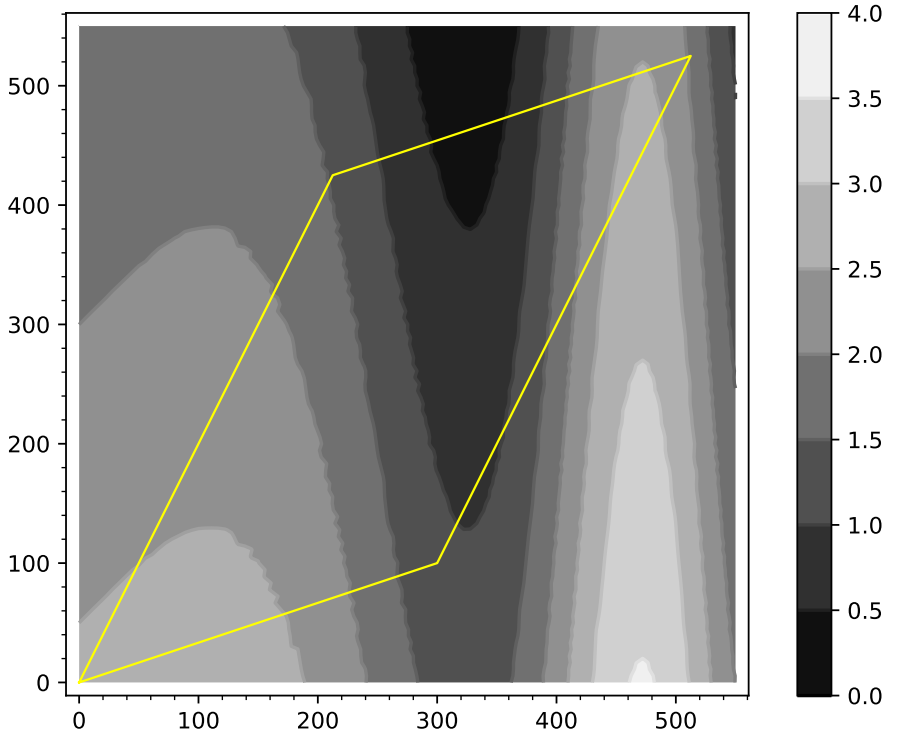


Figure 13.12.1 A plot of land with density of copium deposits and the edges of the plot drawn in yellow

The opening tasks of Preview Activity 13.6.1 had you estimating the amount of copium mined by driving along the edge of a plot of land (drawn in yellow on Figure 13.12.1). This interpretation meant that the scalar function we were using measured the linear density of copium. Thus, the Riemann sum we computed was the product of linear density of copium with the distance traveled.

(a) In this task, we will interpret Figure 13.12.1 as a contour graph of the density of copium per unit area. This will allow us to compute the total amount of copium in our mining area by setting up a double integral. To approximate the total amount of copium available in our mine, do the following:

(i) Break the mining plot (the area inside the yellow segments of Figure 13.12.1) into three pieces. Estimate the area of the three pieces you are using. Write a few sentences explaining your methods of estimating the areas.

(ii) For each of the three pieces you used in part a.i, estimate the average density of copium on the piece. Write a few sentences explaining your methods of estimating the average density on each piece.

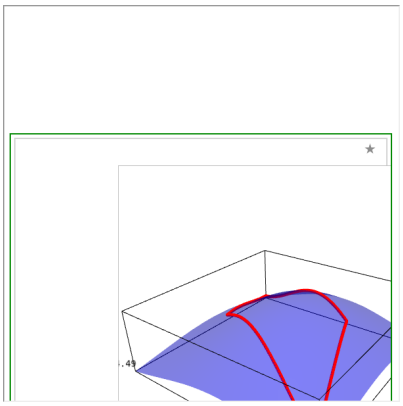
(iii) Give an estimate for the total amount of copium on the mining plot and explain your computation.

(b) What if instead of your mining plot being on a flat piece of land as represented in Figure 13.12.1, your mining plot was on a hill as represented

in Figure 13.12.2. If we had the same copium density plot as a function of the (x, y) coordinates, which of the following would you expect to be true?

- the total amount of copium available on Figure 13.12.2 is greater than on Figure 13.12.1
- the total amount of copium available on Figure 13.12.2 is lesser than on Figure 13.12.1
- the total amount of copium available on Figure 13.12.2 is the same as on Figure 13.12.1

Explain your reasoning for your choice.



Standalone
In Context
Embed

Figure 13.12.2 A 3D plot of the region including our mining area (shown with boundary in red)

In Section 13.11, the idea of a flux integral was introduced by looking at how much of a vector field flows through a given section of a surface in space. In particular, the flux integral measured the accumulated amount of the vector field that is orthogonal to the surface, which changes at different points on the surface. Theorem 13.11.6 gives an efficient way to calculate a flux integral in terms of a parameterization of the surface. This was an extension of the method for computing the surface area given in Section 13.10.

Just as we generalized line integrals of vector fields into scalar line integrals in Section 13.6, we will spend this section examining questions related to the accumulation of scalar-valued functions along surfaces in space as a means of generalizing flux integrals.

13.12.2 Defining surface integrals of scalar functions

As you might expect, we can set a classic calculus approach to calculate the accumulation of our scalar valued function over the points on a curved surface. This process is given in more detail in the proof of Theorem 13.12.3 but proceeds like in our work for calculating surface areas and flux integrals. Namely, we will split our curved surface into smaller piecese, then we will utilize a parameterization to measure the area of the piece and convert the xyz-coordinates of our point and the associated scalar function to a double integral in our parameters, usually s and t.

Theorem 13.12.3 *Let a smooth surface S_1 be parametrized by $\vec{r}(s,t)$ over a region D, and let f be a continuous function on a neighborhood around S_1. The accumulation of $f(x,y,z)$ over S_1 is denoted by $\iint_{S_1} f\,dS$ and is called the* ***scalar surface integral of f over S_1****. The scalar surface integral of f over*

S_1 is computed as

$$\iint_{S_1} f\,dS = \iint_D f(s,t)\,\|(\vec{r}_s \times \vec{r}_t)(s,t)\|\,dA.$$

Proof. In order to create a Riemann sum that will measure the accumulation of $f(x,y,z)$, a scalar valued function, over S_1, a smooth, bounded surface in space, we will break our surface into smaller pieces where we can approximate the value of f and the surface area of each piece. Recall that in Subsection 13.10.4 and Section 13.11, we developed a way to measure surface area given a parameterization for the surface.

Suppose that S_1 is parameterized by $\vec{r}(s,t)$ with $a \le s \le b$ and $c \le t \le d$. In our classic calculus style, we slice our region of interest into smaller pieces. Specifically, we slice $a \le s \le b$ into n equally-sized subintervals with endpoints $s_0, s_1, \ldots, s_n$ and we slice $c \le t \le d$ into m equally-sized subintervals with endpoints $t_0, t_1, \ldots, t_m$. This divides D into nm rectangles of size $\Delta s = \frac{b-a}{n}$ by $\Delta t = \frac{d-c}{m}$. We index these rectangles as $D_{i,j}$. Every $D_{i,j}$ has area (in the st-plane) of $\Delta s \Delta t$. The partition of D into the rectangles $D_{i,j}$ also partitions S_1 into nm corresponding pieces which we call $S_{i,j} = \vec{r}(D_{i,j})$. From Section 13.10 (specifically equation (13.10.3)) the surface area of $S_{i,j}$ is approximated by $SA_{i,j} = \|(\vec{r}_s \times \vec{r}_t)(s_i,t_j)\|\,\Delta s \Delta t$.

We illustrate this approximation of S_1 by rectangular patches of tangent planes, each with area $SA_{i,j} = \|(\vec{r}_s \times \vec{r}_t)(s_i,t_j)\|\,\Delta s \Delta t$ in Figure 13.12.4, where you can adjust the slider to control the number of rectangles. (For simplicity, the figure takes $m = n$.)

Standalone
In Context
Embed

Figure 13.12.4 Pieces of the tangent planes used to approximate surface area

To measure the accumulation A of the output of $f(x,y,z)$ over S_1, we use the following Riemann sum:

$$\begin{aligned} A &= \sum_{i=1}^{n}\sum_{i=1}^{m} f(x^*_{i,j}, y^*_{i,j}, z^*_{i,j}) SA_{i,j} \\ &= \sum_{i=1}^{n}\sum_{i=1}^{m} f(x(s^*_i,t^*_j), y(s^*_i,t^*_j), z(s^*_i,t^*_j)) SA_{i,j} \\ &= \sum_{i=1}^{n}\sum_{i=1}^{m} f(s^*_i,t^*_j)\,\|(\vec{r}_s \times \vec{r}_t)(s_i,t_j)\|\,\Delta s \Delta t. \end{aligned}$$

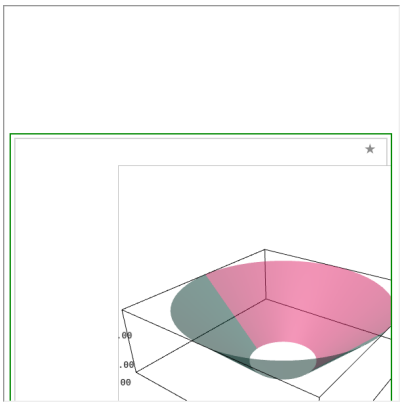
as $\Delta s \to 0$ and $\Delta t \to 0$, a double Riemann sum of this form represents the double integral (over the variables s and t) of the function $f(x(s,t), y(s,t), z(s,t)) \, \|(\vec{r}_s \times \vec{r}_t)(s,t)\|$. This leads to the following theorem. ■

In the following example, we both reason about the value of a scalar surface integral without working through computations and apply Theorem 13.12.3 to compute values of scalar surface integrals.

Example 13.12.5 For this example, we will consider the surface S_1 given by $x^2 + y^2 = z^2$ such that $1 \leq z \leq 3$. This is the same part of the cone used in Example 13.11.9. We also let $f(x, y, z) = x$.

(a) For our first part of our example, we use symmetry to reason about $\iint_{S_1} f \, dS$. Without using any calculations or parameterizations, we can reason that this scalar surface integral will be zero. In a scalar surface integral, we are calculating the accumulation of the scalar output over a surface area in space. In Figure 13.12.6, you can see that we can symmetrically break our surface S_1 into two regions that have exactly the same surface area properties ***and*** symmetric output values for our scalar function, f. Here we denote by S_2 the portion of S_1 with $x \geq 0$ (colored green), while S_3 is the magenta portion of the surface, which has $x < 0$. The only difference between the green and magenta portions of Figure 13.12.6 that is relevant to our surface integral is that the values of f will be positive on the green surface S_2 and negative on the magenta surface S_3 No matter how we break up the Riemann sum or the corresponding scalar surface integral for the green side of our surface, applying the same pattern to the magenta side will give the same result but with an opposite sign. In other words,

$$\int_{S_2} f \, dS = -\int_{S_3} f \, dS.$$



Standalone
In Context
Embed

Figure 13.12.6 The surface S_1 divided into regions S_2 with positive x-coordinate in green and S_3 with negative x-coordinate in red

Therefore,

$$\int_{S_1} f \, dS = \int_{S_2} f \, dS + \int_{S_3} f \, dS = \int_{S_2} f \, dS - \int_{S_2} f \, dS = 0.$$

(b) From a calculation standpoint, we can verify this result using our parameterization from Example 13.11.9 in Theorem 13.12.3. Specifically, we will use the parameterization of S_1 given by $\vec{r}(s,t) = \langle s\cos(t), s\sin(t), s \rangle$ with

bounds $1 \leq s \leq 3$ and $0 \leq t < 2\pi$. We also know from Example 13.11.9 that

$$\vec{w} = \vec{r}_s \times \vec{r}_t = \langle -s\cos(t), -s\sin(t), s \rangle.$$

Hence, we have $\|\vec{w}\| = \sqrt{s^2(\cos(t)^2 + \sin(t)^2) + s^2} = \sqrt{2}s$. Thus, Theorem 13.12.3 gives

$$\int_{S_1} f\, dS = \int_1^3 \int_0^{2\pi} (s\cos(t))\sqrt{2}s\, dt\, ds = 0.$$

Notice that we can evaluate this integral quickly because the inside integral is $\int_0^{2\pi} \cos(t)\, dt = 0$.

(c) If we want to calculate the integral of $f(x, y, z) = x$ over S_2, then we have very little extra work to do. Only the bounds on the integral will change since we are using the same parameterization and function from the previous part. We can use the parameterization given by $\vec{r}(s, t) = \langle s\cos(t), s\sin(t), s \rangle$ with $1 \leq s \leq 3$ and $-\frac{\pi}{2} \leq t < \frac{\pi}{2}$. As before, $\|\vec{w}\| = \sqrt{s^2(\cos(t)^2 + \sin(t)^2) + s^2} = 2s$. Thus, Theorem 13.12.3 gives

$$\begin{aligned}
\int_{S_2} f\, dS &= \int_1^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (s\cos(t))\sqrt{2}s\, dt\, ds \\
&= \int_1^3 \sqrt{2}s^2[\sin(t)]_{\Big|_{t=-\frac{\pi}{2}}^{\frac{\pi}{2}}}\, ds \\
&= \int_1^3 2\sqrt{2}s^2\, ds \\
&= \int_1^3 2\sqrt{s}s^2\, ds = 52\sqrt{2}/3.
\end{aligned}$$

It makes sense that the result of $\int_{S_2} f\, dS$ is positive because the output of our function is non-negative for all points on S_2. Remember that the Riemann sum used in calculating surface integrals of scalar functions is the product of the function's value on each piece times the estimate of the surface area. Thus, if the function's value is positive for all points on the surface, then the scalar surface integral will also be positive.

▲

Before moving on to an activity that gives you a chance to practice reasoning about scalar surface integrals, we consider an additional example that will highlight why our algebraic calcultions are vital in cases where there is not a simple geometric argument.

Example 13.12.7 For this example, we will consider the same surface S_1 as in Example 13.12.5, but we will change the scalar function we are integrating to something more complicated. Our goal is to calculate

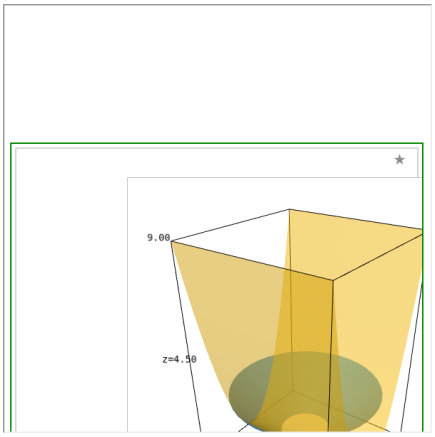
$$\int_{S_1} z - x^2\, dS.$$

We can reuse the parameterization and calculations from above for S_1 here and move directly to setting up the iterated integral. We will take a moment here to note that converting our scalar function $g(x, y, z) = z - x^2$ to a function of s and t will give $g(s, t) = s - s^2\cos(t)^2$. Therefore, we have the following

iterated integral, which we evaluate:

$$\begin{aligned}\int_{S_1} z - x^2 \, dS &= \int_1^3 \int_0^{2\pi} (s - s^2\cos(t)^2) \cdot 2s \, dt \, ds \\ &= \int_1^3 2s^2 \left[t - s\left(\frac{2t + \sin(2t)}{4}\right)\right]_{t=0}^{2\pi} ds \\ &= \int_1^3 2s^2(-\pi(s-2)) \, ds = -\frac{16\pi}{3}\end{aligned}$$

This result is a bit more difficult to to make sense of as being negative. In Figure 13.12.8, we have a plot of the surface S_1 in blue and the parabolic cylinder surface given by $z - x^2 = 0$ is plotted in yellow. The points on the yellow surface correspond to where our scalar function ($g(x, y, z) = z - x^2$) gives an output of zero. The points inside the parabolic cylinder have a z-coordinate greater than x^2, which means that g will have a positive output for these points. The points outside of the parabolic cylinder have a z-coordinate less than x^2, which means that g will have a negative output for these points. In order to make an argument about whether $\int_{S_1} z - x^2 \, dS$ is positive, negative, or zero, we will need to assess whether there is more surface area where g takes on positive values, more with negative values, or an equal amount of positive-negative values. Actually, our problem is even harder than this! Because the Riemann sum is the product of the output for each piece of our surface times the surface area estimate, we will need to think about which points/pieces have larger or smaller outputs for g. This is *very* difficult in general, because we need a systematic way of estimating both the output of g and the surface area of different pieces of the surface.



Standalone
In Context
Embed

Figure 13.12.8 The surface S_1 is plotted with the surface given by $z - x^2 = 0$

In our case, there is more surface area of S_1 inside the parabolic cylinder than outside. For points that are far away from the yellow surface, the magnitude of the value of g for that point will be large. For points that are close to the yellow surface, the z-coordinates are close to x^2, and thus values of g are close to zero. This means that in order to properly estimate the scalar surface integral, we must assess whether there is a greater weighted surface area inside the parabolic cylinder than outside, where the weight of each surface area piece is given by how far the piece is from the yellow parabolic cylinder. This is not an easy argument to make, and is not easy to verify visually for this example. For some examples, you may be able to make a nice geometric argument about the size of scalar surface integrals, but the algebraic calculation given by Theorem 13.12.3 gives an exact value. ▲

In the next activity, we consider some situations where we can reason about the sign of a surface integral of a scalar function.

Activity 13.12.2 In this activity, we will try to understand the scalar surface integral by looking at whether the value of the scalar surface integral will be positive, negative, or zero over common surfaces. In each part below, you are given a function and a surface. For each surface, first draw a plot of the surface and make sure you have labeled a proper scale for each coordinate direction. Then reason if the given surface integral is positive, negative, or zero. Be sure to justify your answers in terms of the function being integrated *and* the particulars of the surface of integration.

(a) For S_1 defined as the top half ($z \geq 0$) of the sphere of radius one centered at the origin, consider the surface integral $\iint_{S_1} x \, dS$.

(b) For S_2 defined as the bottom half ($z \leq 0$) of the sphere of radius one centered at the origin, consider the surface integral $\iint_{S_2} z \, dS$.

(c) For S_3 defined as the disc of radius one centered at $(1, 0, 0)$ on the plane $x = 1$, consider the surface integral $\iint_{S_3} x + z \, dS$.

(d) For S_1 as defined above, consider the surface integral $\iint_{S_1} x + z \, dS$.

13.12.3 Properties of Scalar Surface Integrals

Before stating some useful properties of scalar line integrals, we will recall some convenient notation. If S_1 and S_2 are disjoint surfaces[1], we denote by $S_1 + S_2$ the surface containing every point that is in S_1 or S_2. Also, if S_1 is a surface, then $-S_1$ denotes the same surface but parameterized in such a way that the normal vector points in the opposite direction. The list below summarizes some other properties of scalar surface integrals, each of which has a familiar analogue amongst the properties of other integrals we have studied.

Properties of Scalar Line Integrals.

For a constant scalar k, scalar valued functions f and g, and oriented surfaces S_1 and S_2, the following properties hold:

a. $\displaystyle\iint_{S_1} (kf) \, dS = k \iint_{S_1} f \, dS$

b. $\displaystyle\iint_{S_1} (f + g) \, dS = \iint_{S_1} f dS + \iint_{S_1} g \, dS$

c. $\displaystyle\iint_{-S_1} f \, dS = \iint_{S_1} f \, dS$

d. $\displaystyle\iint_{S_1+S_2} f \, dS = \iint_{S_1} f \, dS + \iint_{S_2} f \, dS$

Activity 13.12.3 Explaining Properties of Scalar Surface Integrals. In this activity, we will be explaining each of the properties from Properties of Scalar Line Integrals in the context of a new analogy. We have just purchased a plot of land that spans two mountains, Sugar Mountain and Spice Mountain.

[1] Technically, the surfaces may intersect, but there are restrictions on the manner in which they can intersect, and we will not go into the details here.

We will label the plot of land on Sugar Mountain S_1 and the plot of land on Spice Mountain S_2. Unfortunately there are two types of toxic organisms on the surface of your new land, which may explain why you paid so little for the land. Let f be the density of the toxic fungus on your new plot of land and let g be the density of toxic bacteria on the new plot.

(a) Explain in your own words what $\iint_{S_1} f\,dS$ means in the above analogy and what exactly would be measured by this scalar surface integral.

(b) Explain in your own words what $\iint_{S_1} (2f)\,dS = 2\iint_{S_1} f\,dS$ means in the new analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

(c) Explain in your own words what $\iint_{S_2} (f+g)\,dS = \iint_{S_2} f\,dS + \iint_{S_2} g\,dS$ means in the new analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

(d) Explain in your own words what $\iint_{-S_2} f\,dS = \iint_{S_2} f\,dS$ means in the new analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

(e) Explain in your own words what $\iint_{S_1+S_2} f\,dS = \iint_{S_1} f\,dS + \iint_{S_2} f\,dS$ means in the new analogy. It may be helpful to describe each side of the equation separately and say why they are equal in the analogy.

We will conclude this section with one more example of computing a scalar surface integral and then an activity that asks you to compute one for yourself using a parameterization of a surface.

Example 13.12.9 Let us return to one of the problems from Activity 13.12.2. Specifically, we will compute the answer to part a of that activity. We will compute $\iint_{S_1} x\,dS$ where S_1 is the top half ($z \geq 0$) of the sphere of radius one centered at the origin. Thinking in terms of spherical coordinates is helpful for determining a parameterization of S_1. Here we can use $\vec{r}(s,t) = \langle \cos(s)\sin(t), \sin(s)\sin(t), \cos(t)\rangle$ where $0 \leq s \leq 2\pi$ and $0 \leq t \leq \pi/2$, which comes from thinking of $s = \theta$ and $t = \phi$ in spherical coordinates (and $\rho = 1$). Using Theorem 13.12.3, we need to compute the partial derivatives of $\vec{r}$ with respect to s and t, then compute $\|\vec{r}_s \times \vec{r}_t\|$. For our parameterization, we get

$$\begin{aligned} \vec{r}_s &= \langle -\sin(s)\sin(t), \cos(s)\sin(t), 0\rangle \\ \vec{r}_t &= \langle \cos(s)\cos(t), \sin(s)\cos(t), -\sin(t)\rangle \\ \vec{r}_s \times \vec{r}_t &= -\sin(t)\langle \cos(s)\sin(t), \sin(s)\sin(t), \cos(t)\rangle \end{aligned}$$

The form given for $\vec{r}_s \times \vec{r}_t$ is simplified. Before simplifying, we have

$$\begin{aligned} &\langle -(\sin(t))^2\cos(s), -(\sin(t))^2\sin(s), \\ &\quad -(\cos(s))^2\cos(t)\sin(t) - (\sin(s))^2\cos(t)\sin(t)\rangle. \end{aligned}$$

Note that $\vec{r}_s \times \vec{r}_t = -\sin(t)\vec{r}(s,t)$, which means that $\|\vec{r}_s \times \vec{r}_t\| = \sin(t)$ because $\|\vec{r}(s,t)\| = 1$ for any s or t and $\sin(t) \geq 0$ for $0 \leq t \leq \pi/2$.

We can now use Theorem 13.12.3 to compute the scalar surface integral as

$$\iint_{S_1} x\,dS = \int_0^{\pi/2}\int_0^{2\pi} [\cos(s)\sin(t)]\,(\sin(t))\,ds\,dt$$

$$= [\sin(s)]_0^{2\pi} \int_0^{\pi} (\sin(t))^2 \, dt = 0 \cdot \int_0^{\pi} (\sin(t))^2 \, dt = 0.$$

This result should match your answer for part a of Subsection 13.12.2 since the values of the function $f(x, y, z) = x$ are symmetric but opposite in sign for the parts of the surface in the octants with $x > 0$ as compared to the octants with $x < 0$. ▲

Activity 13.12.4 Compute the value of the scalar surface integral in part c of Subsection 13.12.2. That is, compute $\iint_{S_3} x + z \, dS$ where S_3 is the disc of radius one centered at $(1, 0, 0)$ on the plane $x = 1$. Explain why your answer makes sense geometrically.

Summary

- Scalar surface integrals are defined in terms of double Riemann sums of the product of the value of a scalar-valued function at a point on the surface and the area patch of the tangent plane to the surface at that point.
- A scalar surface integral measures the total accumulation of a scalar-valued function on the surface.
- Scalar surface integrals can be efficiently computed by parameterizing the surface of integration as $\vec{r}(s, t)$ and then integrating $\iint_D f(s, t) \, \|\vec{r}_s \times \vec{r}_t\| \, dA$, where D is the domain of the parameterization.

13.12.4 Exercises

1. This WebWork problem uses a slightly different notation which you might see in other sources. In this problem, the parameters used are u and v and the parameterization is given by $\vec{\Phi}(u, v)$. The tangent vectors created by taking partial derivatives are notated by $\vec{T}_u$ and $\vec{T}_v$, and the normal vector to the surface is given by $\vec{n}(u, v)$.

Show that $\Phi(u, v) = (4u + 9, u - v, 7u + v)$ parametrizes the plane $2x - y - z = 18$. Then:

(a) Calculate $\mathbf{T}_u$, $\mathbf{T}_v$, and $\mathbf{n}(u, v)$.

(b) Find the area of $\mathcal{S} = \Phi(\mathcal{D})$, where $\mathcal{D} = (u, v) : 0 \le u \le 9, 0 \le v \le 2$.

(c) Express $f(x, y, z) = yz$ in terms of u and v and evaluate $\iint_{\mathcal{S}} f(x, y, z) \, d\mathcal{S}$.

(a) $\mathbf{T}_u =$ _____ , $\mathbf{T}_v =$ _____ , $\mathbf{n}(u, v) =$ _____

(b) $Area(S) =$ _____

(c) $\iint_{\mathcal{S}} f(x, y, z) \, d\mathcal{S} =$ _____

2. Calculate $\iint_{\mathcal{S}} f(x, y, z) \, dS$ For

$$y = 5 - z^2, \qquad 0 \le x, z \le 7; \qquad f(x, y, z) = z$$

$\iint_{\mathcal{S}} f(x, y, z) \, dS =$ _____

3. Evaluate $\iint_S \sqrt{1 + x^2 + y^2} \, dS$ where S is the helicoid: $\mathbf{r}(u, v) = u\cos(v)\mathbf{i} + u\sin(v)\mathbf{j} + v\mathbf{k}$, with $0 \le u \le 3, 0 \le v \le 2\pi$

4. Find the surface integral $\iint_\sigma f(x, y, z) \, dS$ where $f(x, y, z) = \left(x^2 + y^2\right) z$

and σ is the sphere $x^2 + y^2 + z^2 = 25$ above $z = 4$.

Parameterize the surface integral

$$\iint_{\sigma} f(x,y,z)\,dS = \iint_{R} f\left(x(\theta,\phi), y(\theta,\phi), z(\theta,\phi)\right) \left\| \frac{\partial r}{\partial \theta} \times \frac{\partial r}{\partial \phi} \right\| dA$$

$$= _\int__\int _____ \left\| _____ \right.$$

Note: For θ type theta and for ϕ type phi.

5. Consider an empty thin ice cream cone standing with its tip at the origin. The height of the cone is 8 and the radius of the top is 6. Find the center of gravity of the cone by following the steps below. Assume the density of the cone is constant 1.

a. We parametrize the cone S with $s(r,t) = (r\cos(t), r\sin(t),$ _____) where $0 \le t \le 2\pi$ and $0 \le r \le 6$.

b. The partial derivatives are $s_r(r,t) =$ (_____,_____,_____) and $s_t(r,t) =$ (_____,_____,_____).

c. The cross product is $s_r(r,t) \times s_t(r,t) =$ (_____,_____,_____).

d. The norm of the cross product is $\|s_r(r,t) \times s_t(r,t)\| =$ _____.

e. $m = \int\int_S 1 dS =$ _____

f. $m_z = \int\int_S z dS =$ _____

g. The center of gravity is $(0,0,\bar{z})$ where $\bar{z} = \frac{m_z}{m} =$ _____.

6. Show that $\Phi(u,v) = (9u+3, u-v, 17u+v)$ parametrizes the plane $2x - y - z = 6$. Then:

(a) Calculate $\mathbf{T}_u$, $\mathbf{T}_v$, and $\mathbf{n}(u,v)$.

(b) Find the area of $\mathcal{S} = \Phi(\mathcal{D})$, where $\mathcal{D} = (u,v) : 0 \le u \le 2, 0 \le v \le 7$.

(c) Express $f(x,y,z) = yz$ in terms of u and v and evaluate $\iint_{\mathcal{S}} f(x,y,z)\,d\mathcal{S}$.

(a) $\mathbf{T}_u =$ _____ , $\mathbf{T}_v =$ _____ , $\mathbf{n}(u,v) =$ _____

(b) $Area(S) =$ _____

(c) $\iint_{\mathcal{S}} f(x,y,z)\,d\mathcal{S} =$ _____

7. Calculate $\iint_{\mathcal{S}} f(x,y,z)\,dS$ For

$$y = 1 - z^2, \qquad 0 \le x, z \le 9; \qquad f(x,y,z) = z$$

$\iint_{\mathcal{S}} f(x,y,z)\,dS =$ _____

8. Calculate $\iint_{\mathcal{S}} f(x,y,z)\,dS$ For

$$x^2 + y^2 = 4, \quad 0 \le z \le 5; \qquad f(x,y,z) = e^{-z}$$

$\iint_{\mathcal{S}} f(x,y,z)\,dS =$ _____

9. Calculate $\iint_{\mathcal{S}} f(x,y,z)\,dS$ For

Part of the surface $x = z^3$, where $0 \le x, y \le 5^{-\frac{3}{2}}$; $\quad f(x,y,z) = x$

$\iint_{\mathcal{S}} f(x,y,z)\,dS =$ _____

13.13 Stokes' Theorem

Motivating Questions

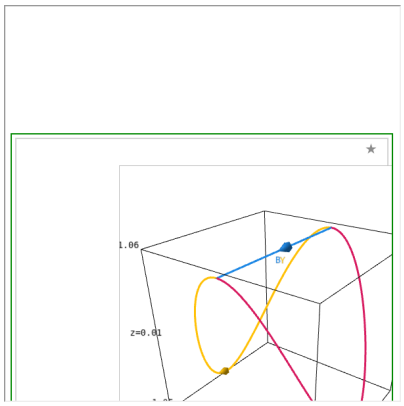
- What is the relationship between the circulation of a vector field $\vec{F}$ along a simple closed curve C in three-dimensional space and the flux integral of curl$(\vec{F})$ through a surface with C as its boundary?
- Why does the flux integral of curl$(\vec{F})$ through a surface with boundary only depend on the boundary of the surface and not the shape of the surface's interior?

13.13.1 Introduction

When we studied Green's Theorem in Section 13.9, we saw how integrating the circulation density over a region in the plane bounded by a simple closed curve is equivalent to calculating the circulation along the boundary curve. When we consider simple closed curves in $\mathbb{R}^3$, the situation gets more complicated. However, there is an interesting, and perhaps surprising, generalization of Green's Theorem for us to examine.

Preview Activity 13.13.1 In this activity, we will look at how we can apply the ideas about circulation along overlapping curves from the beginning of Subsection 13.9.2 to curves in space.

(a) For this part, consider the curves in Figure 13.13.1, where the yellow curve is Y, the blue curve is B, and the magenta curve is M.



Standalone
In Context
Embed

Figure 13.13.1 Three curves in space

You should also go back and refamiliarize yourself with our notation for combining paths (as used in line integrals) from Convention 13.3.13. In our convention, $Y+M$ would be closed loop, but $Y-M$ would not make sense because the segment $-M$ does not begin where Y begins.

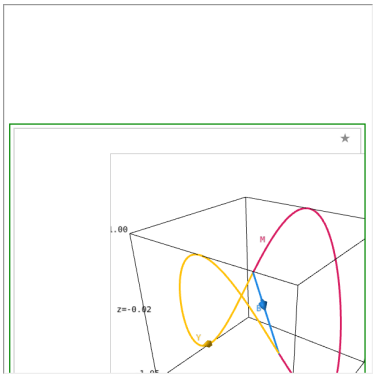
(i) Using the three segments in Figure 13.13.1, write out at least four different closed curves in terms of B, Y, and M. (Remember to consider orientation!)

(ii) Let $C_1 = M + B$ and $C_2 = Y - B$. Describe the curve given by $C_1 + C_2$.

(iii) Write a couple of sentences explaining how the circulation around C_1+C_2 would compare to the circulation around C_1 and the circulation around C_2. Write an equation in terms of $\int_{C_1} \vec{F} \cdot d\vec{r}$, $\int_{C_2} \vec{F} \cdot d\vec{r}$,

and $\int_{C_1+C_2} \vec{F} \cdot d\vec{r}$.

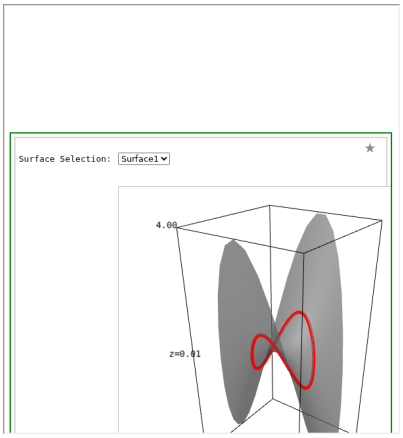
(iv) Explain how your arguments or equations from any of the parts above would or would not change if you considered the curves depicted in Figure 13.13.2.



Standalone
In Context
Embed

Figure 13.13.2 Three slightly different curves in space

(b) Let C be the simple closed curve consisting of the yellow and magenta curves in Figure 13.13.1. You can see C plotted in red in Figure 13.13.3. The drop-down allows you to select three different surfaces. You can visually verify that each of the three surfaces contains C. Notice that the scale on the z-axis changes as you select different surfaces.



Standalone
In Context
Embed

Figure 13.13.3 Surfaces containing a common simple closed curve

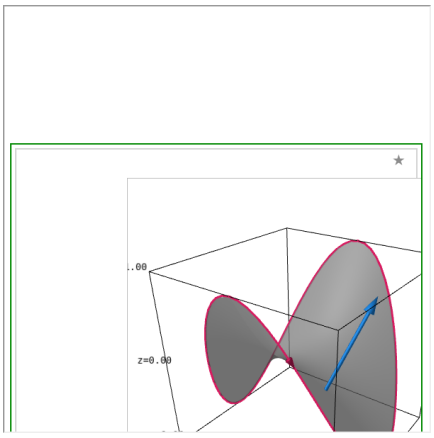
The simple closed curve consisting of the yellow and magenta curves in Figure 13.13.1 can be parameterized by $\langle \cos(t), \sin(t), \cos(2t) \rangle$ with $0 \leq t \leq 2\pi$. Let $C_3 = Y + M$. Use the given parameterization of C_3 to show that C_3 is on each of the following surfaces:

- $x^2 - y^2 = z$
- $z = x^4 - y^4$
- $z = 1 - 2y^2$
- $z = -\cos(\pi\sqrt{x^2 + y^2})(x^2 - y^2)$

13.13.2 Circulation in three dimensions and Stokes' Theorem

In part b of Subsection 13.13.1, we saw that a simple closed curve in $\mathbb{R}^3$ can bound many different surfaces. For now, however, we want to focus on a smooth surface S in $\mathbb{R}^3$ that has a well-defined normal vector $\vec{n}$ at every point and a boundary curve C. We will use the normal vector to define an orientation of C so that if a person were to walk along C in the direction of the orientation with the top of their head pointing in the direction of $\vec{n}$, their left arm would be over the surface S. Notice that this is the same convention that we used with Green's Theorem if we assume that the normal vector being used is $\hat{k}$.

In Figure 13.13.4, we show the curve C from Preview Activity 13.13.1 in magenta as well as a surface S that has C as its boundary. The chosen normal vector $\vec{n}$ to S is shown, as is the orientation of C that matches $\vec{n}$.



Standalone
In Context
Embed

Figure 13.13.4 A surface S (and normal vector) bounded by an oriented simple closed curve C

Thinking back to Green's Theorem, our main idea was that we could calculate the circulation around a simple closed curve in $\mathbb{R}^2$ by taking the double integral of the circulation density over the region bounded by the curve. As we saw in Preview Activity 13.13.1, we can break up $\oint_C \vec{F} \cdot d\vec{r}$ into line integrals around other simple closed curves so that overlapping portions are oriented oppositely just as we did with the square grid for Green's Theorem. To find a three-dimensional analog of Green's Theorem, we require that a simple closed curve C in three dimensions bound a smooth surface S with a normal vector $\vec{n}$. In doing this, we can choose our "smaller" curves similar to the squares we used in Green's Theorem to lie on the surface S. This gives us almost all the ingredients used in Green's Theorem, but we still need to find a suitable replacement for the circulation density.

As we saw in Section 13.8, the curl of a vector field in $\mathbb{R}^3$ measures the rotation of the vector field. Theorem 13.8.22 says that for a unit vector $\vec{v}$, the scalar $(\text{curl}(\vec{F})(a,b,c)) \cdot \vec{v}$ measures the rotational strength of $\vec{F}$ at the point (a,b,c) around the axis defined by $\vec{v}$. When $\vec{v}$ is the normal vector to the surface S at the point (a,b,c), we have the appropriate analog for the circulation density of $\vec{F}$ on S at (a,b,c). Thus, the equivalent idea to integrating the circulation density of a two-dimensional vector field over a region in the plane is calculating the flux integral $\iint_D \text{curl}(\vec{F}) \cdot (\vec{r}_s \times \vec{r}_t)\, dA$, where $\vec{r}(s,t)$ on the domain D that gives a parameterization of the smooth surface S.

A rigorous proof of the following theorem is beyond the scope of this text. However, part a of Subsection 13.13.1 and our discussion of Green's Theorem

provide an intuitive description of why this theorem is true.

Theorem 13.13.5 Stokes' Theorem. *Let S be a smooth surface in $\mathbb{R}^3$ with a simple closed curve C as its boundary. Let $\vec{F}$ be a vector field that is smooth on S and C. Suppose that $\vec{r}(s,t)$ on the domain D gives a parametrization of S for which C is oriented so that a person walking along C in the direction of its orientation and head pointing in the direction of the normal vector $\vec{r}_s \times \vec{r}_t$ would have their left hand above S. In this setting, the circulation of $\vec{F}$ along C is equal to the flux of* $\text{curl}(\vec{F})$ *through S. That is,*

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_D (\text{curl}(\vec{F})) \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

13.13.3 Verifying and Applying Stokes' Theorem

In this subsection, we will look at some examples and activities that will verify Stokes' Theorem by calculation both side for a few different situations.

Example 13.13.6 In this example, we will verify Stoke's Theorem for the curve used in Preview Activity 13.13.1 using the parameterization and surfaces from part b of Subsection 13.13.1. We will use the vector field $\vec{F} = \langle z - x, y + z, xy \rangle$ throughout this problem.

(a) We first calculate the circulation of $\vec{F} = \langle z - x, y + z, xy \rangle$ around the curve C with parameterization given by $\vec{r}(t) = \langle \cos(t), \sin(t), \cos(2t) \rangle$ with $0 \le t \le 2\pi$. Note here that $\vec{r}'(t) = \langle -\sin(t), \cos(t), -2\sin(2t) \rangle$. Applying Theorem 13.4.6 to calculate the circulation, we have

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^{2\pi} \langle \cos(2t) - \cos(t), \sin(t) + \cos(2t), \cos(t)\sin(t) \rangle \cdot \langle -\sin(t), \cos(t), -2\sin(2t) \rangle\, dt$$

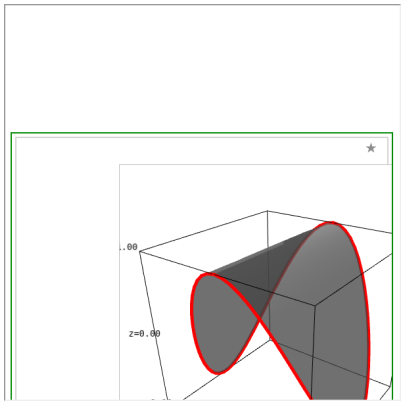
If you were to write out this dot product and combine like terms, then you would be left with four terms. Each of these can be evaluated using a few trig identities and substitutions. In fact, three of the four terms will correspond to functions that have as much area below the axis as above and thus will have integrals of zero.

$$\begin{aligned} \int_C \vec{F} \cdot d\vec{r} = &\int_0^{2\pi} 2\cos(t)\sin(t)\, dt \\ &- \int_0^{2\pi} \cos(2t)\sin(t)\, dt \\ &+ \int_0^{2\pi} \cos(2t)\cos(t)\, dt \\ &- \int_0^{2\pi} 2\cos(t)\sin(t)\sin(2t)\, dt \\ &= 0 - 0 + 0 - \pi = -\pi \end{aligned}$$

The circulation of $\vec{F}$ around C_3 is $-\pi$. This result being negative means that more of the vector field moves opposite the direction of travel given by the parameterization.

(b) In this part, we will calculate the flux of $\text{curl}(\vec{F})$ through a surface with boundary C. As demonstrated by part b of Subsection 13.13.1, there are

several different surfaces that we can use in this problem. For our first case, we will use the part of the surface $z = 1 - 2y^2$ that is bounded by C as shown in Figure 13.13.7. This surface can be parameterized by $\vec{r}(s,t) = \langle s\cos(t), s\sin(t), 1 - 2s^2\sin(t)^2\rangle$ with $0 \le s \le 1$ and $0 \le t \le 2\pi$.



Standalone
In Context
Embed

Figure 13.13.7 A portion of the surface $z = 1 - 2y^2$ bounded by C

Using our parameterization, we have the following for the partial derivative functions and the corresponding normal vector:

$$\begin{aligned}
\vec{r}_s(s,t) &= \langle \cos(t), \sin(t), -4s\sin(t)^2\rangle \\
\vec{r}_r(s,t) &= \langle -s\sin(t), s\cos(t), -4s^2\sin(t)\cos(t)\rangle \\
\vec{w} = \vec{r}_s \times \vec{r}_t &= \langle -4s^2\sin(t)^2\cos(t) + 4s^2\sin(t)^2\cos(t), \\
&\qquad 4s^2\cos(t)^2\sin(t) + 4s^2\sin(t)^2\sin(t), s\cos(t)^2 + s\sin(t)^2\rangle \\
&= \langle 0, 4s^2\sin(t), s\rangle
\end{aligned}$$

In particular, our parameterization yields $\vec{r}_s \times \vec{r}_t\, dA = \langle 0, 4s\sin(t), 1\rangle s\, dsdt$. Since we used polar coordinates as our parameter, $dA = sdsdt$.

If you completed Exercise 24 of Section 13.11, you may recognize our calculation of $\vec{r}_s \times \vec{r}_t$ as fitting the following form: If $z = f(x,y)$, then $\vec{r}_s \times \vec{r}_t\, dA$ is of the form $\langle -\dfrac{\partial f}{\partial x}, -\dfrac{\partial f}{\partial y}, 1\rangle\, dA$.

The next step in setting up the flux integral on the right side of Stokes' Theorem is calculating the curl of $\vec{F}$. Since $\vec{F} = \langle z - x, y + z, xy\rangle$, $\text{curl}(\vec{F}) = \langle x - 1, 1 - y, 0\rangle$ and converting to s and t gives $\text{curl}(\vec{F})(s,t) = \langle s\cos(t) - 1, 1 - s\sin(t), 0\rangle$. Now we are able to apply Theorem 13.11.6, which give the following iterated integral:

$$\int_0^1 \int_0^{2\pi} \langle s\cos(t) - 1, 1 - s\sin(t), 0\rangle \cdot \langle 0, 4s\sin(t), 1\rangle s\, dt\, ds.$$

Evaluating this dot product and computing each integral gives

$$\int_0^1 \int_0^{2\pi} -4s^3\sin(t)^2 + 4s^2\sin(t)\, dtds = -\pi,$$

which matches our result for the calculation of the circulation around C.

(c) As we saw in part b of Subsection 13.13.1 there is more than one surface that has C as a boundary. We will calculate the flux integral (the right side of Stokes' Theorem) for a different surface to help motivate why

it will not matter which surface we use (as long as we have the correct orientation and boundary). For this part we will use the surface $z = x^2 - y^2$, which will be parameterized by $\vec{r}(s,t) = \langle s\cos(t), s\sin(t), s^2(\cos(t)^2 - \sin(t)^2)\rangle$ with $0 \le s \le 1$ and $0 \le t \le 2\pi$. Additionally, we will use the trig identity $\cos(2t) = \cos(t)^2 - \sin(t)^2$ to write the last component of our parameterization as $s^2\cos(2t)$.

Using our parameterization, we have the following for the partial derivative functions and the corresponding normal vector:

$$\begin{aligned}
\vec{r}_s(s,t) &= \langle \cos(t), \sin(t), 2s\cos(2t)\rangle \\
\vec{r}_r(s,t) &= \langle -s\sin(t), s\cos(t), -2s^2\sin(2t)\rangle \\
\vec{w} = \vec{r}_s \times \vec{r}_t &= \langle -2s^2\sin(2t)\sin(t) - 2s^2\cos(2t)\cos(t), \\
&\qquad 2s^2\cos(t)\sin(2t) - 2s^2\sin(t)\cos(2t), s\cos(t)^2 + s\sin(t)^2\rangle \\
&= \langle -2s^2\cos(t), 2s^2\sin(t), s\rangle
\end{aligned}$$

(There are a number of equivalent algebraic simplifications that can be done here by choosing different trigonometric identities.)

Now we are ready to set up the flux integral as the following iterated integral:

$$\int_0^1 \int_0^{2\pi} \langle s\cos(t) - 1, 1 - s\sin(t), 0\rangle \cdot \langle -2s^2\cos(t), 2s^2\sin(t), s\rangle \, dt\, ds$$

You probably noticed that this integral looks slightly more complicated than our work in the previous part, but if we are consistent, we should get the same result. Evaluating this dot product and computing each integral gives

$$\begin{aligned}
&\int_0^1 \int_0^{2\pi} (-2s^2\cos(t))(s\cos(t) - 1) + (2s^2\sin(t))(1 - s\sin(t))\, dt\, ds \\
&= \int_0^1 \int_0^{2\pi} -2s^3\cos(t)^2 + 2s^2\cos(t) + 2s^2\sin(t) - 2s^3\sin(t)^2\, dt\, ds \\
&= \int_0^1 \int_0^{2\pi} -2s^3 + 2s^2\cos(t) + 2s^2\sin(t)\, dt\, ds.
\end{aligned}$$

Because both sine and cosine will integrate to zero over the interval from 0 to 2π, we only need to evaluate

$$2\pi \int_0^1 -2s^3\, ds = -\pi,$$

which gives exactly the same result as our circulation integral *and* our flux integral with the other surface.

▲

We close this subsection with a pair of activities. The first focuses on calculating both of the integrals in Stokes' Theorem. The second asks you to calculate some line integrals along simple closed curves and gives you the discretion to choose the best method to use for this (as well as the best surface to use, if you choose Stokes' Theorem).

Activity 13.13.2 In this activity, we will verify Stokes' Theorem by calculating both a line integral and a flux integral.

(a) Consider the vector field $\vec{F} = \langle x^2, y^2, z^2 \rangle$ and the circle C_1 parameterized as $\vec{r}(t) = \langle \sqrt{2}\cos(t), \sqrt{2}\cos(t), 2\sin(t) \rangle$ for $0 \leq t \leq 2\pi$.

(i) Calculate $\oint_{C_1} \vec{F} \cdot d\vec{r}$ directly using the given parametrization.

(ii) Let S_1 be the hemisphere of the sphere of radius 2 centered at the origin with $y \leq x$. Calculate the flux of $\text{curl}(\vec{F})$ through S_1.

(iii) What could you have observed about $\vec{F}$ that would have gotten you the same answer without doing either of the above calculations?

(b) Consider the vector field $\vec{G} = x\hat{\imath} + y^2 z\hat{\jmath} + x^2\hat{k}$ and the curve C_2, which is the triangle with vertices $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ with orientation corresponding to the order the points are listed here.

(i) Find the circulation of $\vec{G}$ along C_2 by calculating the appropriate line integrals.

(ii) The vertices of C_2 lie in a plane. Let S_2 be the portion of this plane lying in the first octant, i.e., the portion with $x, y, z \geq 0$. Find the flux of $\text{curl}(\vec{G})$ through S_2.

(iii) Write a sentence to explain why the sign of your answer to the previous two parts makes sense.

Activity 13.13.3

(a) Find the circulation of $\vec{F} = \langle 3yz, xz, -xy \rangle$ along the curve C consisting of (given in order of the orientation) the quarter-circle of radius 1 centered at $(0,-2,0)$ in the plane $y = -2$ from $(0,-2,1)$ to $(1,-2,0)$, the line segment from $(1,-2,0)$ to $(1,5,0)$, the quarter-circle of radius 1 centered at $(0,5,0)$ in the plane $y = 5$ from $(1,5,0)$ to $(0,5,1)$, and the line segment from $(0,5,1)$ to $(0,-2,1)$.

(b) Find the circulation of $\vec{G} = 3z^2\hat{\imath} - (z^2 + 2x)\hat{\jmath} + zy\hat{k}$ along the circle in the xy-plane of radius 3 centered at the origin. Assume the counterclockwise orientation of the circle.

In part part b of Subsection 13.13.3, there are two "reasonable" choices for the surface bounded by the circle. If you did not do so while doing the activity, we encourage you to identify both of them and compare which one makes doing the flux integral easier. In general, this will vary depending on the curl of the vector field in question, so we cannot give a rule for determining what surface or coordinate system to use. However, we do encourage you to think about which surface will make evaluating the flux integral easiest.

13.13.4 Practice with Surfaces and their Boundaries

When we looked Green's Theorem, it was generally most useful when we were given a line integral and we calculated it using a double integral. In fact, except in the circumstances described in Exercise 12 and Exercise 14 of Section 13.9, we did not use Green's Theorem to rewrite a double integral as a line integral because of the difficulty of finding a suitable vector field. The situation for Stokes' Theorem will be similar, with the exception of Exercise 14 in this section. However, Stokes' Theorem gives us an interesting additional piece of freedom: selecting the surface S through which we calculate the flux of $\text{curl}(\vec{F})$ from amongst possibly several reasonable surfaces with boundary C. The next two activities focus on the relationships between surfaces and their boundary.

Activity 13.13.4 Because Stokes' Theorem requires us to consider a surface (with normal vector) and the boundary of the surface, this activity will give you a chance to practice identifying the boundary of some surfaces in $\mathbb{R}^3$. For each surface below:

i Describe the boundary in words.

ii Find a parametrization for the boundary.

iii Ensure that a person walking along the boundary in the direction of your parametrization with head pointing in the direction of the surface's normal vector would hold their left hand over the surface.

(a) The surface S_1 is the portion of the sphere $x^2 + y^2 + z^2 = 4$ with $z \geq x$. Assume the outward orientation on the sphere.

(b) The surface S_2 is the portion of the sphere $x^2 + y^2 + z^2 = 4$ with $z \geq 0$. Assume the outward orientation on the sphere.

(c) The surface S_3 is the portion of the hyperbolic paraboloid $z = x^2 - y^2$ with $x^2 + y^2 \leq 1$. Assume the "upward" orientation, e.g., the normal vector at $(0, 0, 0)$ is $\hat{k}$.

(d) The surface S_4 is the portion of the cylinder $x^2 + y^2 = 4$ for which $-2 \leq z \leq 2$, assuming the outward orientation.

Hint. It is fine for the boundary of a surface to be made up of more than one curve. Think carefully about how each piece is oriented!

Activity 13.13.5 In some sense, this activity considers the reverse problem of that considered in Activity 13.13.4. Here, each part of the activity gives you an oriented simple closed curve C in $\mathbb{R}^3$, and your task is to find

- a surface S so that C is the boundary of S and
- a normal vector for the S so that a person walking along C in the direction of the given orientation with head pointing in the direction of your chosen normal vector would have their left hand over S.

You are encouraged to think about multiple possible answers, since as we saw in Preview Activity 13.13.1, there may be more than one reasonable choice of a surface with a particular boundary.

(a) The curve C is the triangle with vertices $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ with orientation corresponding to the order the points are listed here.

(b) The curve C is the circle parameterized as $\vec{r}(t) = \langle \sqrt{2}\cos(t), \sqrt{2}\cos(t), 2\sin(t) \rangle$ for $0 \leq t \leq 2\pi$.

(c) The curve C consists of (given in order of the orientation)

- the quarter-circle C_1 of radius 2 centered at the origin in the xy-plane from $(2, 0, 0)$ to $(0, 2, 0)$,
- the line segment C_2 from $(0, 2, 0)$ to $(0, 2, 2)$,
- the quarter-circle C_3 of radius 2 centered at $(0, 0, 2)$ in the plane $z = 2$ from $(0, 2, 2)$ to $(2, 0, 2)$, and
- the line segment C_4 from $(2, 0, 2)$ to $(2, 0, 0)$.

Summary

- Stokes' Theorem tells us that we can calculate the circulation of a smooth vector field along a simple closed curve in $\mathbb{R}^3$ that bounds a surface (with normal vector) on which the vector field is also smooth by calculating the flux of the curl of the vector field through the surface.
- Given two surfaces S_1 and S_2 with the same boundary C (and assuming normal vectors that give the same orientation on C), the flux of curl($\vec{F}$) through S_1 and through S_2 is the same because Stokes' Theorem tells us that this flux is equal to the circulation of $\vec{F}$ along C.

13.13.5 Exercises

1. Find $\int_C \vec{F} \cdot d\vec{r}$ where C is a circle of radius 2 in the plane $x + y + z = 8$, centered at $(3, 3, 2)$ and oriented clockwise when viewed from the origin, if $\vec{F} = 4y\vec{i} - x\vec{j} + 3(y - x)\,\vec{k}$

$\int_C \vec{F} \cdot d\vec{r} =$ _____

2. Use Stokes' Theorem to evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$ where $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + 5(x^2 + y^2)\mathbf{k}$ and C is the boundary of the part of the paraboloid where $z = 25 - x^2 - y^2$ which lies above the xy-plane and C is oriented counterclockwise when viewed from above.

3. Verify Stokes' theorem for the helicoid $\Psi(r, \theta) = \langle r\cos\theta, r\sin\theta, \theta\rangle$ where (r, θ) lies in the rectangle $[0, 1] \times [0, \pi/2]$, and $\mathbf{F}$ is the vector field $\mathbf{F} = \langle 6z, 2x, 6y\rangle$.

First, compute the surface integral:

$\iint_M (\nabla \times \mathbf{F}) \cdot d\mathbf{S} = \int_a^b \int_c^d f(r, \theta) dr\, d\theta$, where

$a =$_____, $b =$_____, $c =$_____, $d =$_____, and

$f(r, \theta) =$ _____ (use "t" for theta).

Finally, the value of the surface integral is _____.

Next compute the line integral on that part of the boundary from $(1, 0, 0)$ to $(0, 1, \pi/2)$.

$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_a^b g(\theta)\, d\theta$, where

$a =$_____, $b =$_____, and

$g(\theta) =$ _____ (use "t" for theta).

4. Use Stokes' Theorem to evaluate $\iint_M (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$ where M is the hemisphere $x^2 + y^2 + z^2 = 25, x \geq 0$, with the normal in the direction of the positive x direction, and $\mathbf{F} = \langle x^7, 0, y^2\rangle$.

Begin by writing down the "standard" parametrization of ∂M as a function of the angle θ (denoted by "t" in your answer)

$x =$_____, $y =$_____, $z =$_____.

$\int_{\partial M} \mathbf{F} \cdot d\mathbf{s} = \int_0^{2\pi} f(\theta)\, d\theta$, where

$f(\theta) =$_____ (use "t" for theta).

The value of the integral is _____.

5. Use Stokes' theorem to evaluate $\iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$ where $\mathbf{F}(x, y, z) = -16yz\mathbf{i} + 16xz\mathbf{j} + 8(x^2 + y^2)z\mathbf{k}$ and S is the part of the paraboloid $z = x^2 + y^2$ that

lies inside the cylinder $x^2 + y^2 = 1$, oriented upward.

6. Calculate the circulation, $\int_C \vec{F} \cdot d\vec{r}$, in two ways, directly and using Stokes' Theorem. The vector field $\vec{F} = (8x - 8y + 7z)(\vec{i} + \vec{j})$ and C is the triangle with vertices $(0, 0, 0)$, $(3, 0, 0)$, $(3, 5, 0)$, traversed in that order.

Calculating directly, we break C into three paths. For each, give a parameterization $\vec{r}(t)$ that traverses the path from start to end for $0 \leq t \leq 1$.

On C_1 from $(0, 0, 0)$ to $(3, 0, 0)$, $\vec{r}(t) =$ _____

On C_2 from $(3, 0, 0)$ to $(3, 5, 0)$, $\vec{r}(t) =$ _____

On C_3 from $(3, 5, 0)$ to $(0, 0, 0)$, $\vec{r}(t) =$ _____

So that, integrating, we have $\int_{C_1} \vec{F} \cdot d\vec{r} =$ _____

$\int_{C_2} \vec{F} \cdot d\vec{r} =$ _____

$\int_{C_3} \vec{F} \cdot d\vec{r} =$ _____

and so $\int_C \vec{F} \cdot d\vec{r} =$ _____.

Using Stokes' Theorem, we have

curl $\vec{F} =$ _____

So that the surface integral on S, the triangular region on the plane enclosed by the indicated triangle, is

$\int_S \text{curl } \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dy\,dx$,

where $a =$ _____, $b =$ _____, $c =$ _____, and $d =$ _____.

Integrating, we get $\int_C \vec{F} \cdot d\vec{r} = \int_S \text{curl } \vec{F} \cdot d\vec{A} =$ _____

7. Calculate the circulation, $\int_C \vec{F} \cdot d\vec{r}$, in two ways, directly and using Stokes' Theorem. The vector field $\vec{F} = 2y\vec{i} - 2x\vec{j}$ and C is the boundary of S, the part of the surface $z = 16 - x^2 - y^2$ above the xy-plane, oriented upward.

Note that C is a circle in the xy-plane. Find a $\vec{r}(t)$ that parameterizes this curve.

$\vec{r}(t) =$ _____,

with _____ $\leq t \leq$ _____

(Note that answers must be provided for all three of these answer blanks to be able to determine correctness of the parameterization.)

With this parameterization, the circulation integral is

$\int_C \vec{F} \cdot d\vec{r} = \int_a^b$ _____ dt, where a and b are the endpoints you gave above.

Evaluate your integral to find the circulation: $\int_C \vec{F} \cdot d\vec{r} =$ _____

Using Stokes' Theorem, we equate $\int_C \vec{F} \cdot d\vec{r} = \int_S \text{curl } \vec{F} \cdot d\vec{A}$. Find curl $\vec{F} =$ _____.

Noting that the surface is given by $z = 16 - x^2 - y^2$, find

$d\vec{A} =$ _____ $dy\,dx$.

With R giving the region in the xy-plane enclosed by the surface, this gives

$\int_S \text{curl } \vec{F} \cdot d\vec{A} = \int_R$ _____ $dy\,dx$.

Evaluate this integral to find the circulation:

$\int_C \vec{F} \cdot d\vec{r} = \int_S \text{curl } \vec{F} \cdot d\vec{A} =$ _____.

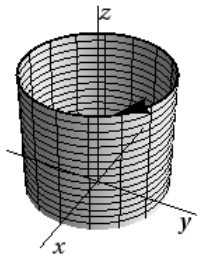
8. Use Stokes' Theorem to find the circulation of the vector field $\vec{F} = xz\vec{i} + (3x + 5yz)\vec{j} + 5x^2\vec{k}$ around the paths. C, is the circle $x^2 + y^2 = 4$, $z = 4$, oriented counterclockwise when viewed from above.

circulation = _____

9. Find $\int_C \vec{F} \cdot d\vec{r}$ where C is a circle of radius 1 in the plane $x + y + z = 2$, centered at $(1, 2, -1)$ and oriented clockwise when viewed from the origin, if $\vec{F} = -5z\vec{j} + 2y\vec{k}$

$\int_C \vec{F} \cdot d\vec{r} =$ _____

10. The figure below open cylindrical can, S, standing on the xy-plane. (S has a bottom and sides, but no top.)



The side of S is given by $x^2 + y^2 = 4$, and its height is 4.

(a) Give a parametric equation, $\vec{r}(t)$ for the rim, C.

$\vec{r}(t) =$ _____, with _____ $\leq t \leq$ _____.

(For this problem, enter your vector equation with angle-bracket notation: $< f(t), g(t), h(t) >$.)

(b) If S is oriented outward and downward, find $\int_S \text{curl } (-4y\vec{i} + 4x\vec{j} + z\vec{k}) \cdot d\vec{A}$.

$\int_S \text{curl } (-4y\vec{i} + 4x\vec{j} + z\vec{k}) \cdot d\vec{A} =$ _____

11. If curl $\vec{F} = (x^2 + z^2)\vec{j} + 5\vec{k}$, find $\int_C \vec{F} \cdot d\vec{r}$, where C is a circle of radius 2, centered at the origin, with

(a) C in the xy-plane, oriented counterclockwise when viewed from above.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

(b) C in the xz-plane, oriented counterclockwise when viewed from the positive y-axis.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

12. ***(a)*** Let $\vec{F} = (3y - 3z)\vec{i} + (3x + 3z)\vec{j} + 4xy\vec{k}$, let C be the circle of radius 4 centered at $(5, 2, 0)$ in the xy-plane, oriented counterclockwise when viewed from above. Compute $\int_C \vec{F} \cdot d\vec{r}$.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

(b) Is $\vec{F}$ path-independent (conservative)? (☐ yes ☐ no)

(Be sure you can explain why this is the case.)

13. Let $\vec{F} = \left(6z + 6x^2\right)\vec{\imath} + \left(5y + 6z + 6\sin\left(y^2\right)\right)\vec{\jmath} + \left(6x + 6y + 5e^{z^2}\right)\vec{k}$.

(a) Find curl $\vec{F}$.

curl $\vec{F} =$ _____

(b) What does your answer to part (a) tell you about $\int_C \vec{F} \cdot d\vec{r}$ where C is the circle $(x-20)^2 + (y-5)^2 = 1$ in the xy-plane, oriented clockwise?

$\int_C \vec{F} \cdot d\vec{r} =$ _____

(c) If C is any closed curve, what can you say about $\int_C \vec{F} \cdot d\vec{r}$?

$\int_C \vec{F} \cdot d\vec{r} =$ _____

(d) Now let C be the half circle $(x-20)^2 + (y-5)^2 = 1$ in the xy-plane

with $y > 5$, traversed from $(21, 5)$ to $(19, 5)$. Find $\int_C \vec{F} \cdot d\vec{r}$ by using your result from (c) and considering C plus the line segment connecting the endpoints of C.

$\int_C \vec{F} \cdot d\vec{r} =$ _____

14. Stokes' Theorem is generally used to turn a line integral into a flux integral. Sometimes it is possible to be given a flux integral and recognize that the given vector field $\vec{F}$ is curl$(\vec{G})$ for some vector field $\vec{G}$, however. When this is the case, we call $\vec{G}$ a **vector potential** for the vector field $\vec{F}$, much like a function f so that $\vec{F} = \nabla(f)$ is called a potential function for $\vec{F}$.

(a) Find a vector field $\vec{F}$ so that curl$(\vec{F}) = \langle x - y^2, 4xy - y - 4, -4xz \rangle$.

Hint. Your experience in finding potential functions for gradient vector fields will be useful to you here, although you will have more flexibility.

(b) When finding an anti-derivative of a function of a single variable, you know that there is an infinite family of anti-derivatives, but that any two anti-derivatives differ by a constant. This is why we write expressions such as $\int \cos(x)\,dx = \sin(x) + C$. A similar phenomenon occurs with (scalar) potential functions for gradient vector fields. Find a *second* vector field $\vec{G}$ with the same curl as in part a, and do so in a way that $\vec{F} - \vec{G}$ is *not* a constant vector. That is, after simplifying fully, $\vec{F} - \vec{G}$ must contain at least one of the variables x, y, z.

(c) Verify that for the vector fields you found above, $\vec{F} - \vec{G}$ is a gradient vector field. Explain why for *every* pair $\vec{F}, \vec{G}$ of vector potentials for a vector field $\vec{H}$, you must have that $\vec{F} - \vec{G}$ is a gradient vector field.

(d) Explain why if $\vec{H}$ is a vector field with a vector potential $\vec{F}$, div$(\vec{H}) = 0$. Such a vector field is called a **solenoidal vector field** or **divergence-free vector field**.

15. Repeat the steps of part 13.13.6.b and part 13.13.6.c with the surface $z = x^4 - y^4$ to verify Stokes' Theorem for another different surface. Your flux integral should calculate to $-\pi$ as in Example 13.13.6.

13.14 The Divergence Theorem

Motivating Questions

- What is a closed surface in $\mathbb{R}^3$?
- How can we efficiently calculate the flux through a closed surface in $\mathbb{R}^3$ when that surface must be parametrized in several pieces?

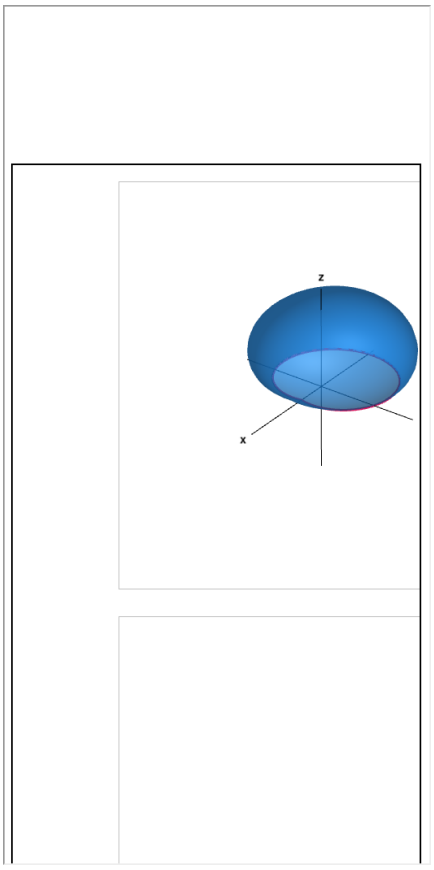
13.14.1 Introduction

In Section 13.7 we examined vector fields to consider how the strength of a vector field changes in different regions. In particular, we developed the divergence of a vector field as a local measurement (based on density) of how the strength of the vector field changes. In particular, we did this by looking

at the flux of the vector field through a closed path in two dimensions and then generalized these ideas to higher dimensions.

In Section 13.11, we measured how much of a vector field flowed through a section of a surface in three dimensions as a generalization of our argument from Section 13.7. In this section, we will connect the ideas of flux of a vector field through a closed surface in three dimensions and the divergence of that vector field.

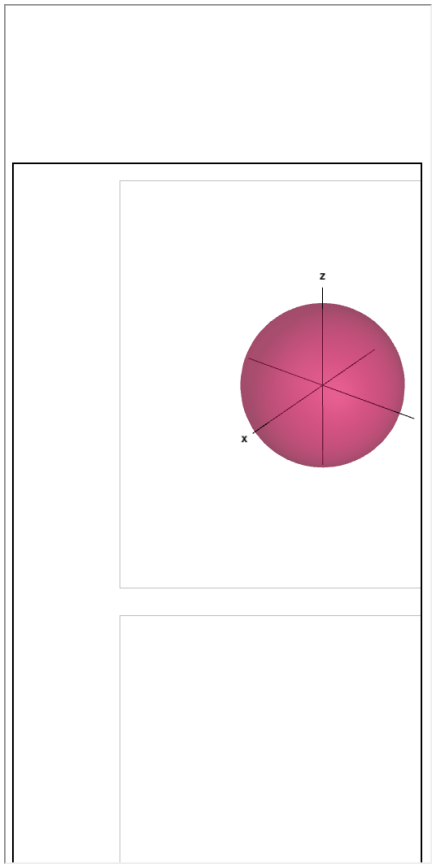
Recalling the surfaces we studied in Section 13.13, where we applied Stokes' Theorem, notice that all of these surfaces had the property that they had a **boundary** along which we calculated circulation. It turns out that giving a precise definition for boundary is challenging. For our purposes, however, you might think of a boundary curve of a surface as being a curve along which you could walk with your arms stretched out on either side with one arm lying over the surface and the other arm not lying over the surface as you walk. In this section, we will study **closed** surfaces in three dimensions, which are those surfaces without boundary. In Section 13.13, *none* of the surfaces were closed, because each had a boundary curve. On the other hand, in Figure 13.14.1, we show two more surfaces that are *not* closed, as demonstrated by the magenta curves marking the edge where the surface ends.



Standalone
In Context
Embed

Figure 13.14.1 Two surfaces that are *not* closed

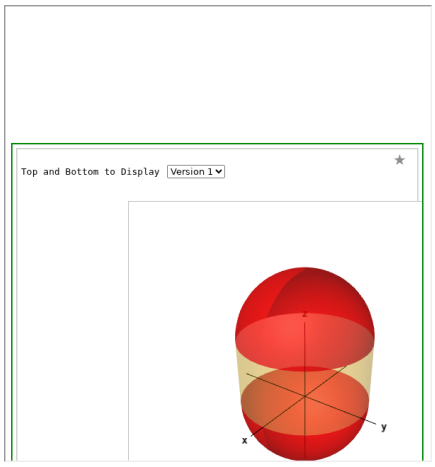
In fact, the yellow cylinder has two edges where the surface ends (does not meet itself). The surfaces in Figure 13.14.2 *are* closed because there is no edge to the surface. Note that the cylinder has been "filled in" with a cap at top and bottom (plotted in gray and green, respectively) to become a closed surface.



Standalone
In Context
Embed

Figure 13.14.2 Two closed surfaces

Closed surfaces can be used to define the boundary of a volume in space. If we have the top and bottom on our cylinder, we have a well-defined volume of space, in that we know which points are "inside" the volume and which are "outside" of the volume. With different top and bottom surfaces, the enclosed volume would be different. Figure 13.14.3 illustrates three different ways to complete the cylindrical surface into a closed surface.



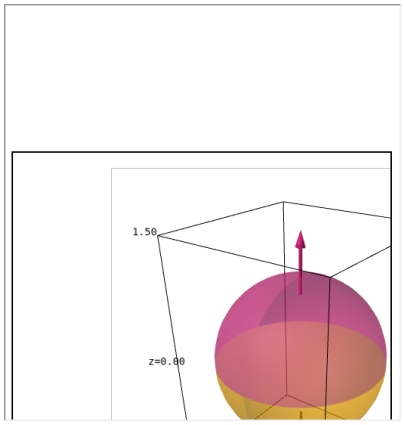
Standalone
In Context
Embed

Figure 13.14.3 Three ways to close a cylindrical surface by adding a top and bottom

Preview Activity 13.14.1 Locating sources of a poisonous gas. We will use three different surfaces to examine the flux through closed surfaces. Let S_{top} be the top half of the unit sphere centered at the origin (graphed in magenta in the figures below). Let S_{bottom} be the bottom half of the unit sphere centered at the origin (graphed in yellow). Finally, let S_{mid} be the unit disk centered at the origin in the xy-plane (graphed in blue). With these definitions, S_{top} and S_{bottom} will make a closed surface given by the unit sphere. The surfaces S_{top} and S_{mid} will enclose the top half of the unit ball, while S_{bottom} and S_{mid} will enclose the bottom half of the unit ball.

In this problem we will be using the surfaces defined above and the flux integrals of a poisonous gas through these surfaces to try to determine whether different regions of space are emitting or absorbing the poisonous gas.

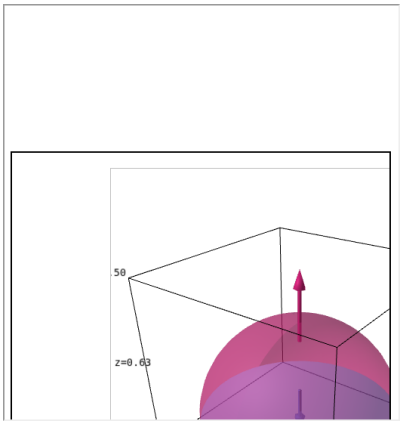
(a) In this part, we will consider the unit ball shown in Figure 13.14.4, with boundary and normal vectors as shown in the plot. If the flux integral of a poisonous gas through S_{top} is 15 and the flux integral of the poison gas through S_{bottom} is -3, is the interior of the sphere emitting or absorbing poisonous gas? Explain your reasoning.



Standalone
In Context
Embed

Figure 13.14.4 Unit ball with boundary given by the combination of S_{top} and S_{bottom}

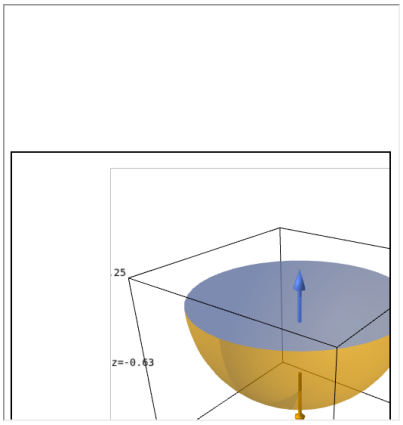
(b) In this part, we will consider the top half of the unit ball shown in Figure 13.14.5, with boundary and normal vectors as shown in the plot. If the flux integral of a poisonous gas through S_{top} is 15 and the flux integral of the poison gas through S_{mid} is -20, is the top half of the unit ball emitting or absorbing poisonous gas? Explain your reasoning.



Standalone
In Context
Embed

Figure 13.14.5 Upper half of the unit ball with boundaries given by S_{top} and S_{mid}

(c) In this part, we will consider the bottom half of the unit ball shown in Figure 13.14.6, with boundary and normal vectors as shown in the plot. Based on the information given in the previous two parts, what will the flux integrals of the poison gas be for S_{bottom} and S_{mid} be in this case? Be sure to pay attention to the orientation of what we consider positive flow. Explain your reasoning.



Standalone
In Context
Embed

Figure 13.14.6 Lower half of the unit ball with boundaries given by S_{mid} and S_{bottom}

(d) Using your answer from the previous part, is the bottom half of the unit ball emitting or absorbing poisonous gas? Explain your reasoning.

The preview activity showed how the flux through a closed surface can be subdivided into the flux through surfaces which combine to be the closed surface (with orientation switches corresponding to additive inverse). The net flux through the closed surface measures the net amount of the vector field that is created or destroyed on the interior of the closed surface.

13.14.2 The Divergence Theorem

The divergence of a vector field was developed as a measurement of the density with which the strength of vector field is changing. In three dimensions, the divergence measures the density per unit volume in which the vector field is being created or destroyed. This means that if we integrate the divergence of a vector field over a volume of space, we will get the net amount of the vector

field that is created or destroyed in that particular volume of space. Since the net amount of a vector field that is created or destroyed in a volume of space is the same as the net flux of the vector field through the closed surface that is the boundary of that volume, we have a correspondence between a triple integral of the divergence of a vector field on the interior of a closed surface and the flux integral of the vector field over the closed surface.

Theorem 13.14.7 Divergence Theorem. *Let S be a closed surface in three dimensional space and let Q be the volume bounded by S. If $\vec{F}$ is a vector field that has continuous first partial derivatives on Q and S, then*

$$\iint_S \vec{F} \cdot \vec{N}\, dS = \iiint_Q \operatorname{div}(\vec{F})\, dV$$

where the integral on the left measures the flux through S in terms of the outward pointing normal vector.

The preview activity and the discussion before the statement of the Divergence Theorem have hopefully given you some intuition as to why the theorem is true. The ideas should also seem similar to the manner in which we approached Green's Theorem and Stokes' Theorem. In the next example, we will verify the Divergence Theorem for a particular case.

Example 13.14.8 In this example we will apply the Divergence Theorem to the vector field $\vec{F} = \langle xy - z, yz + e^x, z(x-y)\rangle$ and the region in the first octant bounded above by $z = 1 - x - y$. We will go through the calculations of the flux integral on the right side and the triple integral on the left side.

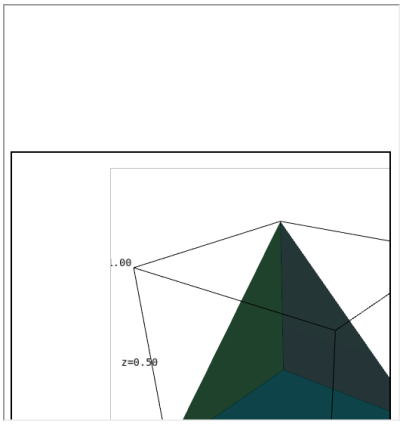
(a) We first calculate the left side of the Divergence Theorem using four flux integrals (one for each of the four boundary surfaces) as illustrated in Figure 13.14.9. We will need to parametrize each of the four faces which we will call S_1 (in magenta), S_2 (in yellow), S_3 (in blue), and S_4 (in green).

$$\begin{aligned}
\vec{r}_1(s,t) &= \langle 0, s, t\rangle \\
\vec{r}_2(s,t) &= \langle s, 0, t\rangle \\
\vec{r}_3(s,t) &= \langle s, t, 0\rangle \\
\vec{r}_4(s,t) &= \langle s, t, 1-s-t\rangle
\end{aligned}$$

where all of these parameterization use $0 \le s \le 1$ and $0 \le t \le 1 - s$. These parameterizations have the corresponding normal vectors:

$$\begin{aligned}
\vec{n}_1 &= \langle 1, 0, 0\rangle \\
\vec{n}_2 &= \langle 0, -1, 0\rangle \\
\vec{n}_3 &= \langle 0, 0, 1\rangle \\
\vec{n}_4 &= \langle 1, 1, 1\rangle
\end{aligned}$$

Notice that $\vec{n}_1$ and $\vec{n}_3$ point into Q while $\vec{n}_2$ and $\vec{n}_4$ point out of Q. We will take into account the idea that we will need to calculate the flow out of Q when we sum our flux integrals later.



Standalone
In Context
Embed

Figure 13.14.9 The region of integration Q with each of the four faces in a different color.

We set up and evaluate each of these flux integrals using Theorem 13.11.6:

$$S_1 : \int_0^1 \int_0^{1-s} \langle (0)s - t, st - e^1, t(0-s)\rangle \cdot \langle 1,0,0\rangle \, dt\, ds$$
$$S_2 : \int_0^1 \int_0^{1-s} \langle s(0) - t, 0(t) - e^s, t(s-0)\rangle \cdot \langle 0,-1,0\rangle \, dt\, ds$$
$$S_3 : \int_0^1 \int_0^{1-s} \langle st - 0, t(0) - e^s, 0(s-t)\rangle \cdot \langle 0,0,1\rangle \, dt\, ds$$
$$S_4 : \int_0^1 \int_0^{1-s} \langle st - (1-s-t), t(1-s-t) + e^s, (1-s-t)(s-t)\rangle \cdot \langle 1,1,1\rangle \, dt\, ds$$

A strategy we can use to make our calculations more efficient is to note that we can subtract the first and third integrals from the second and fourth (remember the direction of flux) and do one integral. We can do this because the bounds on *all* of our parameterizations is the same. If we write out everything, we have:

$$S_1 : \int_0^1 \int_0^{1-s} -t \, dt\, ds$$
$$S_2 : \int_0^1 \int_0^{1-s} -e^s \, dt\, ds$$
$$S_3 : \int_0^1 \int_0^{1-s} 0 \, dt\, ds$$
$$S_4 : \int_0^1 \int_0^{1-s} -1 + 2s + t + e^s - s^2 \, dt\, ds$$

Combining these we have:

$$S_4 - S_1 + S_2 - S_3 : \int_0^1 \int_0^{1-s} -1 + 2s + 2t - s^2 \, dt\, ds$$
$$\int_0^1 (s-1)^2 s \, ds$$

which evaluates to $\frac{1}{12}$.

(b) Next, we calculate the triple integral on the right side of the Divergence Theorem for our example. Since $\vec{F} = \langle xy - z, yz + e^x, z(x-y)\rangle$, we can

calculate

$$\begin{aligned}\operatorname{div}(\vec{F}) &= \frac{\partial}{\partial x}(xy - z) + \frac{\partial}{\partial y}(yz - e^x) + \frac{\partial}{\partial z}(z(x - y)) \\ &= y + z + (x - y) = z + x\end{aligned}$$

The region Q can be described by the bounds $0 \le x \le 1$, $0 \le y \le 1 - x$, and $0 \le z \le 1 - x - y$. Thus, our triple integral I will be set up and evaluated as

$$\begin{aligned}I &= \int_0^1 \int_0^{1-x} \int_0^{1-x-y} z + x \, dz \, dy \, dx \\ &= \int_0^1 \int_0^{1-x} 1/2(1 - x - y) + x(1 - x - y) \, dy \, dx \\ &= \int_0^1 \frac{1}{12}(\frac{x^3}{3}) \, dx\end{aligned}$$

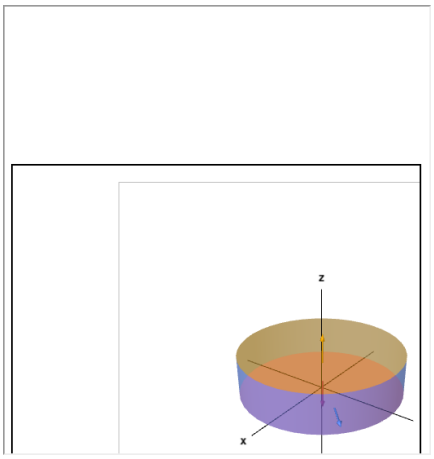
This evaluates to $\frac{1}{12}$, which matches with the flux calculations we did in the previous step to verify the Divergence Theorem for our example.

▲

The next activity walks you through evaluating both the flux integrals necessary to calculate the flux directly and the triple integral given in the Divergence Theorem for a specific vector field and closed surface.

Activity 13.14.2 In this activity, we will look at calculating both sides of a non-trivial example of the Divergence Theorem. We will look at the region inside the right circular cylinder shown in Figure 13.14.10. Let S be the closed surface formed by combining S_{top} (in yellow), S_{sides} (in blue), and S_{bottom} (in magenta). The solid volume Q is the volume bounded by S. The region shown has radius 2, and its height is 1. The vector field we consider in this activity is given by

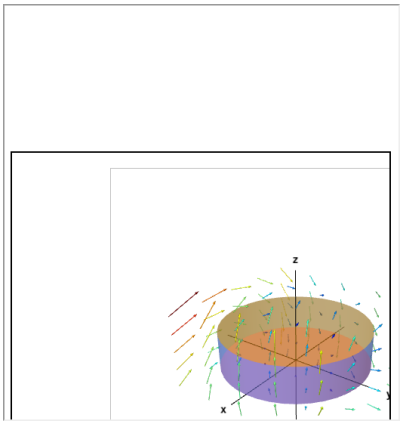
$$\vec{F} = \langle xy - 2z, y^2 - yz, 3x + z^2 \rangle.$$



Standalone
In Context
Embed

Figure 13.14.10 A closed cylindrical surface

(a) Figure 13.14.11 shows the vector field $\vec{F}$ on a region around S. Without doing any computations, write a couple of sentences to explain if you think the flux of $\vec{F}$ through S will be positive, negative, or zero.



Standalone
In Context
Embed

Figure 13.14.11 The vector field $\vec{F} = \langle xy - 2z, y^2 - yz, 3x + z^2 \rangle$ in the region near Q

(b) Parametrize each of the surfaces S_{top}, S_{sides}, and S_{bottom}. Be sure to give bounds for each of your parametrization.

(c) Give inequalities in terms of cylindrical coordinates to describe Q.

(d) Set up and evaluate double integrals to calculate the flux of $\vec{F}$ through S_{top}, S_{sides}, and S_{bottom}.

(e) What is the net flux through the closed surface? Be sure to state if the net flux is in or out.

(f) Compute the divergence of $\vec{F}$ and use this to explain whether you think $\iiint_Q \operatorname{div}(\vec{F})\, dV$ will be positive, negative, or zero.

(g) Set up and compute the triple integral for $\iiint_Q \operatorname{div}(\vec{F})\, dV$.

Hint. Use cylindrical coordinates.

(h) Verify that your answers for part e and part g are the same and thus that the Divergence Theorem holds for this example.

(i) Was the left-hand side or right-hand side of the equation in the Divergence Theorem more tedious to calculate in this example? Do you think this will be true for most other cases where the Divergence Theorem applies?

The next activity asks you to compute flux in some circumstances where it may be wise to apply Divergence Theorem.

Activity 13.14.3

(a) Find the flux of the vector field $\vec{F} = \langle 3x^2 + y^5, 5 + e^{z^3}, z \rangle$ through the surface of the solid cube Q in $\mathbb{R}^3$ with $-2 \leq x \leq 2$, $-2 \leq y \leq 2$, and $-2 \leq z \leq 2$.

(b) Find the flux of the vector field $\vec{G} = \langle x^3, y^3, z^3 \rangle$ through surface consisting of the top half of sphere of radius 3 centered at the origin and the disc of radius 3 in the xy-plane (centered at the origin).

Hint. Spherical coordinates

Summary

- A closed surface is one that has no boundary.
- The flux of a smooth vector field through a closed surface can be computed by integrating the divergence of the vector field over the volume bounded by the closed surface.

13.14.3 Exercises

1. Compute the flux integral $\int_S \vec{F} \cdot d\vec{A}$ in two ways, directly and using the Divergence Theorem. S is the surface of the box with faces $x = 3, x = 4, y = 0, y = 1, z = 0, z = 3$, closed and oriented outward, and $\vec{F} = 3x^2\vec{i} + 4y^2\vec{j} + 4z^2\vec{k}$.

Using the Divergence Theorem,

$\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d \int_p^q$ _____ $dz\,dy\,dx =$ _____,

where $a =$____, $b =$____, $c =$____, $d =$____, $p =$____ and $q =$____.

Next, calculating directly, we have $\int_S \vec{F} \cdot d\vec{A} =$ (the sum of the flux through each of the six faces of the box). Calculating the flux through each face separately, we have:

On $x = 4$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dz\,dy =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

On $x = 3$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dz\,dy =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

On $y = 1$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dz\,dx =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

On $y = 0$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dz\,dx =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

On $z = 3$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dy\,dx =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

And on $z = 0$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _____ $dy\,dx =$ _____
where $a =$____, $b =$____, $c =$____ and $d =$____.

Thus, summing these, we have $\int_S \vec{F} \cdot d\vec{A} =$_____

2. Use the divergence theorem to find the outward flux of the vector field $\mathbf{F}(x, y, z) = 5x^2\mathbf{i} + 5y^2\mathbf{j} + 3z^2\mathbf{k}$ across the boundary of the rectangular prism: $0 \le x \le 4, 0 \le y \le 3, 0 \le z \le 1$.

3. Find the flux of the vector field $\vec{J} = 4y^2z\,\vec{j} + 4x\,\vec{k}$ out of the closed box $0 \le x \le 2$, $0 \le y \le 6$, $0 \le z \le 8$.

flux = _____

4. Let div $\vec{F} = 2(2 - x)$ and $0 \le a, b, c \le 12$.

(a) Find the flux of $\vec{F}$ out of the rectangular solid $0 \le x \le a$, $0 \le y \le b$, $0 \le z \le c$.

flux = _____

(b) For what values of a, b, c is the flux largest?

$a =$ _____

$b =$ _____

$c =$ _____

(c) What is that largest flux?

flux = _____

5. Use the divergence theorem to calculate the flux of the vector field $\vec{F}(x,y,z) = x^3\vec{i} + y^3\vec{j} + z^3\vec{k}$ out of the closed, outward-oriented surface S bounding the solid $x^2 + y^2 \leq 4,\ 0 \leq z \leq 8$.

$\iint_S \vec{F} \cdot d\vec{A} =$ _____

6. Let $\mathbf{F} = (y^2+z^3, x^3+z^2, xz)$. Evaluate $\iint_{\partial W} \mathbf{F} \cdot d\mathbf{S}$ for each of the following regions W:

A. $x^2 + y^2 \leq z \leq 5$ _____

B. $x^2 + y^2 \leq z \leq 5,\ x \geq 0$ _____

C. $x^2 + y^2 \leq z \leq 5,\ x \leq 0$ _____

7. Use the Divergence Theorem to calculate the flux of the vector fields $\vec{F}_1 = -5z\vec{i} + 5x\vec{k}$ and $\vec{F}_2 = -5z\vec{i} + (5-5y)\vec{j} + 5x\vec{k}$ through the surface S given by the sphere of radius a centered at the origin with outwards orientation. Be sure that you are able to explain your answers geometrically.

With W giving the interior of the sphere,
$\int_S \vec{F}_1 \cdot d\vec{A} = \int_W$ _____ $dV =$ _____
and
$\int_S \vec{F}_2 \cdot d\vec{A} = \int_W$ _____ $dV =$ _____

8. The region W lies between the spheres $x^2+y^2+z^2 = 4$ and $x^2+y^2+z^2 = 16$ and within the cone $z = \sqrt{x^2+y^2}$ with $z \geq 0$; its boundary is the closed surface, S, oriented outward. Find the flux of $\vec{F} = x^3\vec{i} + y^3\vec{j} + z^3\vec{k}$ out of S.

flux = _____

9. Find the flux of $\vec{F} = 4z\,\vec{i} + 4y^2\,\vec{j} + 3x\,\vec{k}$ out of the closed cone $y = \sqrt{x^2+z^2}$, with $0 \leq y \leq 3$.

flux = _____

10. A cone has its tip at the point $(0,0,8)$ and its base the disk $D, x^2+y^2 \leq 1$, in the plane $z = 1$. The surface of the cone is the curved and slanted face, S, oriented upward, and the flat base, D, oriented downward. The flux of the constant vector field $\vec{F} = a\vec{i} + b\vec{j} + c\vec{k}$ through S is given by

$$\int_S \vec{F} \cdot d\vec{A} = 5.23.$$

What is $\int_D \vec{F} \cdot d\vec{A}$?
$\int_D \vec{F} \cdot d\vec{A} =$ _____
*(Enter **indeterminate** if it is not possible to find a value given the information provided.)*

Supposed we instead consider the vector field $\vec{F} = a\vec{i} + b\vec{j} + cz\vec{k}$. If we again know

$$\int_S \vec{F} \cdot d\vec{A} = 5.23.$$

What is $\int_D \vec{F} \cdot d\vec{A}$ in this case?
$\int_D \vec{F} \cdot d\vec{A} =$ _____
*(Again, enter **indeterminate** if it is not possible to find a value given the information provided.)*

11. Use the Divergence Theorem to calculate the flux of $\mathbf{F}$ across S, where $\mathbf{F} = 1x^2yz\boldsymbol{i} + 2xy^2z\boldsymbol{j} + 3xyz^2\boldsymbol{k}$ and S is the surface of the box enclosed by the planes $x = 0, x = 3, y = 0, y = 3, z = 0$ and $z = 1$.

$\int\int_S \mathbf{F} \cdot\ d\mathbf{S} =$ _____

12. Use the Divergence Theorem to calculate the flux of $\mathbf{F}$ across S, where $\mathbf{F} = z\boldsymbol{i} + y\boldsymbol{j} + zx\boldsymbol{k}$ and S is the surface of the tetrahedron enclosed by the coordinate planes and the plane

$$\frac{x}{2} + \frac{y}{3} + \frac{z}{4} = 1$$

$\int\int_S \mathbf{F} \cdot \, d\mathbf{S} =$ _____

13. Let $\mathbf{F}(x, y, z) = 1z^2x\mathbf{i} + (\frac{1}{3}y^3 + \tan(z))\mathbf{j} + (1x^2z - 1y^2)\mathbf{k}$. Use the Divergence Theorem to evaluate $\int_S \mathbf{F} \cdot \, d\mathbf{S}$ where S is the top half of the sphere $x^2 + y^2 + z^2 = 1$ oriented upwards.

$\int\int_S \mathbf{F} \cdot \, d\mathbf{S} =$ _____

14. Let $\mathbf{F}(x, y, z) = z\tan(y^2)\mathbf{i} + z^3\ln(x^2 + 1)\mathbf{j} + 4\mathbf{k}$. Use the Divergence Theorem to find the flux of $\mathbf{F}$ across the part of the paraboloid $x^2 + y^2 + z = 2$ that lies above the plane $z = 1$ and is oriented upwards.

$\int\int_S \mathbf{F} \cdot \, d\mathbf{S} =$ _____

15. Let Q be the volume enclosed by $x = 0$, $x = 1$, $y = 0$, $y = 1$, $z = 0$, and $z = 1$. Compute the flux of $\vec{F} = \langle z^2 - xy, 4yz + cos(x/\pi), e^{xyz} \rangle$ through each of the six cube faces.

16. Let Q be the volume given in cylindrical coordinates by $0 \leq z \leq 3$, $1 \leq r \leq 2$, and $0 \leq \theta < 2\pi$. Give an example of a vector field with component functions that are linear in x, y, and z such that the flux of the vector field through the boundary of Q is 17.

17. Let Q be the volume given in spherical coordinates by $0 \leq \rho \leq 3$, $0 \leq \phi \leq \pi/4$, and $0 \leq \theta < 2\pi$. Give an example of a vector field with component functions that are not linear or constant such that the flux of the vector field through the boundary of Q is 25.

18. Calculate the flux of the given vector field $\vec{F}$ through the surface S for each situation below by using an appropriate triple integral from the Divergence Theorem.

(a) The vector field $\vec{F} = \langle 2x + 3\sin(yz), -4y + e^{x^2}, 7z + \arctan(y^5) \rangle$ through the tetrahedron S with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 2, 0)$, and $(0, 0, 2)$.

(b) The vector field $\vec{F} = \langle xy^2, yz^2, zx^2 \rangle$ through S, the portion of the sphere of radius 3 centered at the origin having $x \geq 0$.

Hint. By itself, S is not a closed surface. However, you should think about what additional surface S' you need to add to make a closed surface, then use those pieces to apply the Divergence Theorem.

Appendix A

Supplementary Materials

A.1 Conic Sections

A.1.1 Introduction

In this section, we will introduce some examples of graphs in two dimensions that have nice algebraic properties and offer a variety of interesting geometric features. For our first subsection, we will recall some properties of graphs and equations that will be useful in describing our new examples. While students may have different levels of experience with graphs like circles, parabolas, ellipses, and hyperbolas, we will focus our work on the correspondence between the algebraic presentations of these shapes and important geometric features.

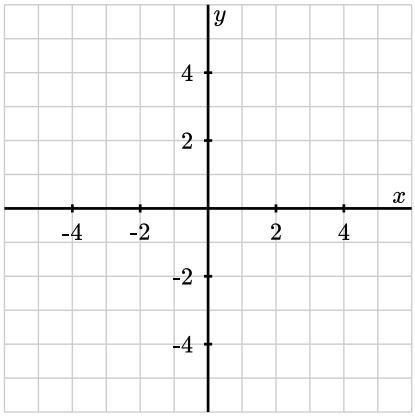
A.1.2 Coordinate Transformations

We have seen at the end of Section 9.2 how the distance formula leads to the equation of a sphere and how the idea of flat one and two dimensional graphs gave rise to lines (Section 9.6) and planes (Section 9.7). Here we introduce some other curves that are used throughout the rest of this text as examples. These objects will allow us to have graphs with a variety of geometric features while still being algebraically simple. You may already be familiar with the curves studied in this section, in which case you can consider this section a review. However, if you have not studied these curves before, we have designed this section to help you understand their key properties.

The next activity will go over a couple of basic coordinate transformation ideas that will be helpful in generalizing our common graphs.

Activity A.1.1 Translation of Coordinates. In this activity we will look at how to translate coordinate systems, which means to use a new coordinate system with axes parallel to to the original set of axes and the same scale. In other words, the new coordinate system has a new origin but does not change the way coordinates are measured. This is useful to simplify the equation of a graph by making the graph's center the new origin.

(a) On the axes below, draw and label the point $P = (4, -1)$. With a different color, draw a new set of axes that are centered at P and label these axes x^* and y^*. The x^*- and y^*-axes should parallel to the x- and y-axes and use the same scale.



(b) Use the plot of the x^* and y^* axes above to give the (x^*, y^*)-coordinates for each of the following points:

- $(x, y) = (4, -1)$
- $(x, y) = (3, -2)$
- $(x, y) = (0, 0)$
- $(x, y) = (6, 3)$

Remember that the locations of these points should not change, but rather the values of the measurements used to describe the location will change.

(c) Generalize your work for part (b) and write x^* and y^* in terms of x, y, and the coordinates of new center $(4, -1)$.

$x^* =$ $\qquad$ $y^* =$

(d) Using the same coordinate systems as in part (a), give the (x, y)-coordinates of the following points:

- $(x^*, y^*) = (0, 0)$
- $(x^*, y^*) = (-5, 7)$
- $(x^*, y^*) = (4, -1)$
- $(x^*, y^*) = (6, 3)$

(e) Generalize your work for part (d) and write x and y in terms of x, y, and the coordinates of new center $(4, -1)$.

$x =$ $\qquad$ $y =$

Activity A.1.2 In this activity, we look at how transform the coordinates of points when changing the scale of your coordinate system.

While walking your dog, you notice that your dog takes three steps for every step of yours. When you sit down at home you realize that your coordinate system for your house is different than your dog's. You decide to make the origin of both your coordinate system and your dog's to be the entrance to your kitchen. Your dog's bed is eight of your steps south and four steps west of the entrance to the kitchen. The dog's water bowl is 5 of your steps north and 3 steps east from the entrance to your kitchen. The front door of your house is 15 of your steps south and 9 steps east of the entrance to the kitchen.

(a) Draw a set of axes and plot the location of your dog's bed, water bowl, and the front door in terms of your steps from the origin (the entrance to your kitchen).

(b) State what the coordinates of your dog's bed, water bowl, and the front door in terms of your dog's steps from the origin will be.

(c) If your dog's favorite toy is 17 dog steps east and 5 dog steps north of the entrance to the kitchen and your dog's collar is 13 dog steps south and 7 dog steps west, give the coordinates of your dog's toy and collar in terms of your steps.

(d) If you call the coordinates in your steps the (x, y)-coordinate system and refer to coordinates in your dog's steps as the (x', y')-coordinate system, give the location in both coordinate systems for your dog's bed, water bowl, the front door, your dog's favorite toy, and your dog's collar.

(e) Generalize your findings from the previous part into equations that transform between (x, y)- and (x', y')-coordinates.

The results of the previous activities are probably not surprising but should offer some insight into why some aspects of coordinate transformations seem backwards. This is because the transformations are about converting back to the measurement of a different coordinate system.

A.1.3 Conic Sections

You may have seen the ideas of coordinate transformation that were central to Activity A.1.1, but hopefully you see precisely where the algebraic transformations come from and how they correspond to simple geometric transformations. For our basic shapes in two dimensions, we will give a brief definition for the shape, but the focus of this presentation will be about applying transformations to a basic shape in order to generalize the possible usage of these shapes.

A **circle** is the set of points that are a fixed distance (called the radius) away from a specific point (called the center). Most often, circles are introduced with the center at the origin and the radius given by a constant R. This means that a point (x, y) that is on the circle will satisfy

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

which can be expressed as the equivalent equation

$$x^2 + y^2 = R^2.$$

Applying the same idea to a circle with center (h, k) leads to the equation

$$(x-h)^2 + (y-k)^2 = R^2,$$

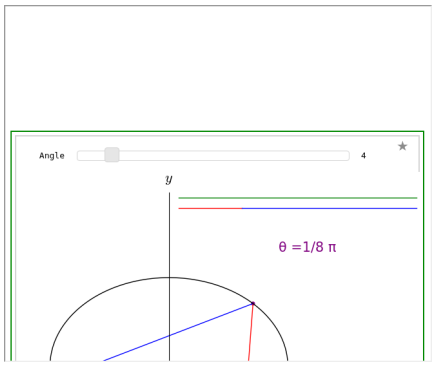
which is often called the standard form of the circle. The standard form is convenient to use because the information needed to graph the circle can be read from this form without needing more algebra. For instance, the circle given by

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

has center $(2, -3)$ and radius $\sqrt{6}$. Notice that the transformation equations from Activity A.1.1 show up in this example as we move a circle from being centered at the origin, where the equation is $x^2 + y^2 = R^2$, to the point (h, k), where the equation is $(x-h)^2 + (y-k)^2 = R^2$.

Definition A.1.1 An **ellipse** is the set of all points such that the sum of the distances from the point (x, y) to a pair of distinct points (called foci) is a fixed constant. ◆

To help illustrate the definition of an ellipse, explore Figure A.1.2. You can use the slider to change the location on the ellipse being displayed. You should see how as you move around the ellipse, the sum of the lengths of the blue and red segments remains constant, as shown compared to the green segment, which is a fixed length.



Standalone
In Context
Embed

Figure A.1.2 A plot of an ellipse with foci in green

While the definition of the ellipse given above has applications in engineering, orbital mechanics, and optics, we will focus on the ellipse as a transformation of a circle. While this is not obvious and the details take a little while to prove, any ellipse in the plane can be obtained by transforming a circle (through translation, horizontal/vertical coordinate stretches that are not the same as each other, and rotations of the coordinate systems).

Activity A.1.3

(a) What transformation is done to convert between the circle given by $x^2 + y^2 = 1$ and the graph of $\frac{x^2}{4} + y^2 = 1$? You should be specific about how the graph of $\frac{x^2}{4} + y^2 = 1$ is different than the graph of $x^2 + y^2 = 1$.

Hint. Look at the x-intercepts of each equation.

(b) What transformations are done to convert between the circle given by $x^2 + y^2 = 1$ and the graph of $\frac{x^2}{4} + \frac{y^2}{9} = 1$? You should be specific about how the graph of $\frac{x^2}{4} + \frac{y^2}{9} = 1$ is different than the graph of $x^2 + y^2 = 1$ and specify if the transformations need to be done in a particular order.

(c) What transformations are done to convert between the circle given by $x^2 + y^2 = 1$ and the graph of $\frac{(x+2)^2}{4} + \frac{(y-3)^2}{9} = 1$? You should be specific about how the graph of $\frac{(x+2)^2}{4} + \frac{(y-3)^2}{9} = 1$ is different than the graph of $x^2 + y^2 = 1$ and specify if the transformations need to be done in a particular order.

(d) Draw a plot of $\frac{(x+2)^2}{4} + \frac{(y-3)^2}{9} = 1$ and label the center of your plot and the points that demonstrate how far the ellipse is stretched in the vertical and horizontal directions.

(e) The graph of the equation $9x^2 + 16y^2 = 400$ is an ellipse. Convert this equation to the form $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$. Then use the idea of transformations from above to graph this ellipse and label all extreme points on your plot.

(f) The graph of the equation $4x^2 + y^2 + 24x - 2y + 21 = 0$ is an ellipse. Convert this equation to the form $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$. Then use the

idea of transformations from above to graph this ellipse and label all extreme points on your plot.

(g) Give the equation of the ellipse shown in Figure A.1.3.

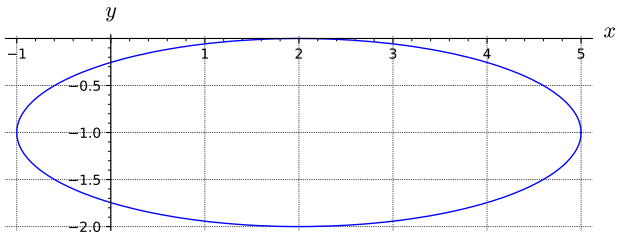
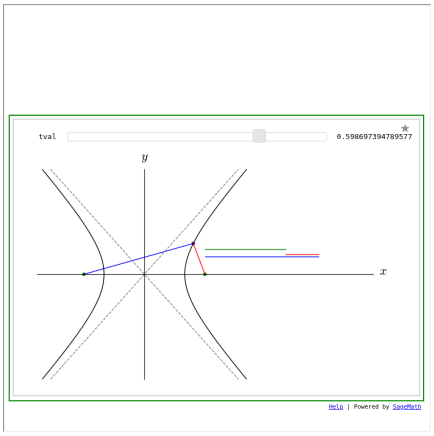


Figure A.1.3 A plot of an ellipse

Having considered the curves where the *sum* of distances to two points is constant, we now turn our attention to curves where the *difference* of the distances is constant.

Definition A.1.4 A **hyperbola** is the set of all points such that the difference of the distances from the point (x, y) to a pair of distinct points (called foci) is a fixed constant. ◆

In Figure A.1.5, you can see how as you move around the hyperbola, the difference of the lengths of the blue and red segments remains constant, as shown in green. Additionally, you can see that a hyperbola has asymptotic behavior in that as you approach the edges of the plot, the hyperbola will get very close to the dashed asymptote lines.



Standalone
In Context
Embed

Figure A.1.5 A plot of an hyperbola with foci in green

Activity A.1.4

(a) The most basic equation for a hyperbola is $x^2 - y^2 = 1$. Make a plot of the hyperbola given by $x^2 - y^2 = 1$. In your plot, be sure to include the asymptotes, which are given by $y = \pm x$, as well as the vertices. These are the points on the hyperbola that are closest to the center. In this case, the vertices are the x-intercepts.

(b) What transformation is done to change the graph of the hyperbola given by $x^2 - y^2 = 1$ to the graph of $\frac{x^2}{4} - y^2 = 1$? Be specific about how the graph of $\frac{x^2}{4} - y^2 = 1$ is different than the graph of $x^2 - y^2 = 1$.

Hint. Look at the x-intercepts of each graph.

(c) What transformation is done to change the graph of the hyperbola given by $x^2 - y^2 = 1$ to the graph of $y^2 - x^2 = 1$? Be specific about how the graph of $y^2 - x^2 = 1$ is different than the graph of $x^2 - y^2 = 1$.

Hint. Look at both intercepts of each graph.

(d) What transformations are done to change the graph of the hyperbola $x^2 - y^2 = 1$ to the graph of $\frac{x^2}{4} - \frac{y^2}{9} = 1$? Be specific about how the graph of $\frac{x^2}{4} - \frac{y^2}{9} = 1$ is different than the graph of $x^2 - y^2 = 1$ and specify if the transformations need to be done in a particular order.

(e) Draw a graph of $\frac{(x+2)^2}{4} - \frac{(y-3)^2}{9} = 1$ and label the center, the vertices, and the asymptote lines for the hyperbola. You will need to apply the transformations from the previous part to the asymptotes $y = \pm x$ of the base hyperbola in order to get the equations of the transformed asymptotes.

(f) The graph of the equation $9x^2 - 16y^2 = 400$ is a hyperbola. Convert this equation to the form $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$. Then use the idea of transformations from above to graph this hyperbola and label the center, the vertices, and the asymptote lines of the hyperbola.

(g) The graph of the equation $4x^2 - y^2 + 24x - 2y + 21 = 0$ is an hyperbola. Convert this equation to the form $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$. Then use the idea of transformations from above to graph this hyperbola and label the center, the vertices, and the asymptote lines of the hyperbola.

(h) Give the equation of the hyperbola shown in Figure A.1.6

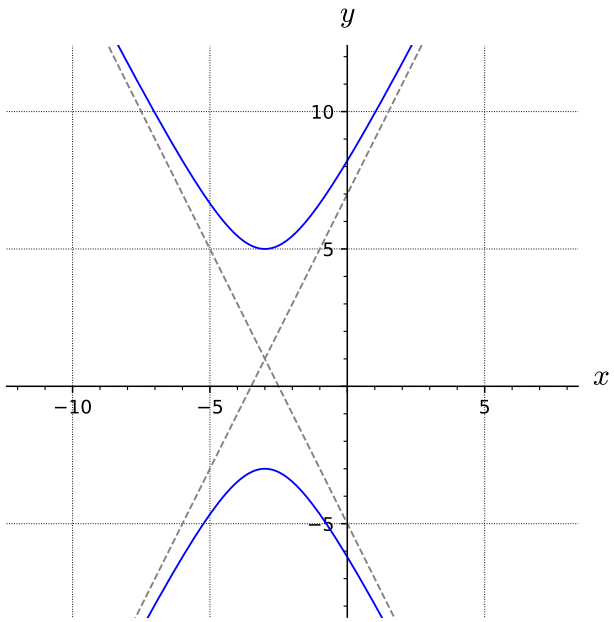


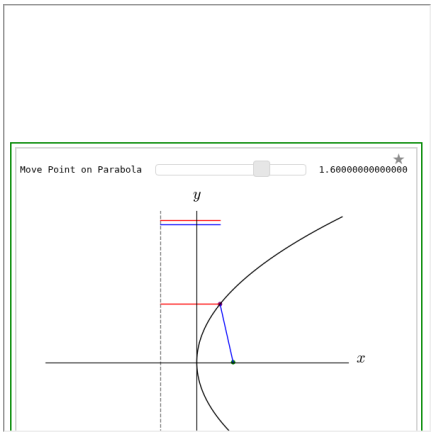
Figure A.1.6 A plot of a hyperbola

The final conic section we will study in this section involves curves whose points have the same distance from a fixed point as from a fixed *line*.

Definition A.1.7 A **parabola** is the set of all points that are the same distance from a point, called the **focus**, and a line, called the **directrix**. ◆

In Figure A.1.8, you can see how as you move around the parabola, the

lengths of the blue and red segments remain the same as each other. The point that is halfway between the focus and the directrix is referred to as the **vertex** of the parabola.



Standalone
In Context
Embed

Figure A.1.8 A plot of a parabola with the focus in green and the directrix shown as a dashed line

Activity A.1.5

(a) Make a plot of the parabola given by $x = y^2$. You should draw and label the vertex and four other points on the parabola.

(b) Draw a plot of the graph for $2x = y^2$ and label the vertex and four other points. What transformation is done to change the graph of the parabola given by $x = y^2$ to the graph of $2x = y^2$? Be specific about how the graph of $2x = y^2$ is different than the graph of $x = y^2$.

(c) Draw a plot of the graph for $y = x^2$ and label the vertex and four other points on the parabola. What transformation is done to change the graph of the parabola given by $y = x^2$ to the graph of $x = y^2$? Be specific about how the graph of $x = y^2$ is different than the graph of $y = x^2$.

(d) What transformations are done to change the graph of $x = y^2$ to the graph of $\frac{x-1}{2} = \left(\frac{y+2}{3}\right)^2$? Be specific about how the graph of $\frac{x-1}{2} = \left(\frac{y+2}{3}\right)^2$ is different than the graph of $x = y^2$ and specify if the transformations need to be done in a particular order.

(e) Draw a plot of $\frac{x-1}{2} = \left(\frac{y+2}{3}\right)^2$ and label the vertex and four other points on the parabola.

(f) The graph of the equation $x^2-8x-8y+8 = 0$ is an parabola. Convert this equation to the form $\left(\frac{x-h}{a}\right)^2 = \frac{y-k}{b}$ and use the idea of transformations from above to graph this parabola. Label the vertex and four other points on the parabola.

(g) Give the equation of the parabola shown in Figure A.1.9

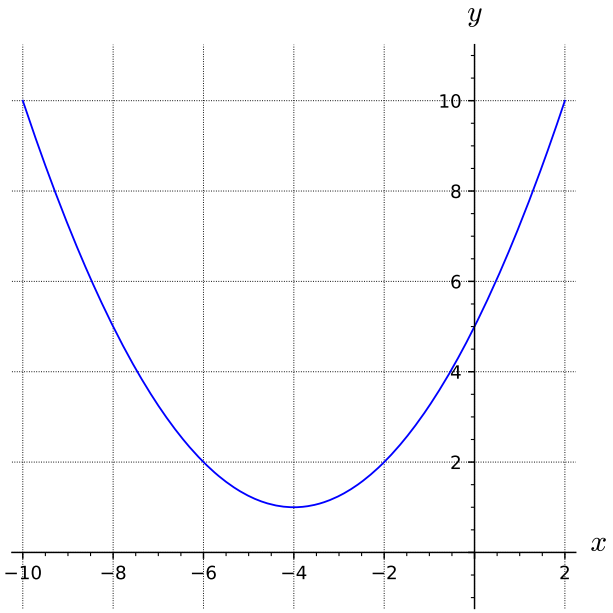


Figure A.1.9 A plot of an parabola

Summary: Summary

- Translated Conic Sections have the following forms:
  - Ellipse: $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ with center (h, k), vertices $(h \pm a, k)$ and covertices $(h, k \pm b)$
  - Hyperbola: $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ with center (h, k), vertices $(h \pm a, k)$, and asymptotes $(y - k) = \frac{b}{a}(x - h)$
  - Parabola: $(x - h) = A(y - k)^2$ with vertex/center (h, k)

  Other versions of these forms that are oriented vertically will have the roles of x and y switched.

A.1.4 Exercises

1. Match each equation below to the curve it represents. Each answer should be A, B, C, D, E, F, or G.

CURVES
A. Circle,
B. Ellipse,
C. Point,
D. Parabola,
E. Empty Set,
F. Intersecting lines,
G. Hyperbola,
EQUATIONS

2. Find the vertices and foci of the conic section $\left(\frac{x}{6}\right)^2 + \left(\frac{y}{5}\right)^2 = 1$

The foci are at $F_1 =$ (_____,_____) (left) and $F_2 =$(_____,_____) (right)

The vertices are :
(_____,_____) (on the positive y-axis) ,

(_____,_____) (negative y-axis),
(_____,_____) (positive x-axis),
(_____,_____) (negative x-axis).

3. Find the vertices and foci of the conic section $\left(\frac{x}{6}\right)^2 - \left(\frac{y}{7}\right)^2 = 1$

The foci are at F_1 = (_____,_____) (left) and F_2 =(_____,_____) (right)

The vertices are :
(_____,_____) (on the positive x-axis) ,
(_____,_____) (negative x-axis),

4. Find the vertices and foci of the conic section $\left(\frac{x-2}{6}\right)^2 - \left(\frac{y-5}{4}\right)^2 = 1$

The foci are at F_1 = (_____,_____) (left) and F_2 =(_____,_____) (right)

The vertices are :
(_____,_____) and (_____,_____) .

Note: Write the vertices from left to right, in an increasing order, starting with the vertex with the lowest x-value.

5. Consider the ellipse

$$\left(\frac{x-2}{8}\right)^2 + \left(\frac{y-9}{2}\right)^2 = 1$$

Find the equation of the ellipse translated so that its center is at the origin.

_____= 1

6. Find the equation of the ellipse with the given properties:

Vertices at $(\pm 81, 0)$ and $(0, \pm 49)$

_____= 1

7. Find the equation of the ellipse with the given properties:

Foci $(0, \pm 5)$ and two vertices at $(\pm 2, 0)$.

_____= 1

8. Find the equation of the hyperbola with the given properties

Vertices $(0, -4)$, $(0, 3)$ and foci $(0, -11)$, $(0, 10)$.

_____= 1

9. Find the equation of the parabola with the given properties

Vertex $(0, 0)$, focus $(3, 0)$.

_____= x

10. Find the equation of the parabola with the given properties

Focus $(4, 0)$, directrix $x = -4$.

_____= x

11. Find the equation of the parabola with the given properties

Focus $(5, 0)$, directrix $x = -5$.

_____= x

12. Find the vertices, foci, axes, center (if an ellipse or a hyperbola) and asymptotes (if a hyperbola) of the conic section

$$x^2 + 6y^2 = 36$$

The conic section is

☐ An ellipse

☐ A Parabola

☐ A hyperbola

The vertices are at

☐ $(0, \pm\sqrt{6})$

☐ $(\pm\sqrt{30}, 0)$

☐ $(0, 0)$

☐ $(0, \pm 36)$

☐ $(\pm 36, 0)$

☐ $(\pm 6, 0)$

The foci are at

☐ $(0, \pm\sqrt{30})$

☐ $(\pm 36, 0)$

☐ $(\pm\sqrt{30}, 0)$

☐ $(\pm 6, 0)$

The focal axis and the conjugate axis are

☐ $y = 0$ and $x = 0$ respectively.

☐ $x = 36$ and $y = 6$ respectively.

☐ $x = 0$ and $y = 0$ respectively.

☐ $y = 6$ and $x = 36$ respectively.

The conic section's center is at

☐ $(0, 0)$

☐ $(-6, -6)$

☐ $(6, 36)$

☐ $(6, 6)$

☐ $(1, 1)$

Index

additive inverse
    of vector, 24
angularly simple, 477
arc length, 178
average value over a solid, 495, 496

CCA, 135
center of mass
    of a lamina, 466
    of a solid, 495, 496
centroid
    of a lamina, 466
change of variable
    double integral, 521
    triple integral, 523
circulation, 549
Classic Calculus Approach, 135
Cobb-Douglas production function, 406
continuity, 264
coordinate, 10
coordinate plane, 6, 10
critical point, 374
cross product, 51
curvature, 203
cylinder surface, 89

differentiable function, 318
differentials, 308
directional derivative, 332, 359
discriminant, 380
dot product, 35
double integral
    average value, 418
    definition, 415
    difference in volumes, 416
    mass of lamina, 462
    probability, 463
double Riemann sum, 414

frame
    Frenet, 196
    TNB, 196
Frenet frame, 196
function
    continuously differentiable, 310
    domain, 230
    graph, 233
    locally linear, 304, 309
    of two variables, 229
    range, 230
    surface, 233
fundamental plane, 10

graph of a vector-valued function
    definition, 143

horizontally simple, 438

inverse
    additive (of vector), 24
iterated integral
    cylindrical coordinates, 505
    polar coordinates, 479
    rectangular coordinates, 425
    spherical coordinates, 506

Jacobian, 520
joint probability density function, 463

Lagrange multiplier, 400
level curve, 239
line
    direction vector, 66
    in space, 66
    parametric equations, 69
    vector equation, 67
line integral
    definition, 544
    differential form of, 560

lines
parallel, 68
skew, 68
locally linear, 62

magnitude, 27
mass
of a solid, 494–496
moments about coordinate axes, 466
multiplier
Lagrange, 400

normal vector, 76

open
definition, 570
oriented curve
closed, 549
orthogonal, 38

parallel
vectors, 26
parallel lines, 68
parameterization
curve, 143
parametric equations for a curve, 143
partial derivatives
first-order, 274
second-order, 290
second-order, mixed, 290
second-order, unmixed, 290
path-connected
definition, 570
path-independent
definition, 565
plane, 6, 10
definition, 76
fundamental, 10
scalar equation, 77
vector equation, 77
polar coordinates, 471
position vector, 19
principal unit normal vector, 212

radially simple, 477
radius of curvature, 212

scalar triple product, 56
set-builder notation, 230
simple
angularly, 477
radially, 477
skew lines, 68
smooth curve, 545
speed
of vector-valued function, 159
sphere
definition, 11
equation of, 13
standard position (of vector), 18
standard unit vector, 23
surface area, 647

TNB frame, 196
trace, 236
triple integral, 488
triple Riemann sum, 488

vector
angle between, 36
component in the direction of, 41
definition, 18
equality, 18
normal (to plane), 76
orthogonal, 38
parallel, 26
position, 19
standard position of, 18
standard unit, 23
subtraction, 23
sum, parallelogram, 25
tail, 18
unit, 28
zero, 24
vector-valued function
antiderivative, 164
definition, 143
derivative, 158
indefinite integral, 164
vectorfield
definition, 532
vertically simple, 437
volume of a solid, 494, 496

zero vector, 24