Active Calculus - Multivariable

Active Calculus - Multivariable

Steve Schlicker
Grand Valley State University

Mitchel T. Keller
University of Wisconsin-Madison

Nicholas Long
Stephen F. Austin State University

Contributing Authors
David Austin
Grand Valley State University
Matt Boelkins
Grand Valley State University

December 6, 2024

Cover Photo: Lars Jensen

Edition: 2025 DRAFT

Website: https://activecalculus.org/[1]

[1]

[2] 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 indeed to be completed by students before a day 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 National 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[3]. 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[4] text by Greg Hartmann et al.

[3] activecalculus.org/multi/
[4] www.apexcalculus.com

Graphics As much as possible, we strive to demonstrate key fundamental ideas visually, and to encourage students to do the same. Throughout the text, we use full-color graphics to exemplify and magnify key ideas, and to use this graphical perspective alongside both numerical and algebraic representations of calculus. When the text itself refers to color in images, one needs to view the .html or .pdf electronically. The figures and the software to generate them have been created by David Austin.

In Chapter 12, the authors have used interactive Sage figure to allow users of the html version to rotate and zoom on dynamic visuals. In some figures, the user can manipulate different parameters to see how a featured element will change.

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.

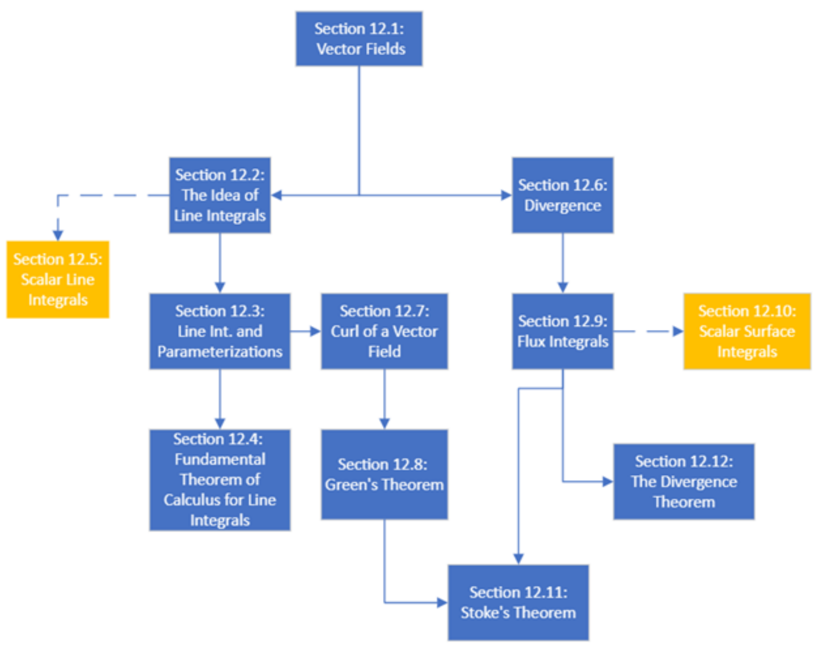
Links to technological tools Many of the ideas of multivariable calculus are best understood dynamically, and we encourage readers to make frequent use of technology to analyze graphs and data. Since technology changes so often, we refrain from indicating specific programs to use in the text. However, aside from computer algebra systems like *Maple*, *Mathematica*, or *Sage*, 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 applets 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`[5], good all-purpose 3D graphing tool
- `Wolfram Alpha`[6], useful for graphing surfaces in 2D and 3D, and for general calculations
- `Wolfram Alpha widgets`[7], searchable site for simple to use programs using Wolfram Alpha
- `GeoGebra`[8], all purpose graphing tool with some computer algebra capabilities. Clicking on the magnifying glass icon allows you to search a large database of GeoGebra applets.
- `A collection of Flash Mathlets`[9] for graphing surfaces, parametric curves in 3D, spherical coordinates and other 3D tools. Requires Flash Player.

[5] c3d.libretexts.org/CalcPlot3D/index.html
[6] www.wolframalpha.com/
[7] www.wolframalpha.com/widgets/gallery/?category=math
[8] www.geogebra.org/
[9] www.math.uri.edu/~bkaskosz/

Vector Calculus Preface

Recognizing that not all institutions will cover all the material in the chapter on vector calculus, we have created the following chart for the depencecies of the topics in Chapter 12.



In this flow chart, the yellow sections are not strictly dependent on the others but use some common arguments that would be helpful for a reader to have seen before working on.

Observant readers may notice that the graphics in Chapter 13 deviate from those in the previous 11. Because so much of vector calculus relies upon visualizing things in three dimensions, we elected to produce almost all of the 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, 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 and dragging.

Our Fall 2023 release includes at least some WeBWorK exercises in each of the sections of Chapter 13, but we welcome suggestions of additional WeBWorK exercises that faculty would consider good inclusions and feedback on any aspect of this chapter. The best way to provide feedback is through the Active Calculus - Multivariable Google Group[10]. We will occasionally give some updates through the blog on https://activecalculus.org/.

Mitchel T. Keller and Nicholas Long
Madison, Wisconsin and Nacogdoches, Texas

[10] groups.google.com/u/0/g/active-calculus-multivariable

Acknowledgments

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[11] to see updates and provide feedback to the authors.

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

[11] groups.google.com/u/0/g/active-calculus-multivariable

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. 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 and text.
- To build in students intuition for why the main ideas in multivariable calculus are natural and true. 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 and explain aloud 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[12], but you can find links to a pdf and hard copy of the text at https://activecalculus.org/[13]. Furthermore, because the text is open-source, any instructor may acquire the full set of source files, which are available on GitHub[14].

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 introduces functions of several independent variables along with tools that will be used to study these functions, namely vectors and vector-valued functions. Chapter 11 studies differentiation of functions of several independent variables in detail, addressing the typical topics including limits, partial derivatives, and optimization, while Chapter 12 provides the standard topics of integration of multivariable functions.

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 live .html 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. The .html version is far superior to the .pdf version; this is especially true for viewing on a smartphone.

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

[12] activecalculus.org/multi/
[13] activecalculus.org/
[14] github.com/StevenSchlicker/AC3PreTeXt/tree/master/src

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 all 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.

Community of Users

Because this text is free and open-source, we hope that as people use the text, they will contribute corrections, suggestions, and new material. At this time, the best way to communicate such feedback is through the Active Calculus - Multivariable Google Group[15].

[15] groups.google.com/u/0/g/active-calculus-multivariable

Contents

Chapter 9

Precalculus of Multivariable Functions

Why Multivariable? In previous calculus courses you have looked at how to measure limits, measuring the change and rate of change of a function, and the accumulation of the output of a function. All of these ideas and much of you algebra of functions has centered around functions that have one number as an input and one number as an output. Much of the meaning and tools for these ideas and associated measurements came from algebra and geometry that you learned in the years before you started calculus. We will need to do a bit of precalculus again before we look at our new calculus ideas.

The calculus you have done can't deal with graphs where $y \neq f(x)$. While you talked about motion in previous calculus courses, we will need to expand our tools, coordinates, and measurements to apply our work to motion in three dimensions, as would be needed to describe realistic physical problems. Many important quantities depend on more than one number as an input. 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 in 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. With our new an expanded view of the world around us, we will also need to describe some new basic examples of our mathematical objects that are algebraically nice to deal with and have a variety of features. While these ideas took years of math classes for you to understand and apply to the calculus of 1-variable in and 1 variable out, we will expand our tools in much less time because of all of the depth and variety we have already learned.

9.1 Three Dimensional Space

Motivating Questions

- How can you describe a location in three dimensional space?
- How do you measure distance between points in three or more dimensions?

As you read in the introduction to this chapter, we will be expanding beyond the idea of measuring change of expressions with one number as an input or output. We first need to think about how to represent the basic ideas of coordinates and functions in a more general setting. The following preview activity asks you to recall some algebraic and geometric ideas from before your single variable calculus course.

Preview Activity 9.1.1 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 9.1.1? Draw segments from the point P to the vertical and horizontal axes to demonstrate the measurements of the coordinates.

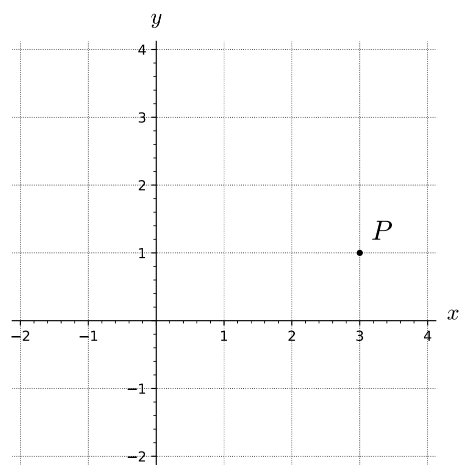


Figure 9.1.1 A 2D plot with point P labeled

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

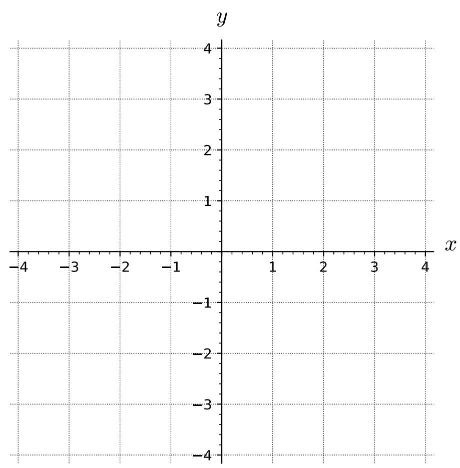


Figure 9.1.2 Two-dimensional axes for plotting points

(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 $-x = y$

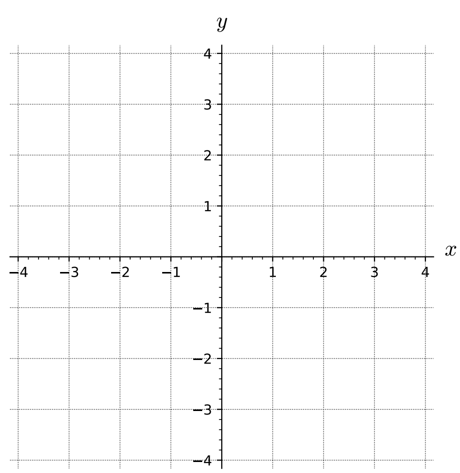


Figure 9.1.3 Two-dimensional axes for plotting equations

(e) What is the formula for the distance from (x_1, y_1) to (x_2, y_2)? Use this formula to find how far $(-1, 2)$ is from $(7, -4)$.

(f) 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. Explain how the right triangle you drew here relates to the calculation in (e).

(g) For each graph in Figure **??**, state whether the graph can be expressed with y as a function of x 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.

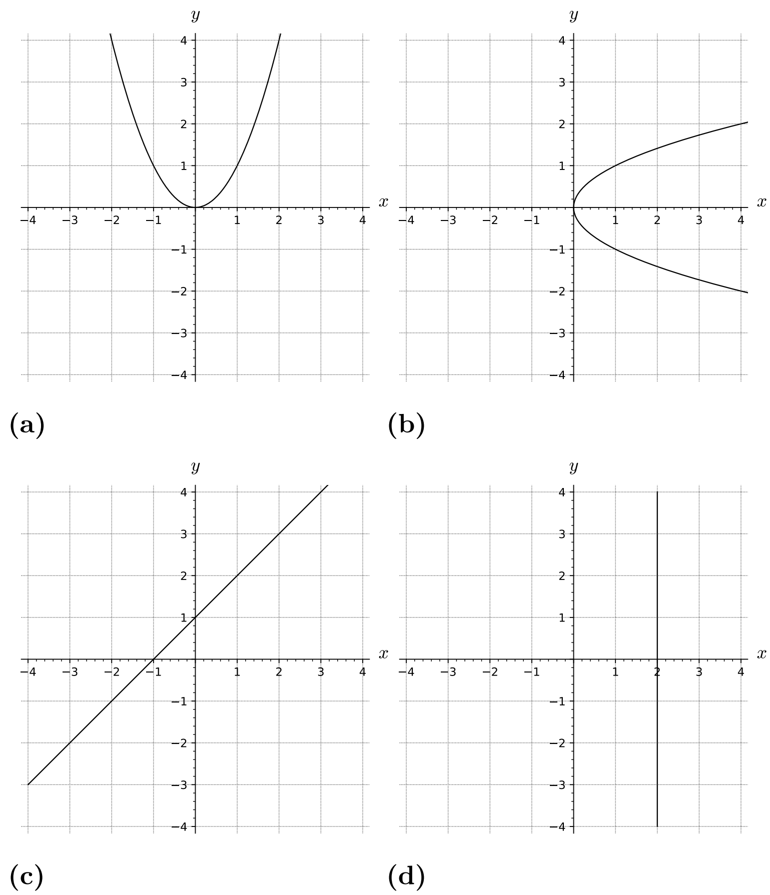
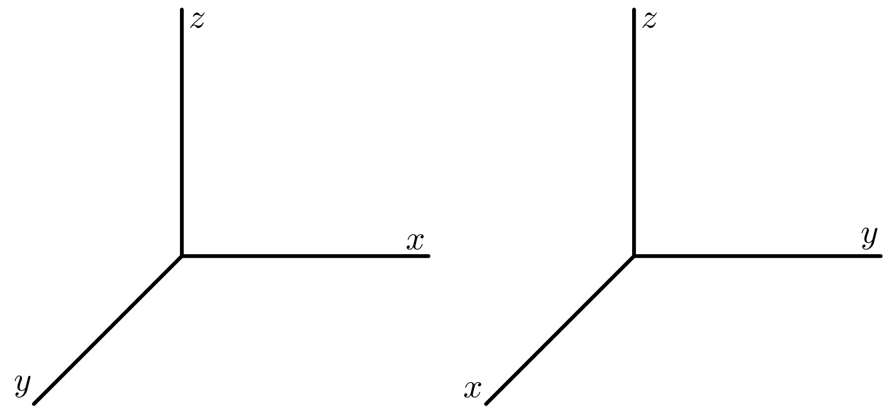


Figure 9.1.4 Four plots to determine if the graphs depict y as a function of x

The tasks in Preview Activity 9.1.1 connect some of the meaning and formulas used in coordinates and graphing for two dimensions to specific measurements. We will use these ideas 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.

9.1.1 Three Dimensional Space and Coordinates

To plot points in three-dimensional space, we need to 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 9.1.8, 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 9.1.5

The distinction between these two figures is subtle, but important. In the coordinate system shown in Figure **??**, 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 **??**, 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 **??** 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.

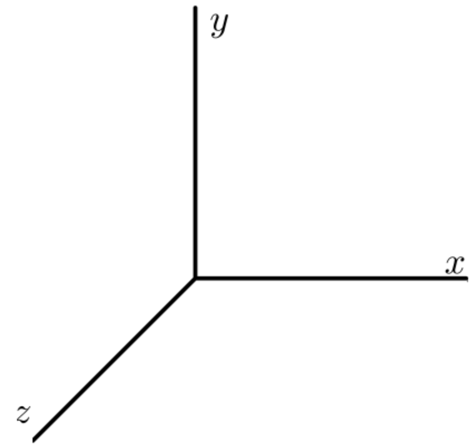


Figure 9.1.6 An alternate orientation of a right handed coordinate system for $\mathbb{R}^3$

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 (minimum) distance to the vertical axis and the vertical coordinate was measured as the (minimum) 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 9.1.10 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.

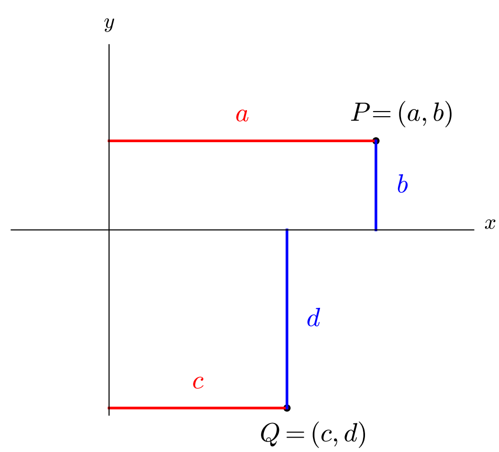
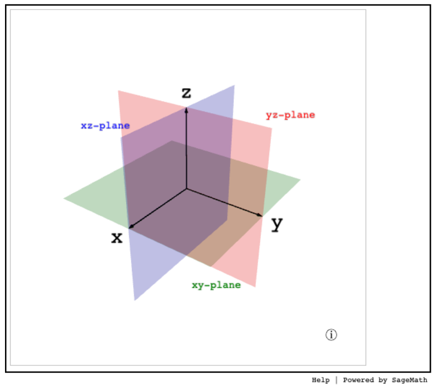


Figure 9.1.7 A plot with the 2D coordinate measurements labeled

In a three dimensional space, we need to pay attention to not just the axes but also the **coordinate planes**coordinate. Figure 9.1.11 shows how the xy-plane is the plane that contains the x- and y-axes. We can define the xz-plane as the plane that contain the x-and z-axes and the yz-plane as the plane that contain the y-and z-axes. 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.



Standalone
Embed

Figure 9.1.8 The coordinate planes in three dimensions

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.

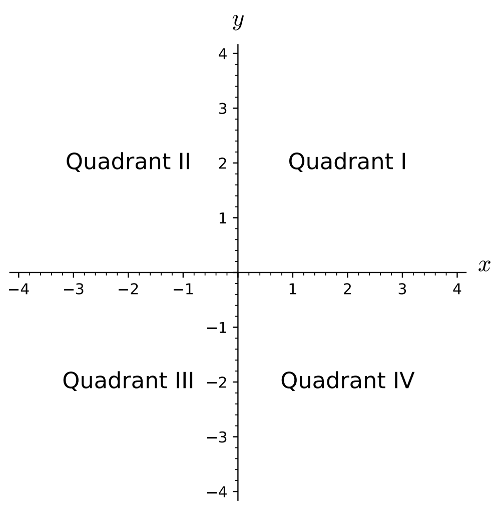
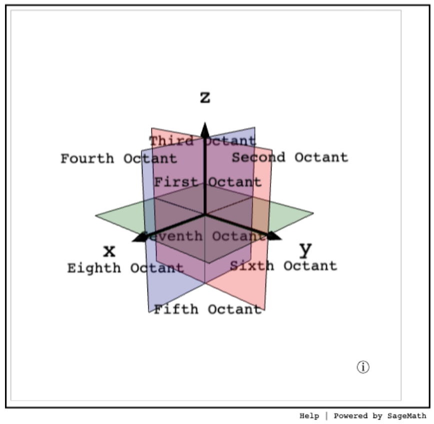


Figure 9.1.9 The four quadrants

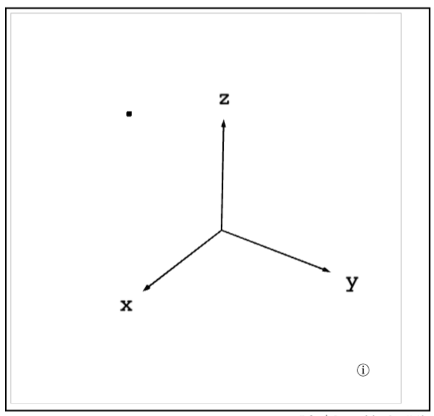
The coordinate planes will divide our three dimensional space up into eight regions, which we will call **octants**. The octants above the xy-plane are ordered just like the quadrants in two dimensions and octants 5 through 8 have the same relationships but below the xy-plane. For example, octant 7 lies below Qudrant III of the xy-plane.



Standalone
Embed

Figure 9.1.10 A plot of the eight 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 9.1.14.

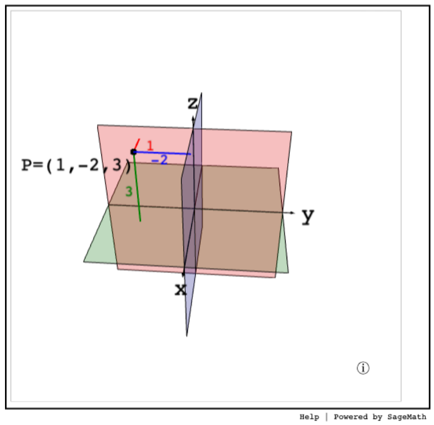


Standalone
Embed

Figure 9.1.11 A plot of the point $(1, -2, 3)$

Notice how hard it is to see where that point is located without turning the axes. In the interactive version of this text, most three-dimensional plots are interactive, which means you click and drag to rotate the perspective. Additionally, you can zoom in and out by either using a scroll wheel on a mouse or pinching on touch screens. Interacting with the three-dimensional plots is an important part of building your understanding.

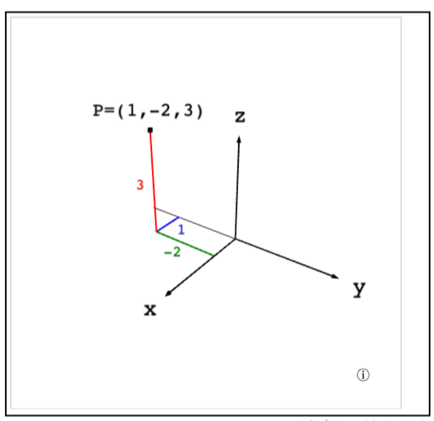
When drawing points (or other plots) in three dimensions, it is often useful to draw segments or other features that are parallel to coordinate axes in order to help the viewer see the orientation and locations. For instance, we hope that you will find it easier to understand the location of the point $(1, -2, 3)$ in Figure 9.1.15 because of the parallel structures shown on the plot.



Standalone
Embed

Figure 9.1.12 A plot of the point $(1, -2, 3)$ with line segments parallel to the coordinate axes

Figure 9.1.15 shows the segments used to determine the coordinates in three dimensions as they are measured from the point to the coordinate planes. Many plots will not show 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. For instance, see how it is easier to understand the location of the point $(1, -2, 3)$ in Figure 9.1.16 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 Figure 9.1.16 are usually reasonable to add to a drawing done by hand as well.



Standalone
Embed

Figure 9.1.13 A plot of the point $(1, -2, 3)$ with line segments parallel to the coordinate axes

Activity 9.1.2 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 9.1.16).

(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 9.1.16).

9.1.2 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. 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 $y \in \mathbb{R}$). In the next activity we consider the three-dimensional analogs.

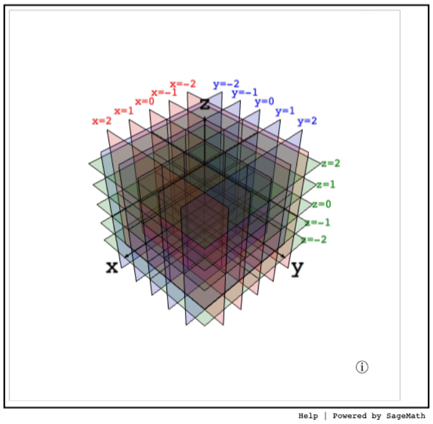
Activity 9.1.3

(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 9.1.3 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 themselvescoordinate. The xy-plane satisfies $z = 0$, the xz-plane satisfies $y = 0$, and the yz-plane satisfies $x = 0$ (see Figure 9.1.11). Planes of the form $x = a$, $y = b$, $z = c$ are called **fundamental planes** are useful in understanding and building structures in three dimensions. In a plot like Figure 9.1.3, you see a grid that helps measure coordinates of points in 2D. Figure 9.1.17 has a lot of visual clutter and is very hard to distinguish fine features, especially something in the middle of the plot. Contrast this to how helpful the grid in Figure **??** is.



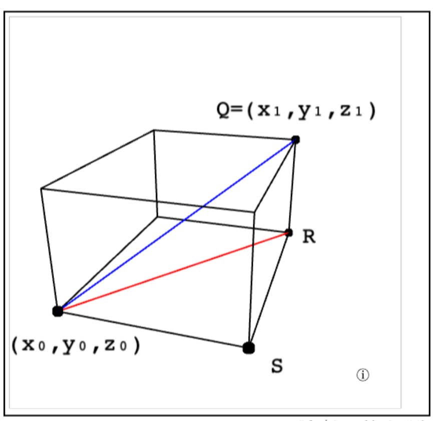
Standalone
Embed

Figure 9.1.14 A grid of fundamental planes in three dimensions

9.1.3 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 9.1.4 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 9.1.18, and the distance between P and Q is the length of the blue diagonal shown in Figure 9.1.18.



Standalone
Embed

Figure 9.1.15 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 9.1.18. 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 **??** 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

$$|PQ| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2} \tag{9.1.1}$$

As Activity **??** 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 (9.1.1) 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 point (x_0, y_0, z_0) is R, the point (x, y, z) will 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

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

9.1.4 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.

(a) Describe the set of points that satisfy $x = 2$ 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 satisfy $x = 2$ 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 satisfy $x = 2$ 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 $x = 2$.

6. 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 $(7, 9, -1)$.

9.1.5 Notes to the Instructor

This section only expects students to have familiarity with measuring distance in two dimensions.

9.2 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?

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 and direction. These quantities are represented by *vectors* and are the study 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 us in describing the behavior of functions of several variables.

Preview Activity 9.2.1 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), then 3) draw lines from $(1, 1)$ to $(1, -1)$ and 4) $(1, -1)$ 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 $\langle x, y\rangle$ (this pair $\langle x, y\rangle$ is a **vector**), where x is the horizontal displacement and y the vertical displacement from one point to another. So, for example, the vector from the origin to the point $(1, 1)$ is represented by $\langle 1, 1\rangle$ but 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.

9.2.1 Representations of Vectors

Preview Activity 9.2.1 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, and we will call such quantity a *vector*.

Definition 9.2.1 A **vector** is a quantity that possesses the attributes of magnitude and direction. $\Diamond$

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 9.2.2.

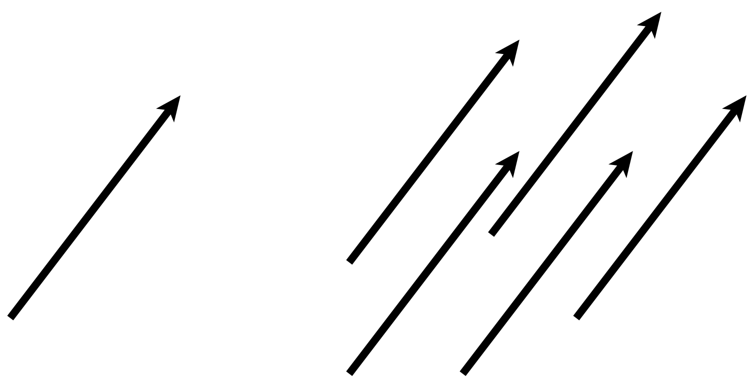


Figure 9.2.2 Left: A vector. Right: Representations of the same vector.

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 at right in Figure 9.2.2. 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, but 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 at left in Figure 9.2.3. 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".

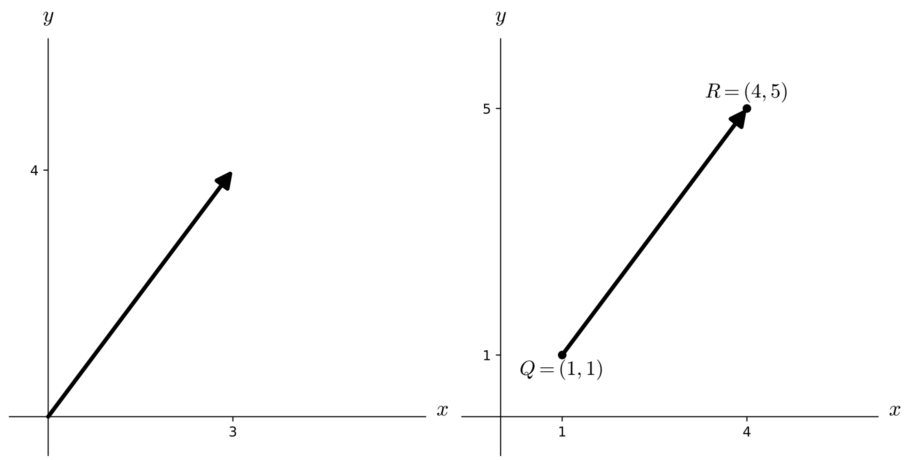


Figure 9.2.3 Left: The vector $\langle 3, 4\rangle$ is standard position. Right: The vector $\langle 3, 4\rangle$ represented from $(1, 1)$ to $(4, 5)$.

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 on right in Figure 9.2.3).

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 9.2.4. 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.

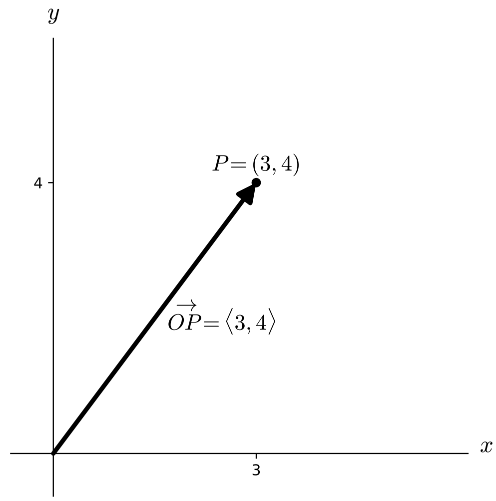


Figure 9.2.4 A point defines a vector in standard position

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 9.2.2 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 $B = (0, 1, 9)$, and your rival Carlos is standing at point $C = (-2, 20, -5)$, all distances are measured in feet.

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

9.2.2 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.

9.2.3 Operations on Vectors

Vectors are not numbers (scalars), but we can now represent them 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 9.2.3 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 1 km west and 4 km north of your current location.

Let $\vec{u}$ be the vector that represents your hike from your car to the expected location of your friend, $\vec{v}$ be the vector that represents the vector from your current location to the ranger's suggested location, and $\vec{w}$ be the vector from your car to the ranger's suggested location. Compute the components of $\vec{u}$, $\vec{v}$, and $\vec{w}$.

(b) Draw a picture that represents $\vec{u}$, $\vec{v}$, and $\vec{w}$ with the reference points C being the location of your car, P being your current position, and S being the ranger's suggested location.

(c) In the context of this problem, explain why you can add the horizontal components of $\vec{u}$ and $\vec{v}$ to get the horizontal component of $\vec{w}$. 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 and says the first ranger made a mistake in their navigation. A drone spotted your friend along the same direction from your car to your current location but your friend is three times as far away you are from your car. 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 $\overrightarrow{CD}$ and how does that compare to $\vec{w}$?

(e) After hiking to the drone's suggested location, you find your friend. Yay! On the walk back to your car with your friend, you decide to help make sure the first ranger understands vectors; in particular, you want to let them know what direction they should have told you to go from P. Use the components of $\overrightarrow{CD}$ and $\vec{u}$ to figure out what the first ranger should have told you.

In Activity 9.2.3 we saw how adding two vectors componentwise gave a vector that represented the total change of traveling along one vector then doing the change described by the other vector. Additionally, we saw that if we wanted to go three times as far in a direction given by a vector, then we could describe the total change by multiplying each component by three. Finally, we saw that we could find the difference between two vectors by subtracting componentwise. This difference vector completed a triangle given by starting the two vectors at the same initial point. The ideas of adding vectors or multiplying a scalar by a vector generalizes beyond the context of Activity 9.2.3.

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 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 9.2.3.e.)

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

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

Here is the cursed i version I made $\hat{i}$

The vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are called the *standard unit vectors* (as we will learn momentarily, unit vectors have length 1), and are commonly used in the physical sciences.

9.2.4 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 $\mathbb{R}^2$; it is straightforward to verify that the rest of the eight properties just noted hold for all vectors in $\mathbb{R}^n$ since all of the operations are done componentwise on the vectors.

Activity 9.2.4 In this activity, you should use the properties of vectors from above 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 vector that fills in the blank: $(2\vec{w} - 3\vec{u}) -$ ______________ $= \vec{0}$

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

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

9.2.5 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 on the left in Figure 9.2.5.

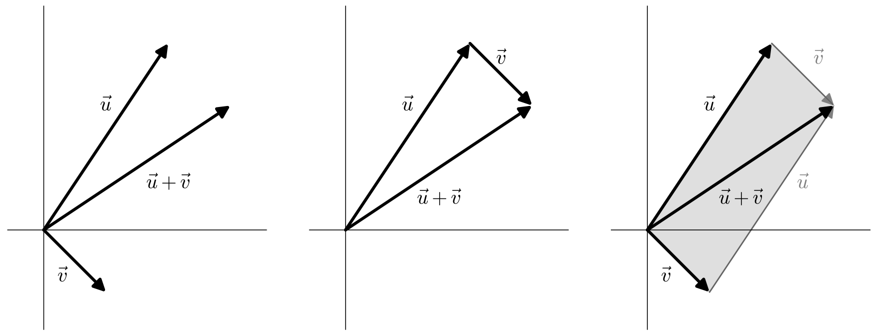


Figure 9.2.5 A vector sum (left), summing displacements (center), the parallelogram law (right).

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 the center of Figure 9.2.5. 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 so we obtain the same sum if we place the tail of $\vec{u}$ at the tip of $\vec{v}$. We therefore see that $\vec{u}+\vec{v}$ appears as the diagonal of the parallelogram determined by $\vec{u}$ and $\vec{v}$, as shown at right in Figure 9.2.5.

Vector subtraction has a similar interpretation. At left in Figure 9.2.6 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 9.2.6. 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 should make sense with our other interpretation of vector subtraction as well; the vector $\vec{w} - \vec{u}$ is what you need to add to $\vec{u}$ to get $\vec{w}$, which is stated algebraically as $\vec{u} + \underline{\vec{w} - \vec{u}} = \vec{w}$.

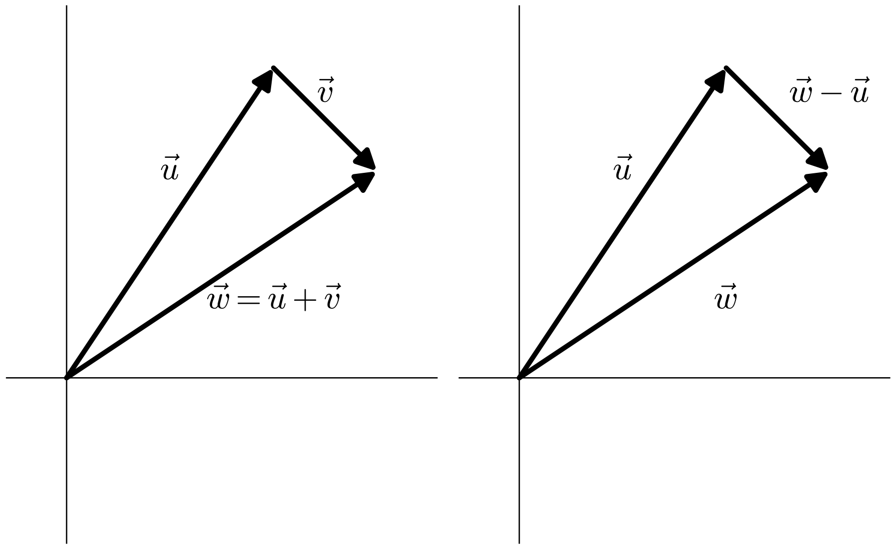


Figure 9.2.6 Left: Vector addition. Right: Vector subtraction.

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 9.2.7, 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 $\vec{v}$.

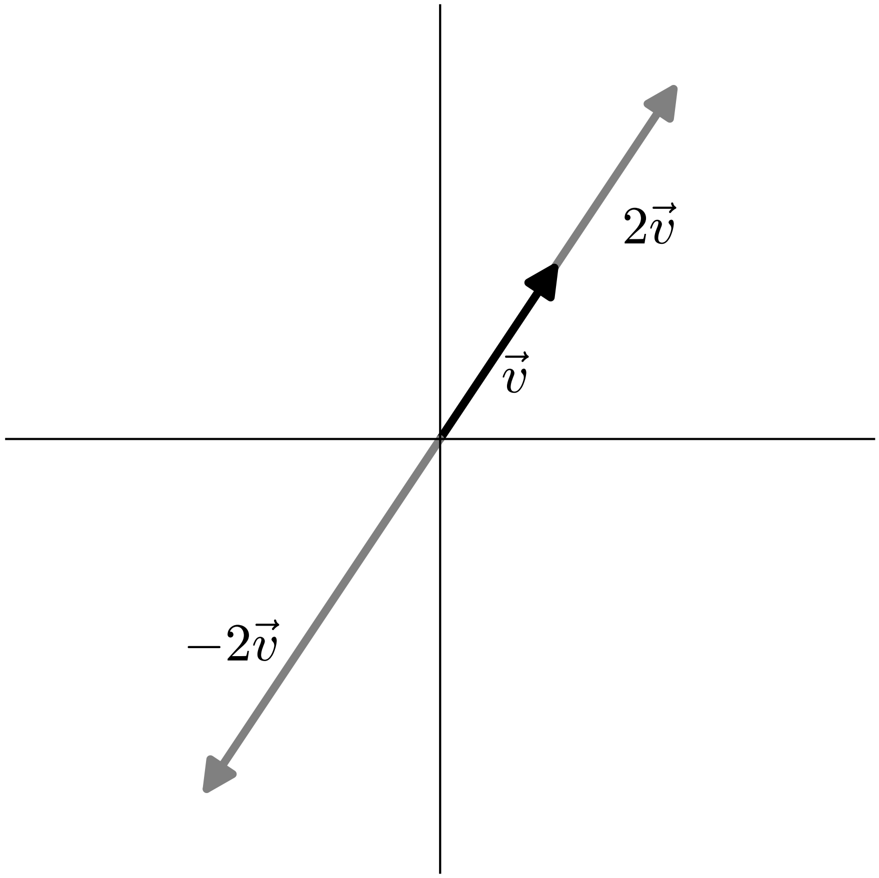


Figure 9.2.7 Scalar multiplication

Figure 9.2.7 motivates the following definition of parallel vectors; Vectors $\vec{v}$ and $\vec{w}$ (in $\mathbb{R}^n$ are parallel if 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 9.2.7. Parallel vectors are extremely useful when trying to classify the orientation of lines, planes, and changes on surfaces.

Activity 9.2.5 Suppose that $\vec{u}$ and $\vec{v}$ are the vectors shown in Figure 9.2.8.

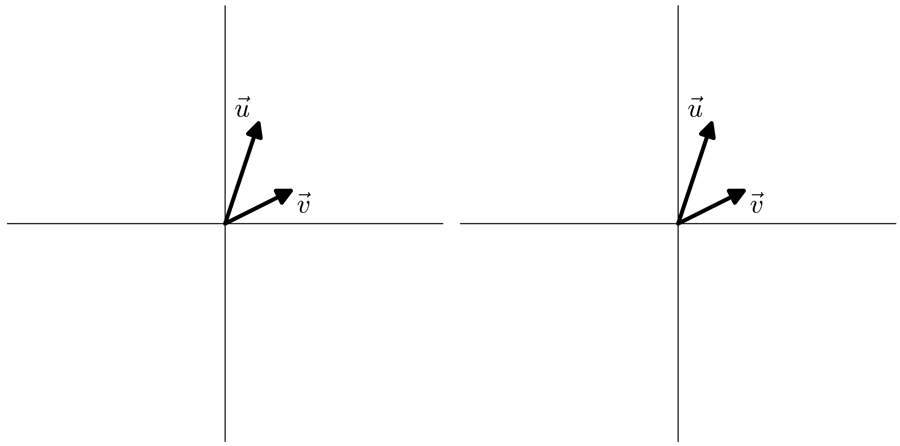


Figure 9.2.8 Left: Sketch sums. Right: Sketch multiples.

(a) On the axes at left in Figure 9.2.8, 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 at right in Figure 9.2.8, 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 at right in Figure 9.2.8, 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 $\hat{k}$. Important questions in the next few math courses include things like "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.

9.2.6 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 9.2.9 The magnitude of a vector is measured as a the length of the directed line segments that represent the vector. For example, the directed line segment $\overrightarrow{AB}$ in Figure 9.2.10 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.

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

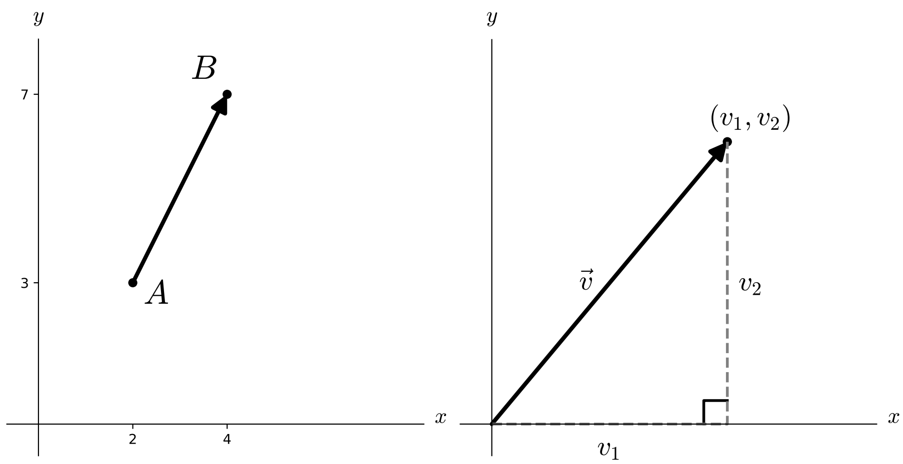


Figure 9.2.10 Left: $\overrightarrow{AB}$. Right: An arbitrary vector, $\vec{v}$.

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 $\vec{v} = \langle v_1,$

Activity 9.2.6

(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 $\|t\vec{w}\|$ 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 $\vec{u} = \langle -2, 3\rangle$.

9.2.7 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

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

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

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

- The magnitude of the vector $\vec{v} = \langle v_1, v_2, \dots, v_n \rangle$ 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 $\vec{v}$.

9.2.8 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) $(-6\,\tilde{i} - 4\,\tilde{j} - 3\,\tilde{k}) - (-8\,\tilde{i} + 5\,\tilde{j} - 3\,\tilde{k}) =$ ____________________

(b) $(-3\,\tilde{i} + 9\,\tilde{j} - 6\,\tilde{k}) - 2(-7\,\tilde{i} + 2\,\tilde{j} - 10\,\tilde{k}) =$ ____________________

2. Find a vector $\mathbf{a}$ that has the same direction as $\langle -6, 9, 6 \rangle$ but has length 5.

Answer: $\mathbf{a} =$ ____________________

3. Let $\mathbf{a} =< -1, 1, -3 >$ and $\mathbf{b} =< 1, 3, -2 >$.

Show that there are scalars s and t so that $s\mathbf{a} + t\mathbf{b} =< 6, 10, -2 >$.

You might want to sketch the vectors to get some intuition.

$s =$ ____________________

$t =$ ____________________

4. Resolve the following vectors into components:

(a) The vector $\vec{v}$ in 2-space of length 5 pointing up at an angle of $3\pi/4$ measured from the positive x-axis.

$\vec{v} =$ __________ $\vec{i} +$ __________ $\vec{j}$

(b) The vector $\vec{w}$ in 3-space of length 3 lying in the xz-plane pointing upward at an angle of $\pi/4$ measured from the positive x-axis.

$\vec{v} =$ __________ $\vec{i} +$ __________ $\vec{j} +$ __________ $\vec{k}$

5. Find all vectors $\vec{v}$ in 2 dimensions having $||\vec{v}|| = 17$ where the $\tilde{j}$-component of $\vec{v}$ is $8\,\tilde{j}$.

vectors: ______________________________

(If you find more than one vector, enter them in a comma-separated list.)

6. Which is traveling faster, a car whose velocity vector is $28\vec{i} + 33\vec{j}$, or a car whose velocity vector is $40\vec{i}$, 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 = ____________________

7. Let $\mathbf{a} = \langle 4, -3, 4 \rangle$ and $\mathbf{b} = \langle 0, 2, -2 \rangle$.

Compute:

$\mathbf{a} + \mathbf{b} =$ (__________,__________,__________)

$\mathbf{a} - \mathbf{b} =$ (__________,__________,__________)

$2\mathbf{a} = (\_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_)$
$3\mathbf{a} + 4\mathbf{b} = (\_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_, \_\_\_\_\_\_\_\_\_\_)$
$|\mathbf{a}| = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$

8. Find the length of the vectors
(a) $5\,\tilde{i} + 4\,\tilde{j} - 3\,\tilde{k}$: length = ________________________
(b) $-2\,\tilde{i} + \tilde{j} + 2.2\,\tilde{k}$: length = ________________________

9. For each of the following, perform the indicated operations on the vectors $\vec{a} = \tilde{j} + 3\,\tilde{k}$, $\vec{b} = 2\,\tilde{i} + 5\,\tilde{j} + 2\,\tilde{k}$, $\vec{z} = \tilde{i} + 4\,\tilde{j}$.
(a) $3\vec{a} + 2\vec{b} =$ ______________________________
(b) $4\vec{a} + 4\vec{b} - 5\vec{z} =$ ______________________________

10. Find the value(s) of a making $\vec{v} = 3a\,\vec{i} - 2\,\vec{j}$ 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.)

11. ***(a)*** Find a unit vector from the point $P = (1, 3)$ and toward the point $Q = (13, 8)$.
$\vec{u} =$ ______________________________
(b) Find a vector of length 26 pointing in the same direction.
$\vec{v} =$ ______________________________

12. A truck is traveling due north at 50 km/hr approaching a crossroad. On a perpendicular road a police car is traveling west toward the intersection at 55 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 $\vec{d} =$ ______________________________

13. 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.

14. 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 $\vec{u} - \vec{v}$.

15. 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 $\vec{v}$.

16.

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.

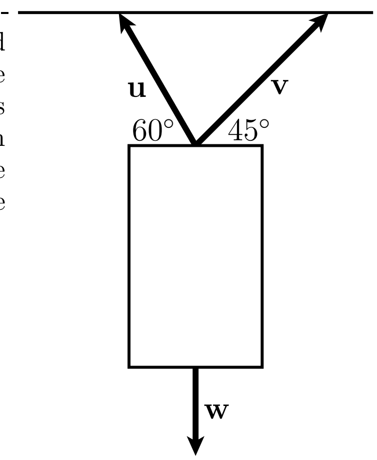


Figure 9.2.11 Forces acting on an object.

As an example, suppose a painting weighing 50 pounds is to be hung from wires attached to the frame as illustrated in Figure 9.2.11. 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 $\vec{v}$.

9.3 The Dot Product

Motivating Questions

- How is the dot product of two vectors defined and what geometric information does it tell us?
- 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?

In Section 9.2, we considered how to do vector addition and scalar multiplication algebraically and what the corresponding geometric meaning is. In this section, we will introduce a type of vector multiplication that will be used to measure how much two vectors move together. First we will discuss what is meant by "move together" in the context of vectors in a one-dimensional space (number line),then we will generalize this idea to two dimensions in Preview Activity 9.3.1.

We will now consider one-dimensional vectors which are denoted $\vec{v} = \langle v_1 \rangle$. Notice that vectors in one dimension will only have one component. Graphically, these vectors will look like an arrow on the number line corresponding to the change described by the component. Remember that vector do not have an inherent starting or ending location, so we can translate vectors (without changing the magnitude or direction).

Figure 9.3.1 A plot of $\langle 2 \rangle$ and $\langle -1.5 \rangle$ in red and blue, respectively

Given two vectors, we want measure how much these vectors move together with a scalar measurement. For instance, if we let $\vec{v} = \langle 2 \rangle$ and $\vec{w} = \langle -3 \rangle$, then we can graphically represent these starting from the origin as in Figure 9.3.2. We want a multiplicative, scalar measure of how much $\vec{v}$ and $\vec{w}$ move together, so let's multiply their (one) components. By multiplying our components, we get $(2)(-3) = -6$. You probably notice the negative in this measurement first, so let's make sense of this right away.

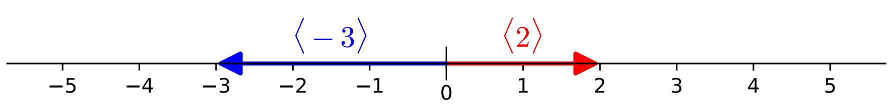


Figure 9.3.2 A plot of $\langle 2 \rangle$ and $\langle -3 \rangle$ in red and blue, respectively

The red and blue vectors in Figure 9.3.2 move completely opposite of each other, which is why we got a negative result. Generalizing this idea for all one-dimensional vectors, we see that taking the product of the components of the vectors will give a positive result when the vectors are in the same direction and a negative result when the vectors move opposite of each other. For example, the pair of vectors $\langle 2 \rangle$ and $\langle 3 \rangle$ will move together exactly as much as $\langle -2 \rangle$ and $\langle -3 \rangle$ will together. Note that the product of components for each of these pairs is 6.

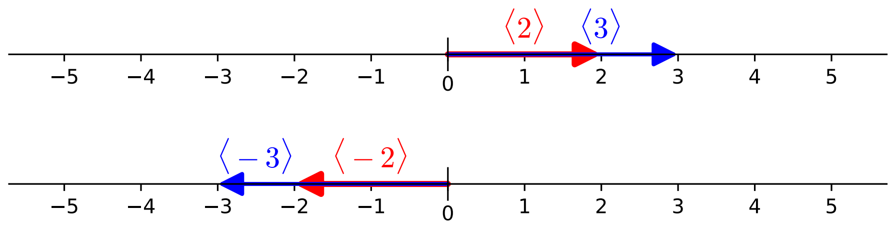


Figure 9.3.3 Plots of $\langle 2\rangle$ and $\langle 3\rangle$ together and $\langle -2\rangle$ and $\langle -3\rangle$ together

Changing the length of the vectors in the pair you are comparing will also affect how much the vectors move together as measured by the product of the components (in one-dimension). For example, the pair of vector $\langle 2\rangle$ and $\langle 1\rangle$ do not move together as much as the pair of vectors $\langle 2\rangle$ and $\langle 3\rangle$ do. In fact, the product of the components says that the pair of vector $\langle 2\rangle$ and $\langle 1\rangle$ move only a third as much together as $\langle 2\rangle$ and $\langle 3\rangle$ do. This should make sense because there are not as long of arrows in the first figure as the second in Figure 9.3.4.

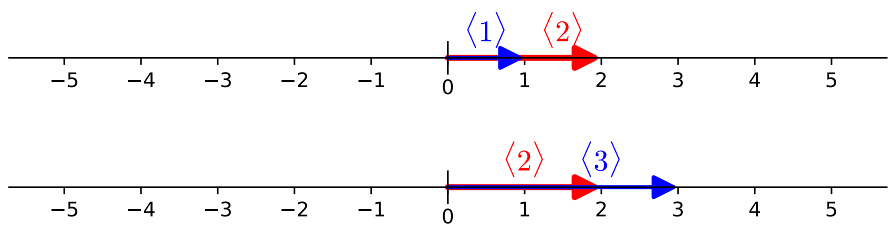


Figure 9.3.4 Plots of $\langle 1\rangle$ and $\langle 2\rangle$ together and $\langle 2\rangle$ and $\langle 3\rangle$ together

Our one dimensional examples above motivate why we will use the product of components to measure how much vectors move together, but we pause now to say something about what "move together" does ***NOT*** mean. Specifically, when we ask how much $\vec{v}$ and $\vec{w}$ move together, we are NOT measuring how much of $\vec{v}$ is in the direction of $\vec{w}$ or how much of $\vec{w}$ is in the direction of $\vec{v}$. Additionally, we are NOT measuring the difference between $\vec{v}$ and $\vec{w}$. We need a scalar result (and not a vector) to measure how much $\vec{v}$ and $\vec{w}$ move together because this measurement of together does not have an associated direction. In the following Preview Activity, we will use our measurement of moving together in one-dimension, as computed by the product of the vector components, to develop a meaning for what "moving together" means in a two-dimensional setting.

Preview Activity 9.3.1

(a) Let's look at some examples to understand what we mean by "moving together" for 2D vectors. In particular, we will use the following:

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

(a) Draw representatives of each of $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ on the plane.

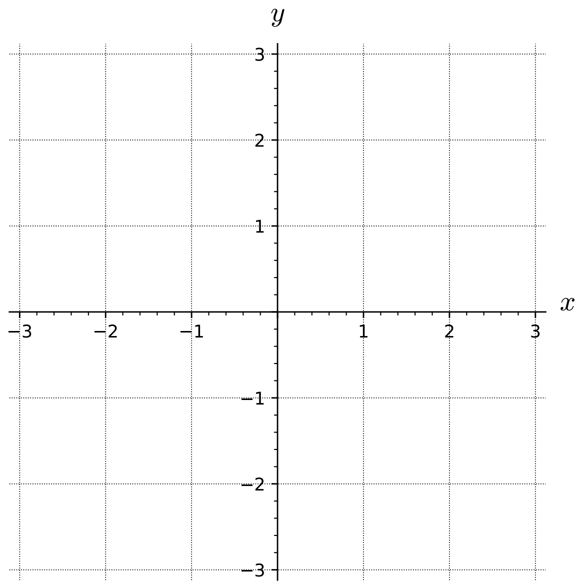


Figure 9.3.5 A plot of the 2D plane

(b) Multiply the first components of $\vec{u}_1$ and $\vec{u}_2$ and use your ideas from one-dimensional vectors to explain what this value says about how much $\vec{u}_1$ and $\vec{u}_2$ move together horizontally.

(c) Multiply the second components of $\vec{u}_1$ and $\vec{u}_2$ and use your ideas from one-dimensional vectors to explain what this value says about how much $\vec{u}_1$ and $\vec{u}_2$ move together vertically.

(d) Add together the results of the previous two parts and explain how this sum relates to how much $\vec{u}_1$ and $\vec{u}_2$ move together. You should also look at a plot of $\vec{u}_1$ and $\vec{u}_2$ starting at the same point.

(b) (a) Multiply the horizontal and vertical components of $\vec{u}_1$ and $\vec{u}_3$ separately and use these results to write about how much $\vec{u}_1$ and $\vec{u}_3$ together horizontally and how much $\vec{u}_1$ and $\vec{u}_3$ move together vertically.

(b) Add together the two results of the previous task and explain how this sum relates to how much $\vec{u}_1$ and $\vec{u}_3$ move together. You should also look at a plot of $\vec{u}_1$ and $\vec{u}_3$ starting at the same point.

(c) Multiply the horizontal and vertical components of $\vec{u}_1$ and $\vec{u}_4$ separately then add the two results. Use this number and a plot of $\vec{u}_1$ and $\vec{u}_3$ starting at the same point to explain how much $\vec{u}_1$ and $\vec{u}_4$ move together.

9.3.1 The Dot Product

We can extend the ideas of Preview Activity 9.3.1 to any dimension which will give us the definition of the dot product.

Definition 9.3.6 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, but it is more convenient from a mathematical perspective to define the dot product algebraically and then view work as an application of this definition.

For instance, we find that

$$\langle 3, 0, 1\rangle \cdot \langle -2, 1, 4\rangle = 3 \cdot (-2) + 0 \cdot 1 + 1 \cdot 4 = -6 + 0 + 4 = -2.$$

Notice that the resulting quantity is a scalar. Our work in Preview Activity 9.3.1 examined dot products of one- and two-dimensional vectors.

Activity 9.3.2 Determine each of the following.

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

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

The dot product is a natural way to define a product of two vectors to measure how much the vectors move together. Additionally, the dot product behaves in ways that are similar to the product 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 addition to the algebraic properties listed above, the dot product gives us valuable geometric information about the vectors and their relative orientation. For instance, let's consider what happens when we dot a vector with itself:

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

In other words, the dot product of a vector with itself gives the square of the length of the vector: $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$.

9.3.2 The Angle between Vectors

As you saw in Preview Activity 9.3.1, the dot product measures how much vectors move together and can help us understand the angle between two vectors. If we are given two vectors $\vec{u}$ and $\vec{v}$, there are two angles that these vectors create, as depicted at left in Figure 9.3.7. 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.

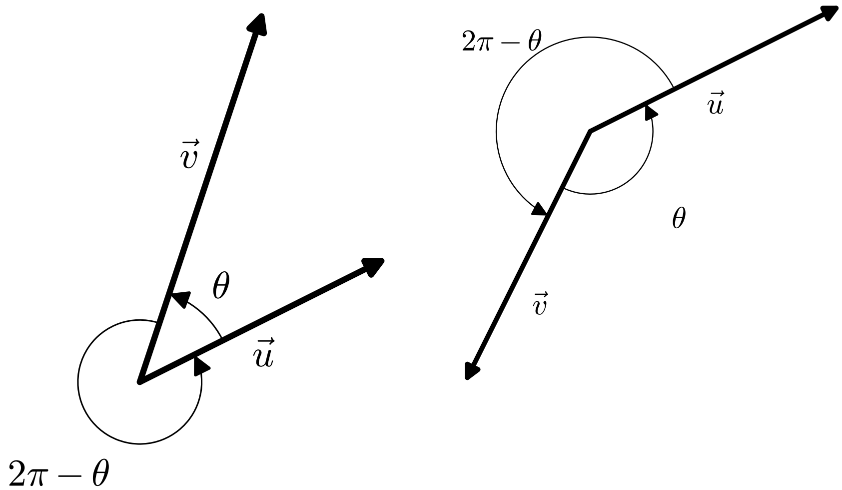


Figure 9.3.7 Left: The angle between vectors for two different configurations of $\vec{u}$ and $\vec{v}$

Key Idea 9.3.8 *Given vectors $\vec{u}$ and $\vec{v}$, the following relationship holds for θ being the smallest angle between the vectors:*

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \|\vec{u}\| \|\vec{v}\| \cos(\theta) \tag{9.3.1}$$

Solving for the angle θ, we get the following:

$$\theta = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|}\right) \tag{9.3.2}$$

Proof. To determine this angle, we may apply the Law of Cosines to the triangle shown at right in Figure 9.3.9.

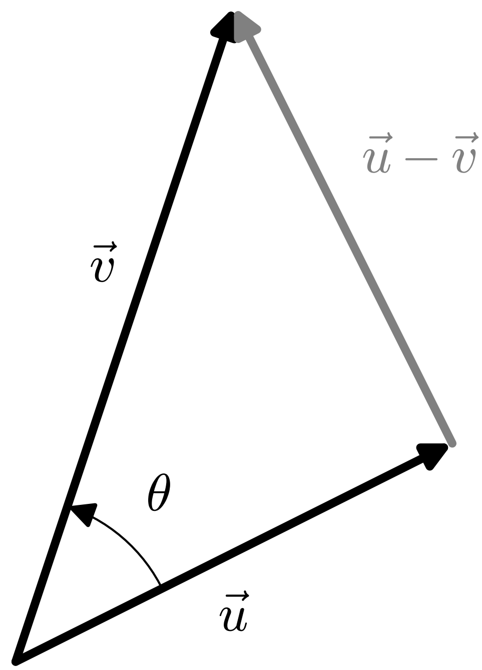


Figure 9.3.9 A triangle of vectors, $\vec{u}$, $\vec{v}$, and $\vec{u}-\vec{v}$

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 real beauty of (9.3.1) is this: the dot product is a very simple algebraic calculation to perform yet it provides us with important geometric information about the angle between the vectors that would be difficult to determine otherwise. This should not be that surprising because we defined the dot product to measure how much vectors moved together, thus the dot product would be related to the angle between the vector inputs. While you normally need to interpret the angle that comes out of an inverse trig function, the result of (9.3.2) does not need to be interpreted because the angle measurement and the range of arccos are both from 0 to π.

Activity 9.3.3 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?

9.3.3 The Dot Product and Orthogonality

When the angle between two vectors is a right angle, it is frequently the case that something important is happening. 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.

As Activity 9.3.3 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 if $\vec{u} \cdot \vec{v} = 0$.

More generally, the sign of the dot product gives us useful information about the relative orientation of the vectors. If we remember that

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

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

$$\begin{aligned} \vec{u} \cdot \vec{v} > 0 &\quad \text{if } \theta \text{ is an acute angle,} \\ \vec{u} \cdot \vec{v} = 0 &\quad \text{if } \theta \text{ is a right angle, and} \\ \vec{u} \cdot \vec{v} < 0 &\quad \text{if } \theta \text{ is an obtuse angle.} \end{aligned}$$

This is illustrated in Figure 9.3.10.

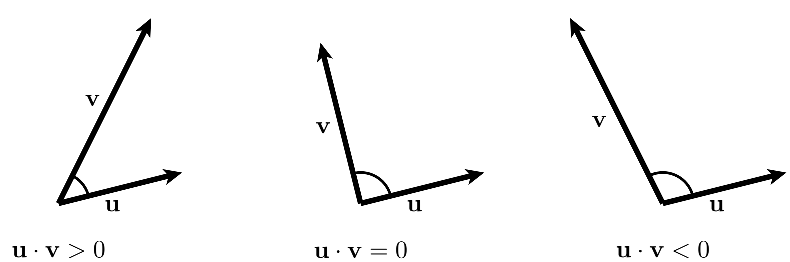


Figure 9.3.10 The orientation of vectors

Earlier in this section we referred to the dot product as measuring how much two vectors move together. We should emphasize that this is a different idea than asking how much of one vector is in the direction of another, which would have a vector as an answer. The relationship between the sign of the dot product and the angle between the vectors reinforces the idea of how much two vectors move together (or apart). We would say two vectors move together a "positive" amount if they have an acute angle between them, meaning that the net amount of how much the components of the vectors move together/apart is positive. Vectors that do not move together at all would have a right angle between them and thus the dot product would be zero. Vectors with a negative dot product will move opposite each other more than together (as a sum of the components moving together/apart), thus they will have an obtuse angle between them.

9.3.4 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 9.3.11, 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 $\overrightarrow{AB}$.

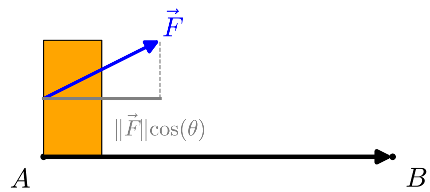


Figure 9.3.11 A diagram showing the force applied to the refrigerator as the refrigerator is moved from A to B

The work required to move the refrigerator the object is

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

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. Consequently, if we are given two vectors $\vec{u}$ and $\vec{v}$, we would like to write $\vec{u}$ as a sum of two vectors, one of which is parallel to $\vec{v}$ and one of which is orthogonal to $\vec{v}$. We will talk about this after the following activity.

Activity 9.3.4 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. In addition, is more work or less work done if the angle to the direction of the object's motion is 60°?

9.3.5 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 the possibilities shown in Figure 9.3.12.

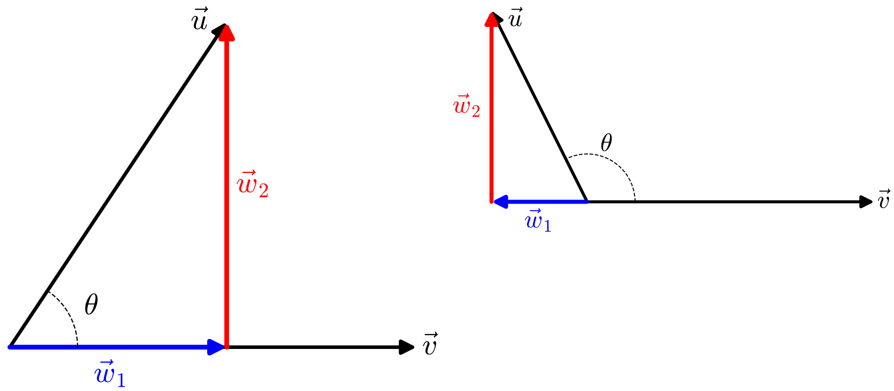
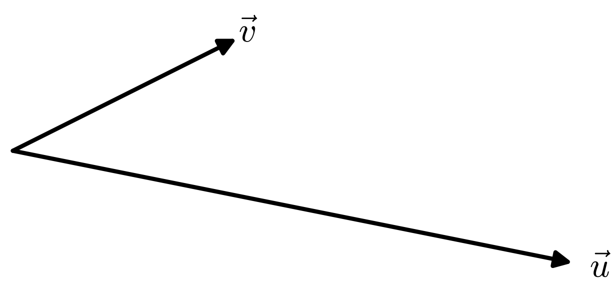


Figure 9.3.12 Different possibilities for spliting $\vec{u}$ into parts that are parallel and orthogonal to $\vec{v}$

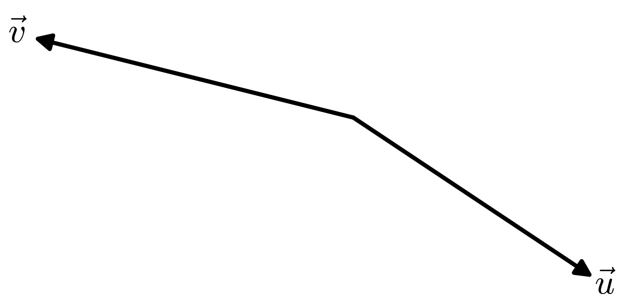
Notice that when the angle between $\vec{u}$ and $\vec{v}$ is obtuse, then the amount of $\vec{u}$ that is parallel to $\vec{v}$ will be in a direction opposite $\vec{v}$. Before we look at how to calculate the vectors $\vec{w}_1$ and $\vec{w}_2$, we will do an activity about drawing the graphical representation of this splitting of $\vec{u}$ into parts that are parallel to $\vec{v}$ and orthogonal to $\vec{v}$

Activity 9.3.5 In this activity, we will be focused on drawing the graphical representation of this splitting of $\vec{u}$ into parts that are parallel to $\vec{v}$ (which we will call $\vec{w}_1$) and orthogonal to $\vec{v}$ (which we will call $\vec{w}_2$).



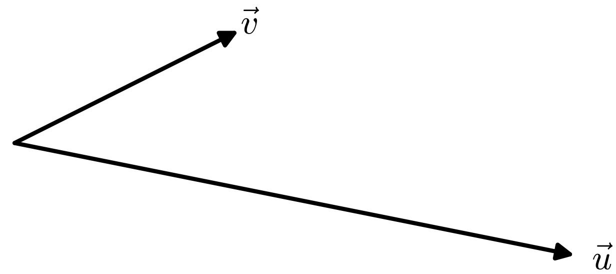
(a) **Figure 9.3.13** A plot of $\vec{u}$ and $\vec{v}$

For the configuration of vectors shown above, draw and label the following vectors: $\vec{w}_1$ and $\vec{w}_2$



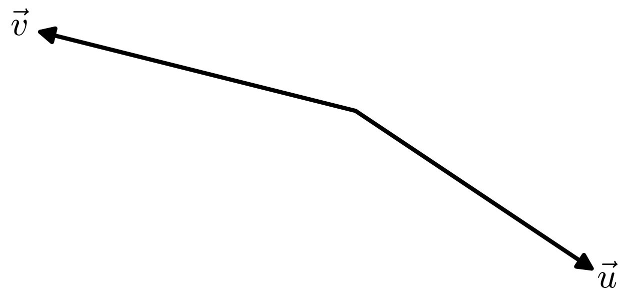
(b) **Figure 9.3.14** A plot of $\vec{u}$ and $\vec{v}$

For the configuration of vectors shown above, draw and label the following vectors: $\vec{w}_1$ and $\vec{w}_2$



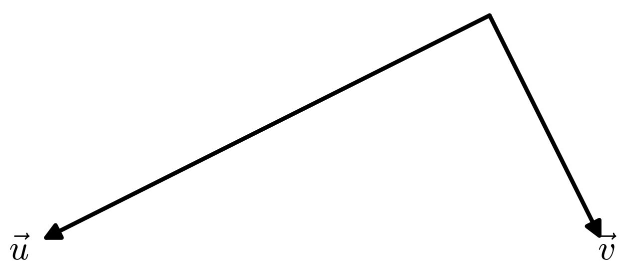
(c) **Figure 9.3.15** A plot of $\vec{u}$ and $\vec{v}$

We want to switch the roles of $\vec{u}$ and $\vec{v}$ for the examples above. Specifically, for the configuration of vectors shown above, 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 the figure above, draw $\vec{z}_1$ and $\vec{z}_2$.



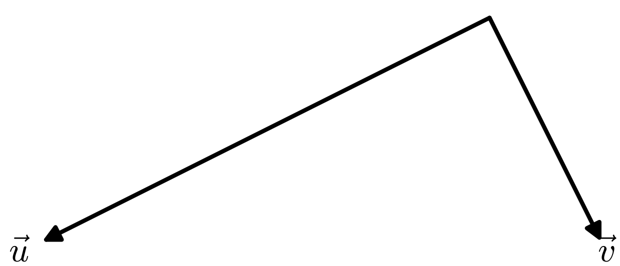
(d) **Figure 9.3.16** A plot of $\vec{u}$ and $\vec{v}$

For the configuration of vectors shown above, draw and label the following vectors: $\vec{z}_1$ and $\vec{z}_2$



(e) **Figure 9.3.17** A plot of $\vec{u}$ and $\vec{v}$

For the configuration of vectors shown above, draw and label the following vectors: $\vec{w}_1$ and $\vec{w}_2$



(f) **Figure 9.3.18** A plot of $\vec{u}$ and $\vec{v}$

For the configuration of vectors shown above, draw and label the following vectors: $\vec{z}_1$ and $\vec{z}_2$

We want to compute $\vec{w}_1$ and $\vec{w}_2$ in terms of $\vec{u}$ and $\vec{v}$. Let's take our sum $\vec{u} = \vec{w}_1 + \vec{w}_2$ and take the dot product of each side with $\vec{v}$

$$\begin{aligned}\vec{u}\cdot\vec{v} =& (\vec{w}_1 + \vec{w}_2)\cdot\vec{v}\\ =& (\vec{w}_1)\cdot\vec{v} + (\vec{w}_2)\cdot\vec{v}\\ =& (\vec{w}_1)\cdot\vec{v}\end{aligned}$$

Notice that because $\vec{w}_2$ must be orthogonal to $\vec{v}$, the second term will contribute nothing to the dot product $\vec{u}\cdot\vec{v}$. Because $\vec{w}_1$ is parallel to $\vec{v}$, then $\vec{w}_1 = k\vec{v}$. Since $\vec{u}\cdot\vec{v} = (k\vec{v})\cdot\vec{v}$, we can solve for $k = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}$. In particular, this means that we can calculate $\vec{w}_1 = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$.

Vector subtraction will also give us a way to solve for the piece of $\vec{u}$ that is orthogonal to $\vec{v}$, namely $\vec{w}_2 = \vec{u} - \vec{w}_1 = \vec{u} - \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$. Using the property that a vector's dot product with itself is the vector's magnitude squared, we can write $\vec{w}_1 = \left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|}\right)\frac{\vec{v}}{\|\vec{v}\|}$.

Definition 9.3.19 Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^n$. The component of $\vec{u}$ in the direction of $\vec{v}$ is the scalar

$$\mathrm{comp}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|},$$

and the projection of $\vec{u}$ onto $\vec{v}$ is the vector

$$\mathrm{proj}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v} = \left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{v}\|}\right)\frac{\vec{v}}{\|\vec{v}\|}$$

The orthogonal projection of $\vec{u}$ onto $\vec{v}$ is

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

◊

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 $\mathrm{proj}_{\vec{v}}\vec{u}$. Specifically, we can write $\mathrm{proj}_{\vec{v}}\vec{u} = (\mathrm{comp}_{\vec{v}}\vec{u})\frac{\vec{v}}{\|\vec{v}\|}$, which separates the projection of $\vec{u}$ onto $\vec{v}$ into the (scalar) how much of $\vec{u}$ is parallel to $\vec{v}$ times the unit vector in the direction of $\vec{v}$. This is a very useful idea that will be used in the analysis of many vector calculations, namely separating the magnitude of a vector and the direction of a vector (as described by a unit vector). The first expression for $\mathrm{proj}_{\vec{v}}\vec{u} = \frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$ has a nice conceptual interpretation in that the projection of $\vec{u}$ onto $\vec{v}$ will need to be parallel to $\vec{v}$ and $\frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}$ will be length of $\mathrm{proj}_{\vec{v}}\vec{u}$ as a proportion of $\vec{v}$.

Example 9.3.20 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 9.3.19, we get the following:

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

Using the other formula for the projection, we get

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

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

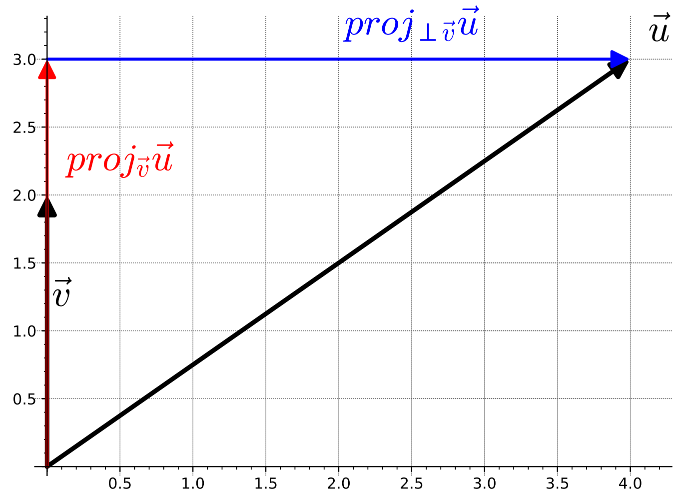


Figure 9.3.21 A plot of $\vec{u}$, $\vec{v}$, $proj_{\vec{v}\vec{u}}$, and $proj_{\perp\vec{v}}\vec{u}$

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

$$proj_{\perp\vec{v}}\vec{u} = \vec{u} - proj_{\vec{v}\vec{u}=\langle 4,3\rangle - \langle 0,3\rangle = \langle 4,0\rangle}$$

All of these calculations should make sense now:

- 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 $\vec{u}$

□

Activity 9.3.6 Let $\vec{u} = \langle 2,6\rangle$.

(a) Let $\vec{v} = \langle 4,-8\rangle$. Find $\text{comp}_{\vec{v}}\vec{u}$, $\text{proj}_{\vec{v}}\vec{u}$ and $\text{proj}_{\perp\vec{v}}\vec{u}$, and draw a picture to illustrate. 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 $\vec{v} = \langle -2,4\rangle$. Without doing any calculations, find $\text{proj}_{\vec{v}}\vec{u}$. Explain your reasoning.

Hint. Refer to the picture you drew in the previous part

(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.

Exercise 9.3.7.12 will show how switching the vectors in the projection formula will give a different result; in general $\text{proj}_{\vec{v}}\vec{w} \neq \text{proj}_{\vec{w}}\vec{v}$.

9.3.6 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 the 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

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

9.3.7 Exercises

1. Find $\mathbf{a} \cdot \mathbf{b}$ if
$\mathbf{a} = \langle 1, -4, -2 \rangle$ and $\mathbf{b} = \langle 3, 1, -2 \rangle$
$\mathbf{a} \cdot \mathbf{b} =$ ____________________
Is the angle between the vectors "acute", "obtuse" or "right"?

2. Determine if the pairs of vectors below are "parallel", "orthogonal", or "neither".
$\mathbf{a} = \langle 2, 3, 3 \rangle$ and $\mathbf{b} = \langle 8, 12, -52/3 \rangle$ are
$\mathbf{a} = \langle 2, 3, 3 \rangle$ and $\mathbf{b} = \langle 8, 12, 12 \rangle$ are
$\mathbf{a} = \langle 2, 3, 3 \rangle$ and $\mathbf{b} = \langle -8, -12, -13 \rangle$ are

3. Perform the following operations on the vectors $\vec{u} = \langle 3, -1, -3 \rangle$, $\vec{v} = \langle -3, 1, -2 \rangle$, and $\vec{w} = \langle 1, 1, 0 \rangle$.
$\vec{u} \cdot \vec{w} =$ ______________________________
$(\vec{u} \cdot \vec{v})\vec{u} =$ ______________________________
$((\vec{w} \cdot \vec{w})\vec{u}) \cdot \vec{u} =$ ______________________________
$\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{w} =$ ______________________________

4. Find $\mathbf{a} \cdot \mathbf{b}$ if $|\mathbf{a}| = 9$, $|\mathbf{b}| = 1$, and the angle between $\mathbf{a}$ and $\mathbf{b}$ is $\frac{\pi}{2}$ radians.
$\mathbf{a} \cdot \mathbf{b} =$ ____

5. What is the angle in radians between the vectors
$\mathbf{a} =$ (10, 5, -10) and
$\mathbf{b} =$ (4, 1, -4)?
Angle: ____
(radians)

6. Find $\mathbf{a} \cdot \mathbf{b}$ if $|\mathbf{a}| = 2$, $|\mathbf{b}| = 4$, and the angle between $\mathbf{a}$ and $\mathbf{b}$ is $-\frac{\pi}{1}$ radians.
$\mathbf{a} \cdot \mathbf{b} =$ ____

7. A constant force $\mathbf{F} = -6\mathbf{i} - 9\mathbf{j} + 7\mathbf{k}$ moves an object along a straight line from point $(-6, -3, 8)$ to point $(-8, -13, 18)$.
Find the work done if the distance is measured in meters and the

magnitude of the force is measured in newtons.

Work: ______________________ Nm

8. A woman exerts a horizontal force of 4 pounds on a box as she pushes it up a ramp that is 7 feet long and inclined at an angle of 30 degrees above the horizontal.

Find the work done on the box.

Work: ______________________ ft-lb

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

- ⊙ positive
- ⊙ any real number
- ⊙ zero
- ⊙ negative

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

11. 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.

12. In this exercise, we will show that the roles of $\vec{u}$ and $\vec{v}$ are not symmetric in projection formula (Definition 9.3.19). We will use the same values as Example 9.3.20, 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 9.3.20).

(b) Use Definition 9.3.19 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 9.3.21.

13. 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 $\vec{b}$.

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

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

iii. Find the area of the given triangle.

14. 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

15. 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 $\mathbb{R}^n$.

16. 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.

17.

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.

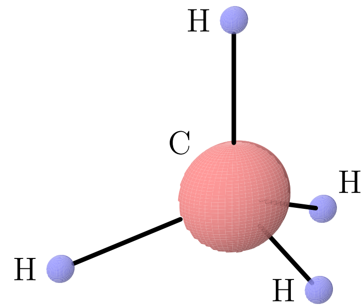


Figure 9.3.22 A methane molecule.

As an example of the molecular geometry of a molecule, consider the methane CH_4 molecule, as illustrated in Figure 9.3.22. 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.

9.4 The Cross Product

Motivating Questions

- How and when is the cross product of two vectors defined?
- What geometric information does the cross product provide?

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-hand coordinate system, as described in Section 9.1. In particular, recall that the vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are oriented as shown below in Figure 9.4.1. We would like to think of a right handed coordinate system in terms of vectors that are not just $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

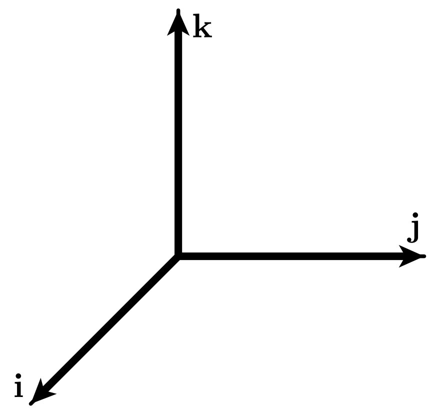


Figure 9.4.1 Basis vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

We will call an ordered list of vectors $\{\vec{u}, \vec{v}, \vec{w}\}$ a **right-handed coordinate system** (or right handed orientation) if when you put the fingers of your right hand in the direction of $\vec{u}$ and curl your fingers in the direction of $\vec{v}$, then the thumb of your right hand will point 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 $\vec{w}$.

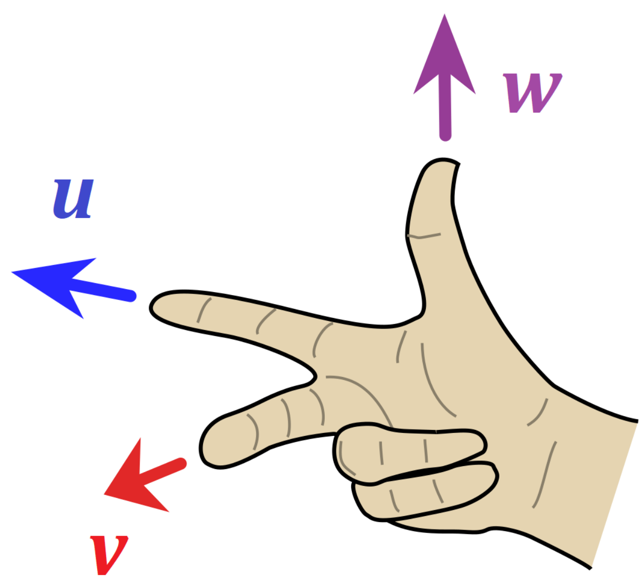


Figure 9.4.2 A right handed orientation of vectors $\{\vec{u}, \vec{v}, \vec{w}\}$

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. So if $\vec{u}$ and $\vec{v}$ are not parallel, we want the cross product of $\vec{u}$ with $\vec{v}$ (denoted $\vec{u} \times \vec{v}$) to make $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ a right-handed coordinate system. Additionally, we would like the cross product (as with all of our vector operations) 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 9.4.1

(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 pay attention to when you have vectors that are in the opposite directions of $\hat{\imath}$ or $\hat{\jmath}$ or $\hat{k}$.

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

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

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

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

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

(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 in the introduction, we would like the cross product to make sense over vector sums $((\vec{v} + \vec{u}) \times \vec{w} = (\vec{v} \times \vec{w}) + (\vec{u} \times \vec{w}))$. Compute each of the following geometrically. You should focus on the direction of the vector needed to complete the right handed coordinate system and not worry about the length of the vector you are using.

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

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

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

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) We also want to make sure that the cross product works well with scalar multiplication.

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

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

(c) 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 $(3\hat{k}) \times \hat{\jmath}$ be?

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

9.4.1 Computing the cross product

As we have seen in Preview Activity 9.4.1, 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 want the cross product to distribute over linear combinations ($(k\vec v + \vec u) \times \vec w = k(\vec v \times \vec w) + (\vec u \times \vec w)$), so 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 9.4.3 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

$$(u_2v_3 - u_3v_2)\hat\imath - (u_1v_3 - u_3v_1)\hat\jmath + (u_1v_2 - u_2v_1)\hat k. \tag{9.4.1}$$

$\lozenge$

At first, this may look intimidating and difficult to remember. However, if we rewrite the expression in Equation (9.4.1) using determinants, important structure emerges. 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 (9.4.1) 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 9.4.4 We will use the definition of cross product to check some of our work from Preview Activity 9.4.1. 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

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

$$= \langle 0, 1, 0 \rangle$$

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.

□

Activity 9.4.2 Suppose $\vec{u} = \langle 0, 1, 3 \rangle$ and $\vec{v} = \langle 2, -1, 0 \rangle$. Use the formula (9.4.1) 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) Multiplication of real numbers is *associative*, which means, for instance, that $(2 \cdot 5) \cdot 3 = 2 \cdot (5 \cdot 3)$. 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 9.4.1.b will apply here.

The cross product satisfies the following properties, some of which were illustrated in Preview Activity 9.4.1 or may be verified from the definition (9.4.1).

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 wanted the cross product to measure. The result of the cross product should complete a right handed coordinate system and the cross product should work well with linear combinations (distribute 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 hand 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. Properties 2 and 3 are exactly what we meant for the cross product to work well with linear combinations and can be verified algebraically from (9.4.1), and Property 5 should make sense in the context of Activity 9.4.2.d.

9.4.2 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 (9.4.1)) satisfies our claim in the introduction to this section.

Key Idea 9.4.5 *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 9.4.1 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 Exercise 9.4.7.16). ■

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

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\|\|\vec{v}\| \sin(\theta) \tag{9.4.2}$$

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 9.3.2. 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 9.4.6 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, which implies that $\sin(\theta)=0\Rightarrow\vec{u}\times\vec{v}=\vec{0}$ by Equation (9.4.2).

9.4.3 Applications of the Cross Product

Equation (9.4.2) is also related to the parallelogram formed by two vectors $\vec{u}$ and $\vec{v}$, as shown in Figure 9.4.7. We say that $\vec{w}$ is the outward direction of the parallelogram formed by $\vec{u}$ and $\vec{v}$ if $\vec{w}$ is perpendicular (in $\mathbb{R}^3$) to the parallelogram and $\{\vec{u},\vec{v},\vec{w}\}$ has a right hand orientation. You should now convince yourself that in Figure 9.4.7 the outward direction is coming directly out of the page (or screen). Remember that 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.

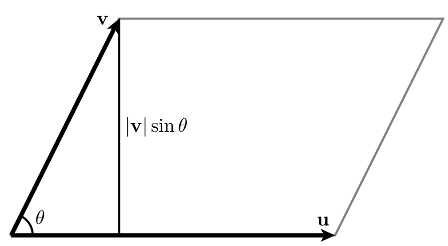


Figure 9.4.7 The parallelogram formed by $\vec{u}$ and $\vec{v}$

Key Idea 9.4.8 *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. In general, the area of a parallelogram is the product of its base and height. Based on Figure 9.4.7, 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

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

Note also that 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 (9.4.1) 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 $\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$. So 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 9.4.5.
■

Activity 9.4.3

(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 9.4.8. Suppose $\vec{u}$, $\vec{v}$, and $\vec{w}$ are vectors in $\mathbb{R}^3$ that are not coplanar ($\vec{w}$ is not a linear combination of $\vec{u}$ and $\vec{v}$). Then, $\vec{u}$, $\vec{v}$, and $\vec{w}$ define a three-dimension parallelepiped as shown in Figure 9.4.9.

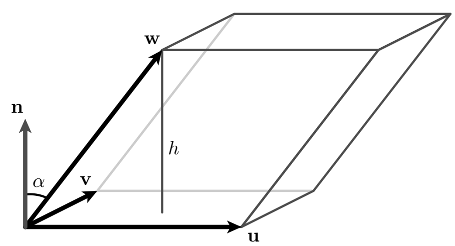


Figure 9.4.9 The parallelepiped determined by $\vec{u}$, $\vec{v}$, and $\vec{w}$

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

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

and therefore

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 *triple scalar product*.

Activity 9.4.4 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 9.4.10, sketch the cross products $\vec{u} \times \vec{v}$ and $\vec{v} \times \vec{u}$.

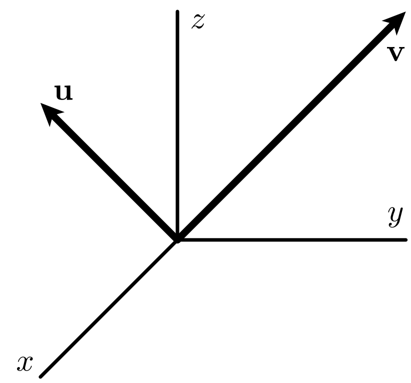


Figure 9.4.10 Vectors $\vec{u}$ and $\vec{v}$

(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.

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 9.4.11. 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**, τ, to be

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

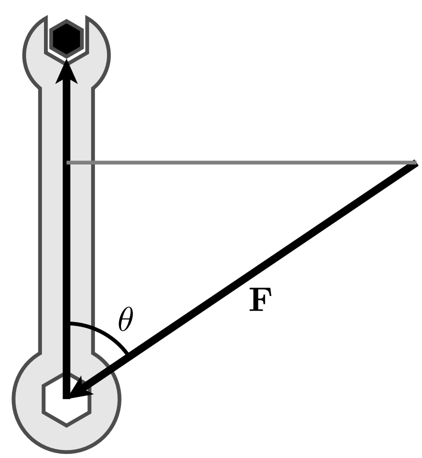


Figure 9.4.11 A force applied to a wrench

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 $\|\tau\| = \|\vec{F} \times \vec{r}\| = \|\vec{F}\|\|\vec{r}\|\|\sin(\theta)$. Thus, to produce a larger torque, we can increase either $\|\vec{F}\|$

or $\|\vec{r}\|$, which you may know if you have ever removed lug nuts when changing a flat tire. Increasing $\|\vec{F}\|$ 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 normal because you have half of $\|\vec{r}\|$.

9.4.4 Comparing the dot and cross products

In this subsection we 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.

9.4.5 Object Types and Vector Notation

In this final subsection, we are taking some time to make sense of what kind of objects are described by different expressions involving vectors. It is very easy to get caught up in all of the manipulations of symbols and new operations and write an expression that "looks" reasonable but does not actually make sense. For example $(\vec{u} \cdot \vec{v}) \times \vec{w}$ looks like a nice and reasonable calculation involving vectors (and looks a lot like the triple product...), but $(\vec{u} \cdot \vec{v}) \times \vec{w}$ doesn't make sense because the quantity in the parentheses is a scalar and you can't take the cross product of a scalar and a vector.

Activity 9.4.5 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 will be a scalar, a vector, or the expression does not make sense. You should write a sentence or two about each to explain your reasoning.

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

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

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

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

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

(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 does not make sense to perform. You can use other operations if you want.

9.4.6 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 $\vec{w}$.

9.4.7 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 $||\vec{v} \times \vec{w}|| = 2$, and the angle between $\vec{v}$ and $\vec{w}$ is θ. Find
(a) $\tan\theta =$ ______________________________
(b) $\theta =$ ______________________________

3. You are looking down at a map. A vector $\mathbf{u}$ with $|\mathbf{u}| = 9$ points north and a vector $\mathbf{v}$ with $|\mathbf{v}| = 10$ 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 $\mathbf{a} = \mathbf{i} + 10\mathbf{j} + \mathbf{k}$ 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 $O, P = (2, 3, 0)$ and $Q = (0, 3, 7)$ and compute its area using cross products.
Area=______________

6. Let $A = (-3, 4, 1)$, $B = (2, -1, 3)$, 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 $\mathbf{a} = \langle -2, 1, -5 \rangle$ and $\mathbf{b} = \langle -1, -1, -2 \rangle$

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) $2\vec{i} \times (\vec{i} + \vec{j}) =$ ____________________

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

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

10. A bicycle pedal is pushed straight downwards by a foot with a 36 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. 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 (9.4.1) 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 $\vec{v}$.

12. 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 $\vec{y}$.

13. 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 $\angle QPR$.

14. 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 $\mathbb{R}^3$.

15. 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 eminate 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.

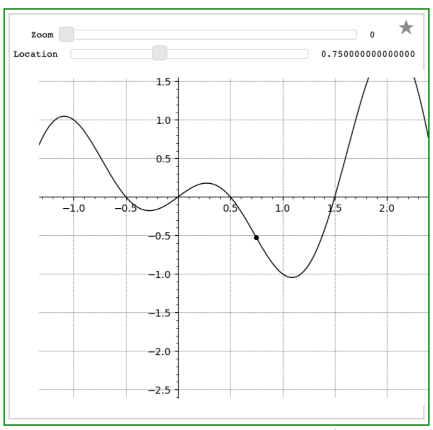
16. In Subsection 9.4.2, we stated “Because the cross product works well with linear combinations, $\vec{u} \times \vec{v}$ is perpendicular to any linear combination of $\vec{u}$ and $\vec{v}$.” We want to verify this algebraically here; Specifically, show that $\vec{u} \times \vec{v}$ will be orthogonal to any vector of the form $a\vec{u} + b\vec{v}$.

9.5 Lines in Space

Motivating Questions

- How are lines in $\mathbb{R}^3$ similar to and different from lines in $\mathbb{R}^2$?
- How can vectors make describing lines in $\mathbb{R}^3$

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 9.5.1 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 you are examining. You should visually verify at a several points that this function is differentiable because at any point shown, the graph is locally linear.



Standalone
Embed

Figure 9.5.1 A plot with the ability to zoom in on a point of $y = f(x)$, where f is a nonlinear, differentiable function

Linear functions (of the form $f(x) = ax + b$) 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. Additionally, as we study functions of two variables, 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.

Consequently, it is important for us to understand both lines and planes in space, as these correspond to graphs of linear expressions in $\mathbb{R}^2$ and $\mathbb{R}^3$. (Recall that a function is linear if it is a polynomial function whose terms all have degree less than or equal to 1. For example, $x - 1$ defines a single variable linear expression and $x + y + 2$ a two variable linear expression. But xy is not linear since it has degree two, the sum of the degrees of its factors.) We will study planes and lines in $\mathbb{R}^3$ as flat graphs before we look at curves and surfaces in general. In your study of 1-variable functions, you probably worked with ideas like slope, direction, and measuring change using lines several times before you generalized those ideas for a curve in general. Similarly, we will want to have a good idea about 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.

Preview Activity 9.5.1 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 9.5.2.

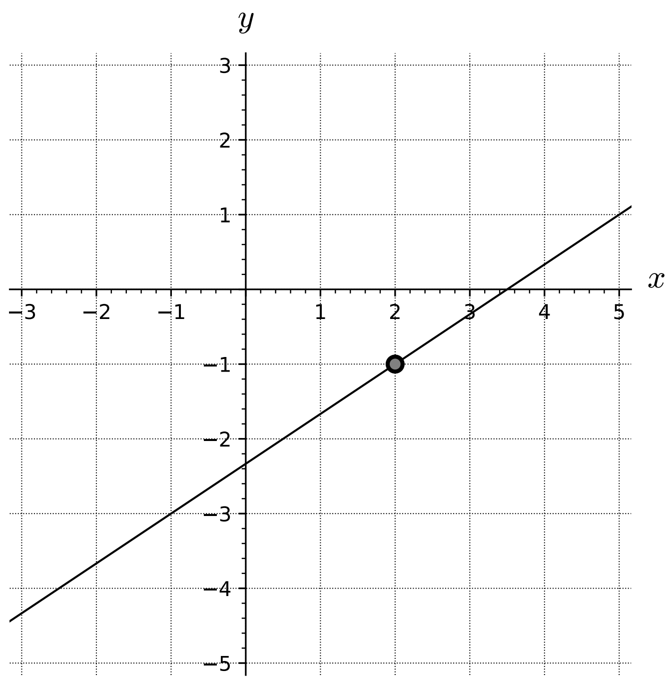


Figure 9.5.2 The line through $(2, -1)$ with slope $\frac{2}{3}$

(a) Suppose we increase x by 1 from the point $(2, -1)$. How does the y-value 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) $t = -\dfrac{1}{2}$

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

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

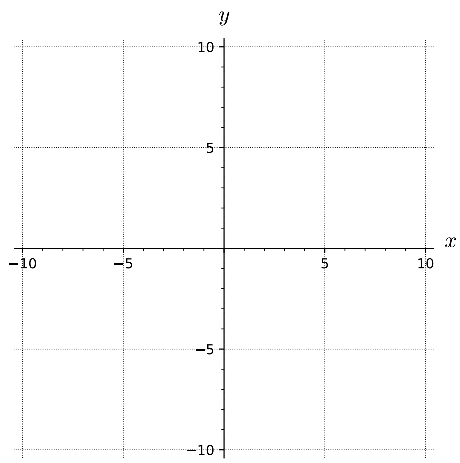


Figure 9.5.3 A blank 2D set of axes

(f) Write a few sentence 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 9.5.2). In particular you should address what aspects of the line are related to $\langle 2, -1\rangle$ or $\langle 3, 2\rangle$.

9.5.1 Lines in Space

The way most people draw a line is more connected to the ideas in Preview Activity 9.5.1 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.

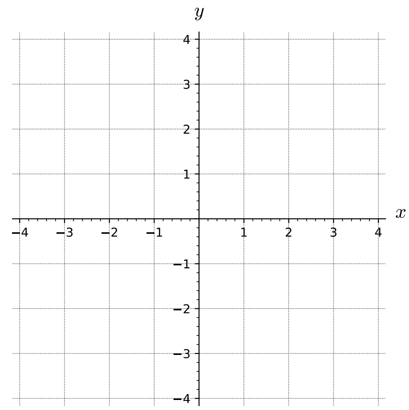


Figure 9.5.4 A blank 2D set of axes

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 almost certainly 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 end of you plot
4. went back 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 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 9.5.7. 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 space ($\mathbb{R}^n$) to provide direction, this same idea of a point and a direction determining a line works in $\mathbb{R}^n$ for any n.

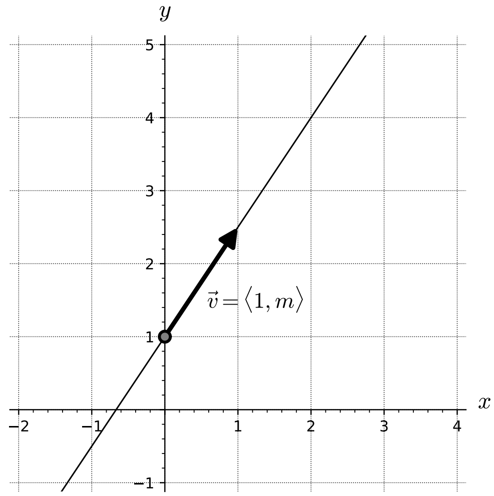


Figure 9.5.5 The line through $(0, 1)$ with slope m

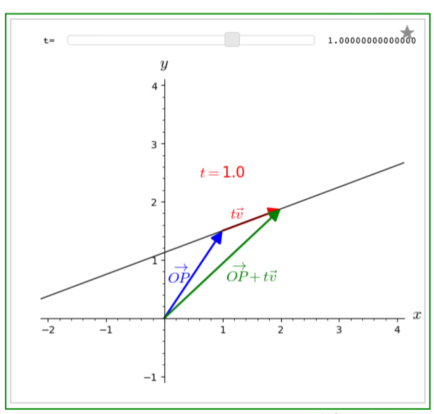
Definition 9.5.6 A **line** in space is the set of terminal points of vectors starting at a given point P that are parallel to a fixed vector $\vec{v}$. ♢

The vector $\vec{v}$ in Definition 9.5.6 is called a *direction vector* for the line. As we saw in Preview Activity 9.5.1, to find an equation for a line through point P in the direction of vector $\vec{v}$, any vector parallel to $\vec{v}$ will have the form $t\vec{v}$ for some scalar t. So, any vector emanating from the point P in a direction parallel to the vector $\vec{v}$ will be of the form

$$\overrightarrow{OP} + \vec{v}t \tag{9.5.1}$$

for some scalar t (where O is the origin).

In the figure below, 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.



Standalone
Embed

Figure 9.5.7 A line in 2-space with vector definition

Figure 9.5.7 shows the plot of a line in two-space in which we can identify the vector $\overrightarrow{OP}$ (going from the origin to our starting point P) and the vector $t\vec{v}$ as in Equation (9.5.1). Here, $\overrightarrow{OP}$ is the fixed vector shown in blue, while the direction vector $\vec{v}$ is the vector parallel to the vector shown in red (that is, the red vector represents $t\vec{v}$, and the line is traced out by the terminal points of the green vector). In other words, the tips (terminal points) of the green vectors (the vectors of the form $\overrightarrow{OP} + t\vec{v}$) trace out the line as t changes.

In particular, the terminal points of the vectors of the form in (9.5.1) define a linear function $\vec{r}$ in space of the following form, which is valid and gives some geometric intuition into each part for in any dimension.

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

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}, \tag{9.5.2}$$

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

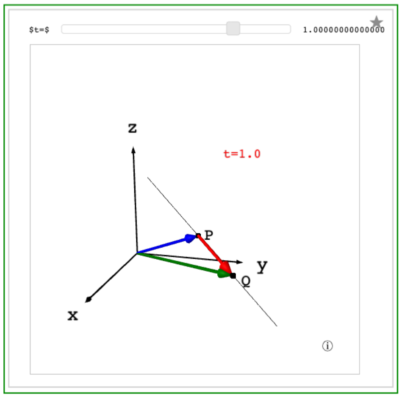
Example 9.5.8

(a) Lines in the xy-plane are commonly described using the slope-intercept equation $y = mx + b$. 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 like $\vec{v} = \langle 0, 1\rangle$, 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) Let's give 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 and we can use $\hat{\jmath} = \langle 0, 1, 0\rangle$ as our direction vector. This give 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 $\vec{r}_1(t) = \langle 0, t, 0\rangle$. This form should not be suprising since the y-axis consists of points with x- and z-coordinates of zero while the y-coordinate can be any real number.

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.

(c) If you wanted to look at a line that is less familiar you could select the line in $\mathbb{R}^3$ that goes through $(1, 2, 1)$ and moves in the direction $\langle 1, 1, -\frac{1}{2}\rangle$. The corresponding vector form would be given by $\vec{r}(t) = \langle 1, 2, 1\rangle + \langle 1, 1, -\frac{1}{2}\rangle$. Figure 9.5.9 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.



Standalone
Embed

Figure 9.5.9 A line in 2-space with vector definition

(d) You can pretty easily verify visually (using Figure 9.5.9) that L_1, the line in $\mathbb{R}^3$ that goes through $(1,2,1)$ and moves in the direction $\langle 1,1,-\frac{1}{2}\rangle$ does not intersect the y-axis. In two dimensions, two lines were parallel if they did not intersect but like slope this idea does translate to higher dimensions because L_1 is not parallel to the y-axis. In general, two lines in $\mathbb{R}^n$ are **parallel** if they can be described with direction vectors that are parallel. Remember that two vectors are parallel if they are non-zero scalar multiples of each other. This should make sense from the perspective that parallel lines should move in the same direction, but the same direction can be described by *ANY* parallel vector.

□

Activity 9.5.2 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 9.5.9 shows a similar example of a line in 3D defined by two points.

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

(b) Give 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.

9.5.2 The Parametric Equations of a Line

The vector form of a line, $\vec{r}(t) = \vec{r}_0 + t\vec{v}$ in Equation (9.5.2), describes a line as the set of terminal points of the vectors $\vec{r}(t)$. If $vr_0 = \langle x_0, y_0, z_0\rangle$ 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

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

We should note here that there are *many* different parametric equations for the same line. For example, choosing another point P on the line or another 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. In mathematical terms, we say that the parameterization of a line is not unique.

Activity 9.5.3 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, which is the same line as in Activity 9.5.2.

(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 to go and stay on the graph (forward/backward). 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) but 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 dimensions you can move in while staying on the graph (forward/backward or left/right), but the direction(s) that the graph extends does not change when you move locations on the plane.

9.5.3 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.

9.5.4 Exercises

1. Rewrite the vector equation $\mathbf{r}(t) = (3 - 3t)\mathbf{i} + (-4 - 2t)\mathbf{j} + (2 - 3t)\mathbf{k}$ 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(4, 3, -4) and parallel to the vector $5\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

Vector Form: $\mathbf{r} = \langle$________, ________, -4 $\rangle + t\langle$ ________, ________, 3 $\rangle$

Parametric form (parameter t, and passing through P when t = 0):
$x = x(t) =$______________________________
$y = y(t) =$______________________________
$z = z(t) =$______________________________

3. Consider the line which passes through the point P(5, 2, -2), and which is parallel to the line $x = 1 + 2t, y = 2 + 3t, z = 3 + 6t$

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 $\langle -4, -4, 5\rangle + t\langle -1, -2, 3\rangle$ intersects the plane $-4x + 4y - 5z = -158$.

(____________________, ____________________, ____________________)

5. 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 9.5.2.

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.

6. 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.

e. Find an equation for the plane that contains both of the lines described in this problem.

9.6 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?

In Section 9.5, we saw how to describe a line in $\mathbb{R}^n$ by setting an initial point and allowing as much movement as you want in a direction given by the direction vector. In our 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 9.6.1 In this Preview Activity, we will be looking at what happens when we allow movement on a figure 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 $\langle 1, 2, 3\rangle$.

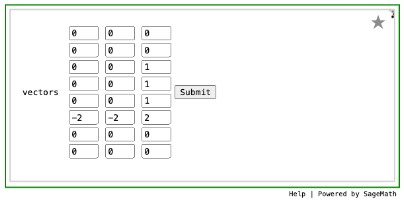
(c) Find values for c_1 and d_1 such that $\langle c_1, d_1, 1\rangle$ is orthogonal to $\langle 1, 2, 3\rangle$.

(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 interact 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.



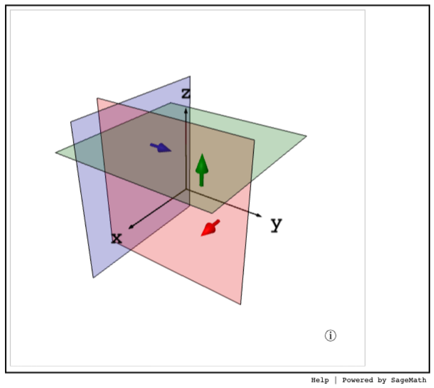
Standalone
Embed

Figure 9.6.1 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.

9.6.1 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 9.1, we studied the coordinate planes and planes parallel to them. In particular, fundamental planes were of the form *coordinate* $=$ *constant*, like $x = 1, y = -2, \text{or} z = \sqrt{3}$.

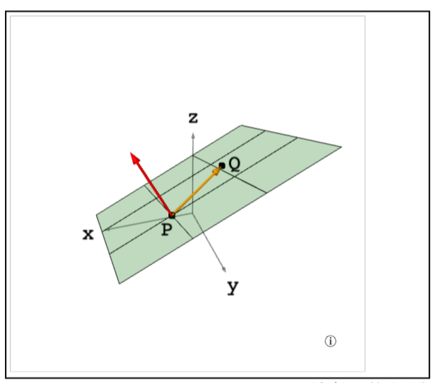


Standalone
Embed

Figure 9.6.2 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 9.6.2, 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 9.6.3 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 9.6.4. ♢



Standalone
Embed

Figure 9.6.4 A plot of a plane with normal vector $\vec{n}$ in red and vector on the plane $\overrightarrow{PQ}$ 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 $\vec{n} = \langle a, b, c\rangle$, $P = (x_0, y_0, z_0)$, and that $Q = (x, y, z)$ is an arbitrary point on the plane. Since the vector $\overrightarrow{PQ}$ lies in the plane, it must be perpendicular to $\vec{n}$. This means that

$$0 = \vec{n} \cdot \overrightarrow{PQ}$$

$$\begin{aligned}
&= \vec{n} \cdot \big(\langle x, y, z\rangle - \langle x_0, y_0, z_0\rangle\big) \\
&= \vec{n} \cdot \langle x - x_0, y - y_0, z - z_0\rangle \\
&= a(x - x_0) + b(y - y_0) + c(z - z_0).
\end{aligned}$$

The fixed vector $\vec{n}$ perpendicular to the plane is frequently called a *normal vector* to the plane. 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 $P = (x_0, y_0, z_0)$ is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \tag{9.6.1}$$

- The *vector equation* of the plane with normal vector $\vec{n} = \langle a, b, c\rangle$ containing the points $P = (x_0, y_0, z_0)$ and $Q = (x, y, z)$ is

$$\vec{n} \cdot \overrightarrow{PQ} = 0. \tag{9.6.2}$$

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 is not a unique way to give an equation for a 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 (have the same set of (x, y, z) points that make the equation a true statement.) 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 vector directions. While the angle between two vectors (as defined in Subsection 9.3.2) 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 (like when we want to measure the angle between planes), 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 equations for the plane. We will look at specifics of this in our next example.

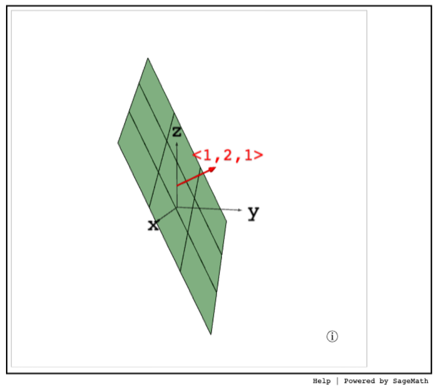
Example 9.6.5 If we would like to describe the plane passing through the point $P = (4, -2, 1)$ and perpendicular to the vector $\vec{n} = \langle 1, 2, 1\rangle$, we have

$$\langle 1, 2, 1\rangle \cdot \langle x, y, z\rangle = \langle 1, 2, 1\rangle \cdot \langle 4, -2, 1\rangle$$

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
Embed

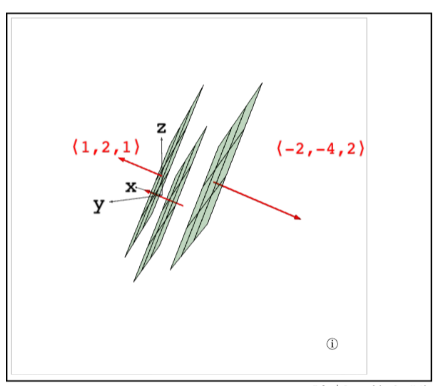
Figure 9.6.6 A plot of a plane with normal vector $\vec{n}$ in red and vector on the plane $\overrightarrow{PQ}$ in orange

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.



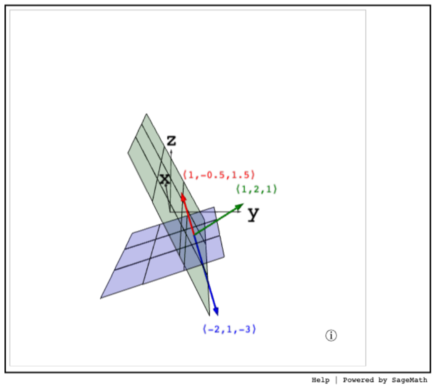
Standalone
Embed

Figure 9.6.7 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 $\langle 1, 2, 1 \rangle \cdot \langle -2, 1, -3 \rangle = -3$, so the angle between these vectors is obtuse, but 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 $\langle 1, -\frac{1}{2}, \frac{3}{2} \rangle$, which makes an acute angle with $\langle 1, 2, 1 \rangle$.

So the angle between $x + 2y + z = 1$ and $-2x + y - 3z = 4$ is

$$\arccos\left(\frac{\langle 1,2,1\rangle \cdot \langle 1,\frac{1}{2},\frac{3}{2}\rangle}{\sqrt{7}\sqrt{\frac{7}{2}}}\right) \approx 1.2373$$



Standalone
Embed

Figure 9.6.8 A plot of planes $x+2y+z = 1$ and $-2x+y-3z = 4$ with normal vectors

□

Activity 9.6.2

(a) Write a scalar equation of the plane p_1 passing through the point $(0, 2, 4)$ and perpendicular to the vector $\vec{n} = \langle 2, -1, 1\rangle$.

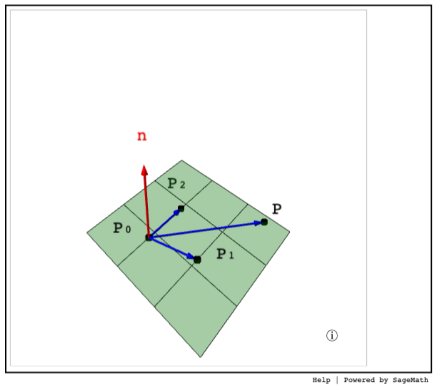
(b) Is the point $(2, 0, 2)$ on the plane p_1?

(c) Write a scalar equation of the plane p_2 that is parallel to p_1 and passing through the point $(3, 0, 4)$. (Hint: Compare normal vectors of the planes.)

(d) Write a parametric description of the line 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 l intersects the plane p_3.

Example 9.6.9 Just as two distinct points in space determine a line, three non-collinear points in space determine a plane. Consider three points P_0, P_1, and P_2 in space, not all lying on the same line as shown in Figure 9.6.10.



Standalone
Embed

Figure 9.6.10 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 P is any other point on p, it then follows that $\overrightarrow{P_0P}$ will be perpendicular to $\vec{n}$, and we have the equation:

$$\vec{n} \cdot \overrightarrow{P_0P} = 0 \tag{9.6.3}$$

□

Activity 9.6.3 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 the plane p.

(c) Find a scalar equation of the plane 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 plane q, as well as a vector $\vec{m}$ that is normal to q.

(e) The angle between two planes is the acute angle between their respective normal vectors. What is the angle between planes p and q?

9.6.2 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
$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0.$$

9.6.3 Exercises

1. Find an equation of a plane containing the three points (-3, 3, -2), (-7, 1, -5), (-7, 2, -3) in which the coefficient of x is -1.

______________________________ = 0.

2. Find an equation for the plane containing the line in the xy-plane where $y = 2$, and the line in the xz-plane where $z = 3$.

equation: ______________________________

3. Find the angle in radians between the planes $2x + z = 1$ and $5y + z = 1$.

4. 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 160 lb person and a 200 person both go 10 miles, the first at 9 mph and the second at 8 mph.

How many calories does the 160 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 160 lb person burn? ______________

How many calories per pound does the 200 lb person burn? ______________

5. 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.

6. 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.

9.7 Common Graphs in Two and Three Dimensions

Motivating Questions

- What are some sample graphs in two and three dimensions that can be used throughout our study of multivariable ideas?

In this section, we will introduce some examples of graphs in two and three dimensions that have nice algebraic properties and a variety of interesting geometric features. For our preview activity, 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 the important geometric features.

Preview Activity 9.7.1 For each of the following equations, you should:

1. Find all x-intercepts
2. Find all y-intercepts
3. Identify the shape of the graph (if you know it)
4. Identify all points where the graph will intersect with the vertical line $x = 2$
5. Identify all points where the graph will intersect with the horizontal line $y = -4$

(a) $\frac{x^2}{9} + \frac{y^2}{25} = 1$

(b) $\frac{x}{3} + \frac{y}{5} = 1$

(c) $\frac{y^2}{4} - \frac{x^2}{1} = 1$

(d) $x = 2y^2$

(e) $\frac{y^2}{9} - \frac{x^2}{25} = 0$

9.7.1 Coordinate Transformations

We have seen 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 and planes. In this section, we will introduce some other curves and surfaces that will be 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. The next activity will go over a couple of basic coordinate transformation ideas that will be helpful in generalizing our common graphs.

Activity 9.7.2 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 very 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^* 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:

(a) $(x, y) = (4, -1)$

(b) $(x, y) = (0, 0)$

(c) $(x, y) = (3, -2)$

(d) $(x, y) = (6, 3)$

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 9.7.2.b and write x^* and y^* in terms of x, y, and the coordinates of new center $(4, -1)$.

$x^* =$ ______________________ $y^* =$ ______________________

(d) Using the same coordinate systems as in part 9.7.2.a, give the (x, y)-coordinates of the following points:

(a) $(x^*, y^*) = (0, 0)$

(b) $(x^*, y^*) = (4, -1)$

(c) $(x^*, y^*) = (-5, 7)$

(d) $(x^*, y^*) = (6, 3)$

(e) Generalize your work for part 9.7.2.d and write x and y in terms of x, y, and the coordinates of new center $(4, -1)$.

$x =$ ______________________ $y =$ ______________________

Activity 9.7.3 In this activity, we look at how transform the coordinates of points when changing the scale of your coordinate system.

(a) 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 entrace 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.

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 denote the coordinates in your steps the (x, y)-coordinate system and your dog's steps with (x', y') coordinates, 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. Generalize your findings into equations that tranform between (x, y)- and (x', y')-coordinates.

The results of the previous activities are probably not suprising 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.

9.7.2 Conic Sections

You may have seen the ideas of coordinate transformation that were central to Activity 9.7.2, but hopefully you see precisely where the algebraic transformations come from and how they correspond to simple geometric transformations. For our basic shapes in 2D, we will give a very 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) means you will get an equation like

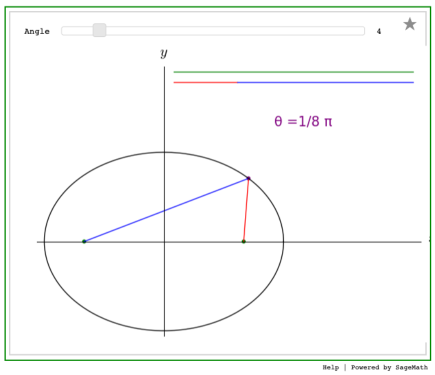
$$(x-h)^2 + (y-k)^2 = R^2$$

, 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

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

would have center $(2,-3)$ and radius $\sqrt{6}$. Notice that the transformation equations from Activity 9.7.2 show up in this example as we move a circle from being centered at the origin($x^2+y^2=R^2$) to a point (h,k) ($(x-h)^2+(y-k)^2=R^2$).

Definition 9.7.1 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. In the interact below, 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
Embed

Figure 9.7.2 A plot of an ellipse with foci in green

◊

While the definition of the ellipse given above has some great 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 9.7.4

(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 $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ and 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 $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$ and 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 9.7.3

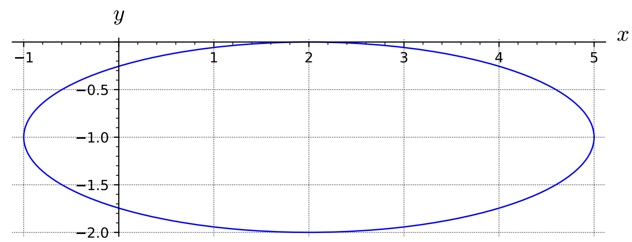
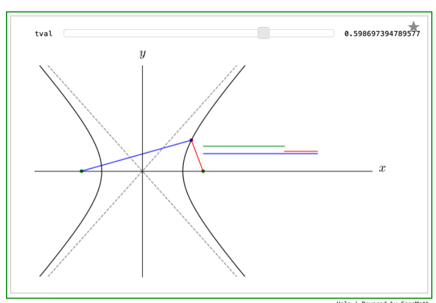


Figure 9.7.3 A plot of an ellipse

Definition 9.7.4 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.



Standalone
Embed

Figure 9.7.5 A plot of an hyperbola with foci in green

In the interact above, you can see how as you move around the hyperbola, the difference of the lengths of the blue and red segments remains constant (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 asymptote lines (shown with a dashed line). ◊

Activity 9.7.5

(a) The most basic equation for a hyperbola is $x^2 - y^2 = 1$. Make a plot of the hyperbola given by $x^2 - y^2 = 1$. You should draw and label the asymptotes ($y = \pm x$) and the vertices,the points on the hyperbola that are closest to the center which are the x-intercepts in this case.

(b) What transformation is done to convert between the hyperbolas 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

(c) What transformation is done to convert between the hyperbolas given by $x^2 - y^2 = 1$ and the graph of $y^2 - x^2 = 1$? You should 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 equation

(d) What transformations are done to convert between $x^2 - y^2 = 1$ 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.

(e) Draw a plot of $\frac{(x+2)^2}{4} - \frac{(y-3)^2}{9} = 1$ and label the center, the vertices of the hyperbola, and the asymptote lines. You will need to apply the transformations from the previous part to the asymptotes of the base hyperbola (given by $y = \pm x$) in order to get the equations of the transformed asymptotes.

(f) The graph of the equation $9x^2 - 16y^2 = 400$ is an hyperbola. Convert this equation to the form $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ and use the idea of transformations from above to graph this hyperbola and label the center, the vertices of the hyperbola, and the asymptote lines.

(g) The graph of the equation $4x^2 - y^2 + 24x - 2y + 21 = 0$ is an hyperbola. Convert this equation to the form $\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1$ and use the idea of transformations from above to graph this hyperbola. Remember to plot the vertices and asymptotes of the transformed hyperbola.

(h) Give the equation of the hyperbola shown in Figure 9.7.6

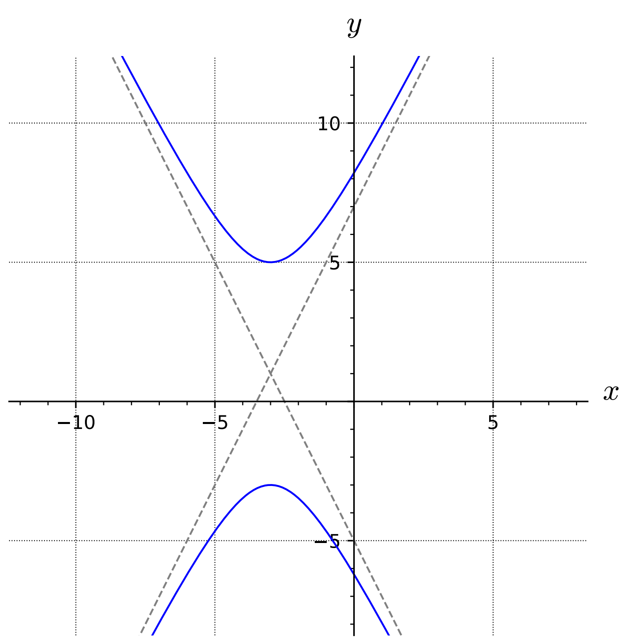
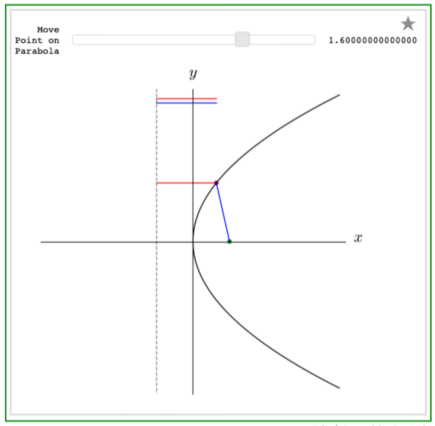


Figure 9.7.6 A plot of an hyperbola

Definition 9.7.7 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).



Standalone
Embed

Figure 9.7.8 A plot of an parabola with the focus in green and the directrix shown as a dashed line

In the interact above, you can see how as you move around the parabola, the lengths of the blue and red segments remains the same as each other. The point that is half way between the focus and the directrix is referred to as the center and the vertex of the parabola. ◊

Activity 9.7.6

(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 convert between the parabola given by $x = y^2$ and the graph of $2x = y^2$? You should 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 convert between the parabolas given by $y = x^2$ and the graph of $x = y^2$? You should 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 convert between $x = y^2$ and the graph of $\frac{x-1}{2} = \left(\frac{y+2}{3}\right)^2$? You should 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. Be sure to label the vertex and four other points on the parabola.

(g) Give the equation of the parabola shown in Figure 9.7.9

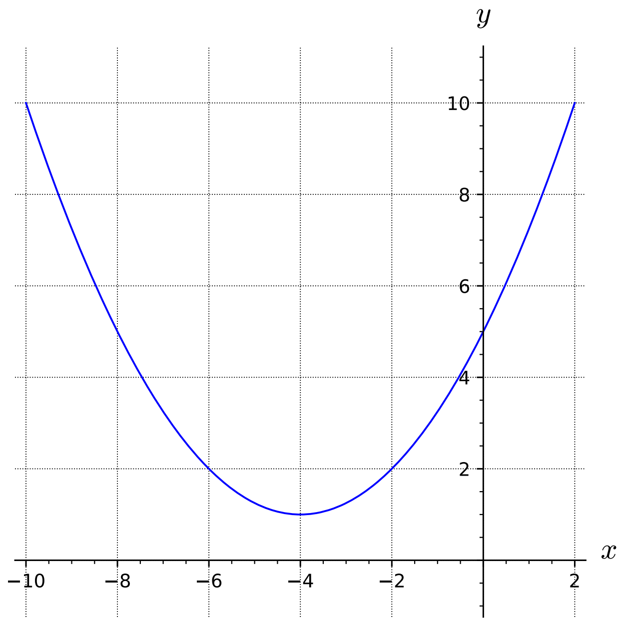


Figure 9.7.9 A plot of an parabola

9.7.3 Cylinder Surfaces

Our previous examples in this section show how to create a variety of interesting curves on the xy-plane. What happens if we consider these same equations in three dimensions?

Example 9.7.10 Let's look at an example like $(x-1) = \frac{(y+2)^2}{2}$. If we consider the graph of $(x-1) = \frac{(y+2)^2}{2}$ in 2D, then we get a parabola centered at $(1,-2)$.

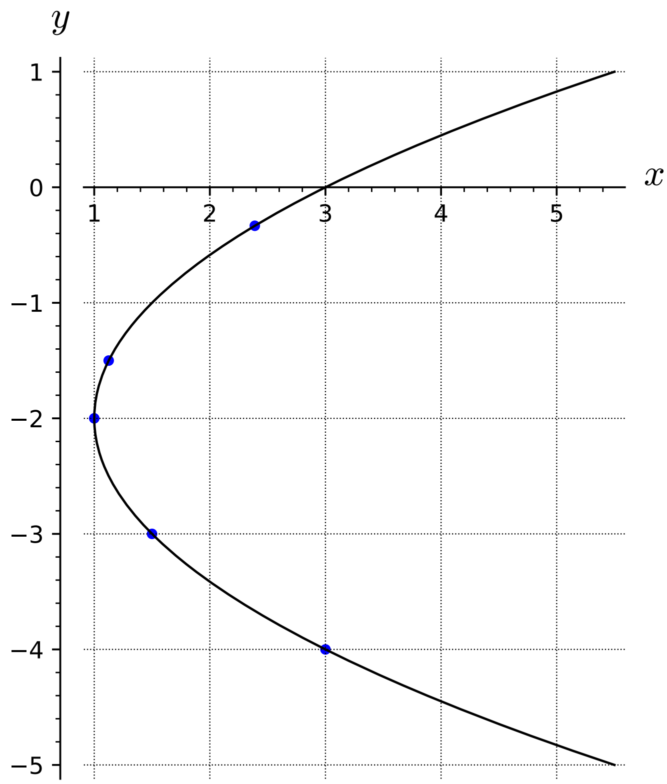
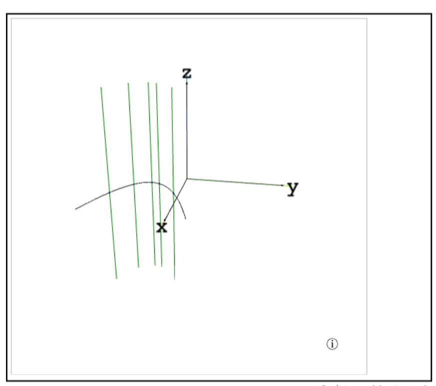


Figure 9.7.11 A plot of $(x-1) = \frac{(y+2)^2}{2}$ in the xy-plane with five points highlighted

If we want to look at the graph of $(x-1) = \frac{(y+2)^2}{2}$ in 3D, we are considering all of the (x, y, z) points that will satisfy this equation. So if we pick x and y values that satisfy $(x-1) = \frac{(y+2)^2}{2}$, then ANY choice of z will satisfy the given question. So for the each of the highlighted points on Figure 9.7.11, we can extend the graph of $(x-1) = \frac{(y+2)^2}{2}$ parallel to the z-axis. Figure 9.7.12 shows how the points on the parabola (in black) can be extended to include any z coordinate.

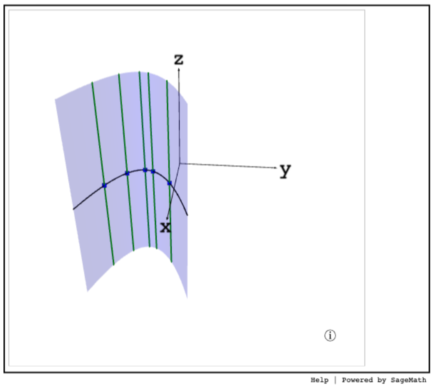


Standalone
Embed

Figure 9.7.12 A plot of $(x-1) = \frac{(y+2)^2}{2}$ generating curve and several rulings

Extending all points from the circle $(x-1) = \frac{(y+2)^2}{2}$ in the xy-plane parallel to the z-axis will give a surface. This kind of surface is called a **cylinder**

surface, the two dimensional curve you use to make the surface is called the **generating curve**, and the lines that extend in the direction of the missing variable are called **rulings**. In Figure 9.7.13, 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
Embed

Figure 9.7.13 A plot of the cylinder surface $(x-1)=\frac{(y+2)^2}{2}$ with generating curve and several rulings

□

Activity 9.7.7

(a) Draw the graph of $2x-y+1=0$ on the xy-plane.

(b) Draw the graph of $2x-y+1=0$ in xyz-space. This is called a linear cylinder surface.

(c) Draw the graph of $(x-1)^2+(y+2)^2=4$ in xyz-space. This is called a right-circular cylinder surface.

(d) Draw the graph of $\frac{x^2}{9}+\frac{z^2}{4}=1$ in xyz-space. This is called an elliptic cylinder surface.

(e) Draw the graph of $x^2-y^2=1$ in xyz-space. This is called an hyperbolic cylinder surface.

9.7.4 Quadric Surfaces

In this section, we will look at a set of surfaces with algebraic equations that are quadratic in x, y, and z. This will give us a category of example surfaces that are algebraically simple but exhibit a variety of interesting characteristics.

Activity 9.7.8 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. Draw a plot of this intersection on the xy-plane (this should be a 2D plot).

(c) 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 yz-plane. Draw a plot of your intersection on the yz-plane (this should be a 2D plot).

(d) 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 xz-plane. Draw a plot of your intersection on the xz-plane (this should be a 2D plot).

(e) 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.

(a) $z=1$

(b) $z=-2$

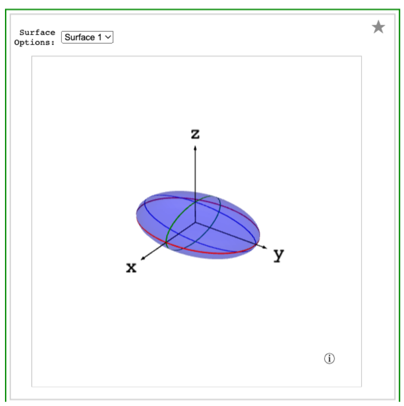
(c) $z=\sqrt{3}$

(d) $x=3$

(e) $x=-1$

(f) Draw each of these intersections on the proper fundamental planes in 3D.

(g) 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
Embed

Figure 9.7.14 A plot surfaces to select from

Activity 9.7.9

(a) 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.

(a) $z=0$

(b) $y=0$

(c) $x=0$

(d) $z=1$

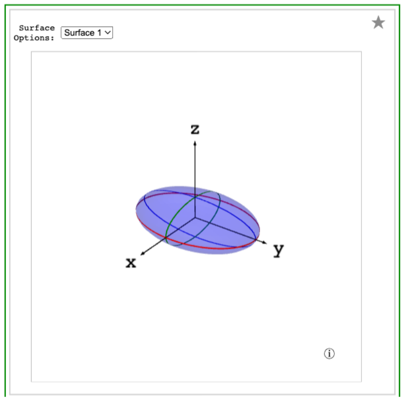
(e) $y=-2$

(f) $z=\sqrt{3}$

(g) $x=3$

(h) $x=-1$

(b) 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
Embed

Figure 9.7.15 A plot surfaces to select from

Activity 9.7.10

(a) 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.

(a) $z = 0$

(b) $y = 0$

(c) $x = 0$

(d) $z = 1$

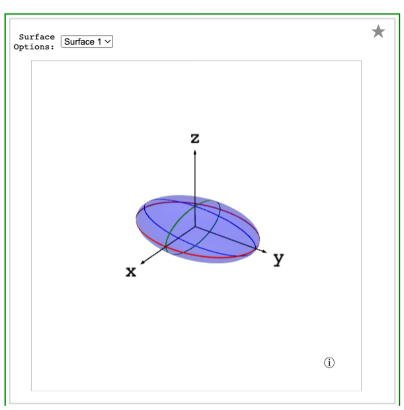
(e) $z = -2$

(f) $y = \sqrt{3}$

(g) $x = 3$

(h) $x = -1$

(b) 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
Embed

Figure 9.7.16 A plot surfaces to select from

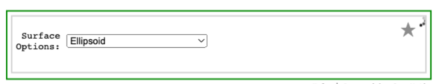
The surfaces given by $\frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{1} = 1$, $\frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$, and $\frac{x^2}{4} - \frac{y^2}{9} - \frac{z^2}{1} = 1$

are examples of **quadric surfaces**. Quadric surfaces are the surfaces generated by polynomials that are quadratic in the three coordinate variables x, y, and z. There are six main shapes of quadric surfaces; Below are 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, but in our later work, we will often consider versions of these oriented along other coordinate directions with the center not at 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
Embed

Figure 9.7.17 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 Figure 9.7.17, any vertical line (parallel to the z-axis) will intersect the saddle surface at only one place, thus 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 $\frac{x^2}{a^2} - \frac{y^2}{b^2} = z$.

Compare this to the graph of a Hyperboloid of 1-Sheet in Figure 9.7.17, where a line that is parallel to the z-axis will intersect our 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 you cannot express the graph of a hyperboloid of one sheet with none 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.

9.7.5 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.

9.7.6 Exercises

1. (a) Describe the set of points whose distance from the y-axis equals the distance from the xz-plane.

⊙ A cylinder opening along the y-axis

⊙ A cylinder opening along the x-axis

⊙ A cone opening along the x-axis

⊙ A cone opening along the y-axis

⊙ A cylinder opening along the z-axis

⊙ A cone opening along the z-axis

(b) Find the equation for the set of points whose distance from the y-axis equals the distance from the xz-plane.

⊙ $x^2 + z^2 = r^2$

⊙ $y^2 + z^2 = r^2$

⊙ $z^2 = x^2 + y^2$

⊙ $x^2 + y^2 = r^2$

⊙ $y^2 = x^2 + z^2$

⊙ $x^2 = y^2 + z^2$

2. 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) $x + y + z = 1$

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

(c) $z = -\sqrt{4 - x^2 - y^2}$

(d) $z = 3$

(e) $z = x^2 + y^2$

9.7.7 Notes to Instructors and Dependencies

This section includes a lot of basic transformation ideas that students may have a range of experience with. We included these activities as a way to make sure that students have some exposure and they get to practice simple algebraic procedures that will be used with more complexity later in the text. A more thorough approach to the conic sections can be found at (insert link to IBL materials here). You may want to have students work through activities like Activity 9.7.2 and Activity 9.7.3 on their own during times where you don't have a preview activity to prepare for new material.

9.8 Polar, Cylindrical, and Spherical Coordinates

Motivating Questions

- What are the polar coordinates of a point, and how are they related to rectangular coordinates?
- How can we convert coordinates of points or equations between rectangular and polar 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 Cartesian coordinates?
- How can we convert coordinates of points or equations between rectangular and spherical coordinates?

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 9.8.1 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 the sine of θ be positive? For which quadrants will the output of sine be negative? What directions/angles will corresponds to the output of sine being zero?

(b) If an angle θ is in standard position, for which quadrants will the cosine of

θ be positive? For which quadrants will the output of cosine be negative? What directions/angles will corresponds to the output of cosine being zero?

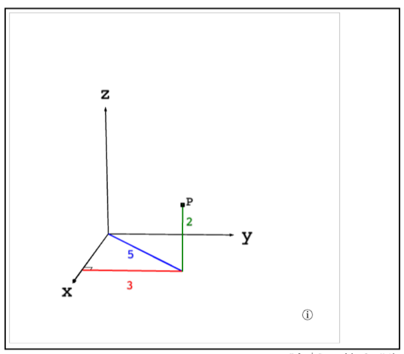
(c) If an angle θ is in standard position, for quadrants will the tangent of θ be positive? For which quadrants will the output of tangent be negative? What directions/angles will corresponds to the output of tangent being undefined? What directions/angles will corresponds to the output of tangent being zero?

(d) Given the plot of point P in Figure 9.8.1, use trigonometry and the distance formula to find the following:

(a) Find the x-coordinate of P.

(b) Find the angle between the x-axis and the blue line segment (in the xy-plane).

(c) Find the distance between the origin and P.



Standalone
Embed

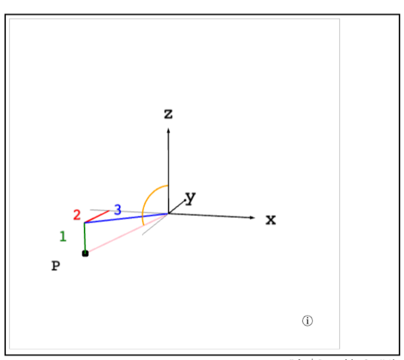
Figure 9.8.1 A plot of point P with y, z, and r labeled

(e) Given the plot of point P in Figure 9.8.2, use trigonometry and the distance formula to find the following:

(a) Find the x-coordinate of P.

(b) Find the angle between the z-axis and the pink line segment.

(c) Find the distance between the origin and P.



Standalone
Embed

Figure 9.8.2 A plot of point P with y, z, and r labeled

In this section we will define and work with a few 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.

9.8.1 Polar Coordinates

We define the rectangular coordinates 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.

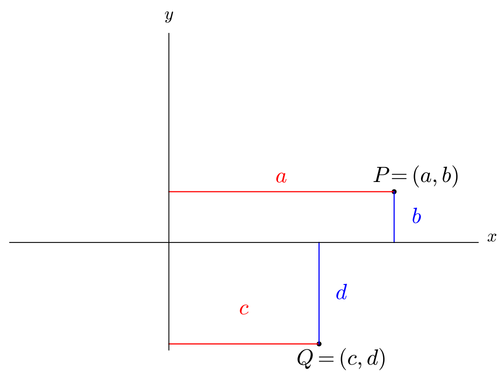


Figure 9.8.3 A plot with the 2D coordinate measurements labeled

We define the **polar coordinates** of a point in two dimensions to be (r, θ), where r is the signed distance from the origin to the point and θ is the angle from the positive horizontal axis and the line segment connecting the origin and P. The angle θ is measured with the counterclockwise direction being positive. The sign on the r will describe whether you need to move forward or backward when facing the θ direction.

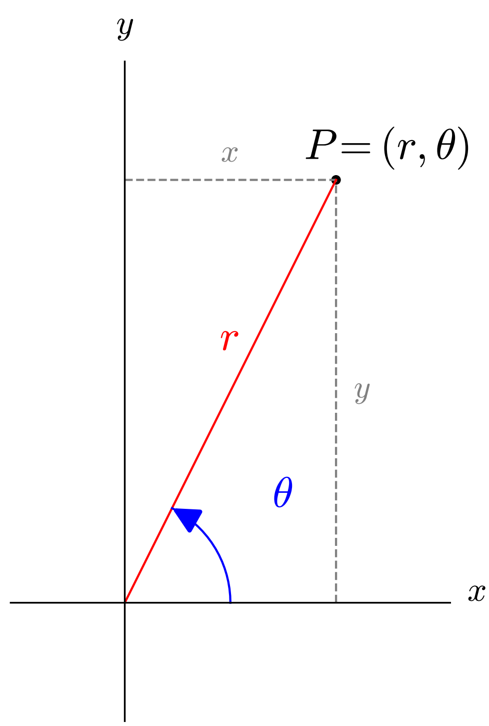


Figure 9.8.4 A plot with polar coordinate measurements labeled

A helpful way to visualize a location based on polar coordinates is to 1) stand at the origin, facing the positive horizontal axis, 2) turn θ (with positive being the counterclockwise direction), and 3) move r in the direction you are facing.

Example 9.8.5 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: 1) go to the origin and face the positive x-axis, 2) turn $\frac{\pi}{2}$ counterclockwise (to the positive y axis), and 3) move two units.

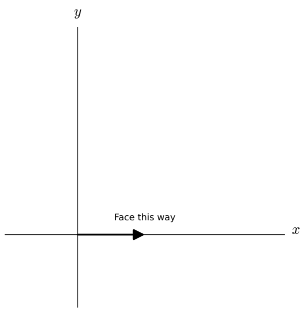


Figure 9.8.6 Starting orientation of polar coordinates

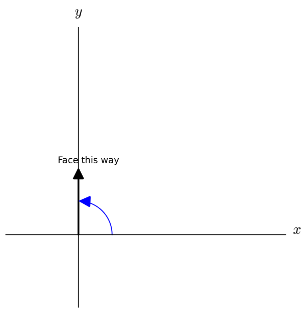


Figure 9.8.7 Rotate by $\frac{\pi}{2}$ counterclockwise

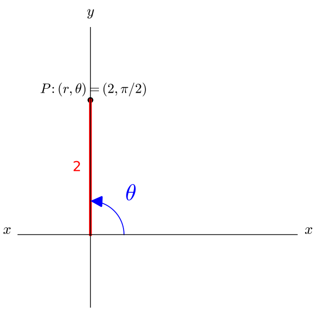


Figure 9.8.8 Move two units in this direction

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. In the next part, 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:

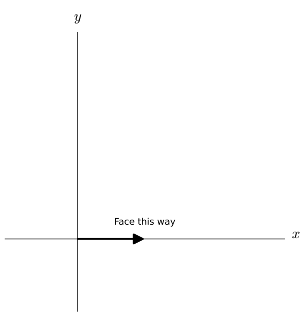


Figure 9.8.9 Starting orientation of polar coordinates

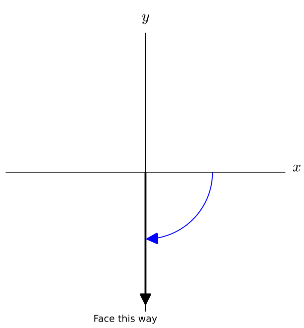


Figure 9.8.10 Rotate by $\frac{\pi}{2}$ clockwise

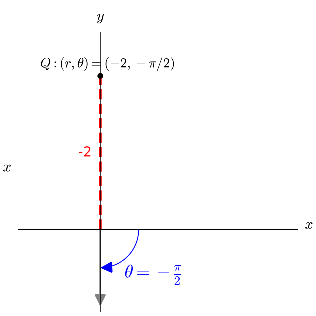


Figure 9.8.11 Move two units backwards while facing this direction

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. □

Converting between Recangular and Polar coordinates.

- *Converting from Polar to Rectangular.*

  If the polar coordinates of a point P are (r, θ), then (x, y), the rectangular coordinates of P, satisfy

$$x = r\cos(\theta) \qquad y = r\sin(\theta)$$

- *Converting from rectangular to polar.*

  If the rectangular coordinates of a point P are (x, y), then (r, θ), the polar coordinates 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 ($x = r\cos(\theta)$ and $y = r\sin(\theta)$) 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 because we are not able to solve explicitly 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 $(r, \theta) = (-\sqrt{2}, -\frac{7\pi}{4})$ or ... 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 9.8.12 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 either $\pm\sqrt{13}$. We will need to pick an appropriate angle 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 any 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 you want to choose $r = \sqrt{13}$ then you will 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 you wanted to not have to use the complementary angle to $\arctan(-\frac{3}{2})$, you will need to use the negative value of r to get $P = (r, \theta) = (-\sqrt{13}, \arctan(-\frac{3}{2}))$.

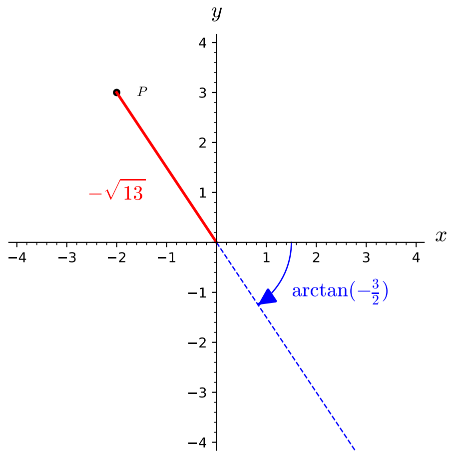


Figure 9.8.13 A visualization of P with polar coordinates $(\sqrt{13}, \arctan(-\frac{3}{2}))$

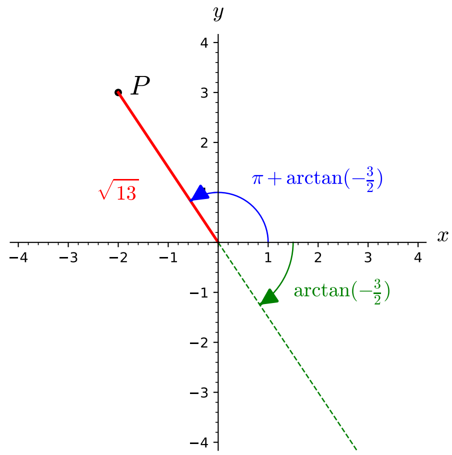


Figure 9.8.14 A visualization of P with polar coordinates $(-\sqrt{13}, \pi + \arctan(-\frac{3}{2}))$

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 $\theta = \arctan(\frac{2}{3}) + \pi$. 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 $\theta = \arctan(\frac{2}{3})$.

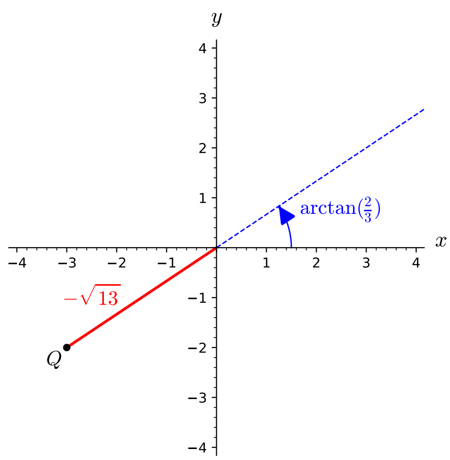


Figure 9.8.15 A visualization of Q with polar coordinates $(\sqrt{13}, \pi + \arctan(\frac{2}{3}))$

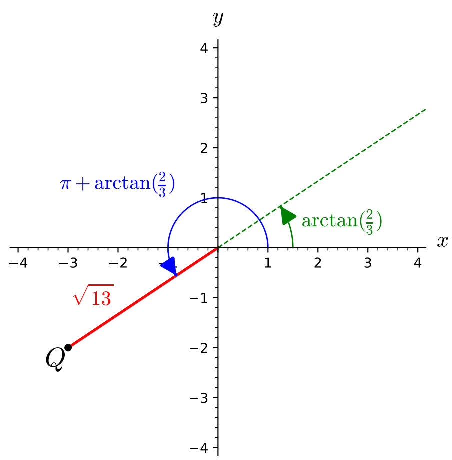


Figure 9.8.16 A visualization of Q with polar coordinates $(-\sqrt{13}, \arctan(\frac{2}{3}))$

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. □

Activity 9.8.2 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. You should simplify the trig function values of any common angles you encounter, like $\sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}$.

(a) $(x, y) = (-3, -7)$

(b) $(x, y) = (7, -3)$

(c) $(x, y) = (\sqrt{5}, -2)$

(d) $(x, y) = (\frac{3}{2}, -\frac{3}{2})$

(e) $(x, y) = (\frac{3}{2}, \frac{3}{2})$

Activity 9.8.3 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. You should simplify the trig function values of any common angles you encounter, like $\sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}$.

(a) $(r, \theta) = (1, -\frac{\pi}{6})$

(b) $(r, \theta) = (\sqrt{3}, \arctan(\frac{\sqrt{2}}{2}))$

(c) $(r, \theta) = (-\sqrt{5}, 0)$

(d) $(r, \theta) = (7, \frac{3\pi}{4})$

(e) $(r, \theta) = (3, \pi - \frac{\pi}{6})$

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, so when we say that we want to convert an equation, like $x = y^2$ to polar coordinates, we meant 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 for x and y ($x = r\cos(\theta)$ and $y = r\sin(\theta)$) and do not require interpretation.

Example 9.8.17 We can convert the equation $x = y^2$ to polar coordinates by substituting $x = r\cos(\theta)$ and $y = r\sin(\theta)$ to get

$$\begin{aligned}
x &= y^2 \\
r\cos(\theta) &= (r\sin(\theta)^2) \\
r\cos(\theta) &= r^2\sin(\theta)^2 \\
\frac{\cos(\theta)}{\sin(\theta)^2} &= r \\
r &= \cot(\theta)\csc(\theta)
\end{aligned}$$

So, the (r, θ) points that satisfy $r = \cot(\theta)\csc(\theta)$ 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. So 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 like

$$r = \begin{cases} 0 \\ \theta = 0 \\ \cot(\theta)\csc(\theta) \\ \theta \neq 0 \end{cases}$$

This examples 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. □

Example 9.8.18 In this example, we will convert the equation $r = 3\cos(\theta)$ into rectangular coordinates, identify the shape of its graph, and plot the graph of this equation. Remember that we do not have explicit conversion equations for r and θ, but we do have some equations, like $r^2 = x^2 + y^2$ which will allow us to convert to an expression in x and y that does not require intepretation. Let's try to square 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 conver to rectangular coordinates.

$$r = 3\cos(\theta) \Rightarrow r^2 = 9(\cos(\theta))^2 \Rightarrow x^2 + y^2 = 9(\cos(\theta))^2$$

Our last equation is a mix of rectangular and polar coordinates so we are not done with our coversion to rectangular (just x and y) 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 wil require some extra algebra and interpretation, so let's try a slightly different approach.

Instead of squaring both sides of our polar equation, what if we multiply both sides by r. You may have seen this as a better approach from the start or gained this insight after our work above. Remember, there are multiple paths to insight and being able to incorporate new ideas and information to form new strategies is a vital skill to solving interesting and realistic problems.

So, if we multiply both sides of our original equation by r, then we get

$$r = 3\cos(\theta) \Rightarrow r^2 = 3r\cos(\theta)$$

This equation is great because *both* the left and right hand sides have conversions to rectangular coordinates that do NOT require interpretation or cases. So, the (r, θ) points that satisfy $r = 3cos(\theta)$ 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 $r = 3cos(\theta)$) be? We can do some algebra (complete the square technique) to transform our equation into the the standard form of a circle ($(x - h)^2 + (y - k)^2 = R^2$).

$$\begin{aligned} x + y^2 = 3x \Rightarrow \quad & x - 3x & +y^2 &= 0 \\ & x - 3x + \frac{9}{4} & +y^2 &= 0 + \frac{9}{4} \\ & (x - \frac{3}{2})^2 & +(y - 0)^2 &= \left(\frac{3}{2}\right)^2 \end{aligned}$$

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 $x + y^2 = 3x$ or $(x - \frac{3}{2})^2 + (y - 0)^2 = \left(\frac{3}{2}\right)^2$ which is plotted below.

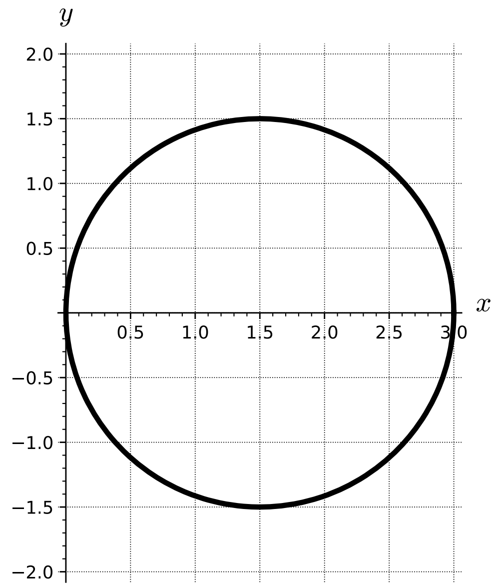


Figure 9.8.19 A plot of $r = 3\cos(\theta)$

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 9.8.2.d, part 9.8.2.e, and part 9.8.3.b will satisfy the equations of Example 9.8.18. The plot below shows how you can visualize part 9.8.3.b.

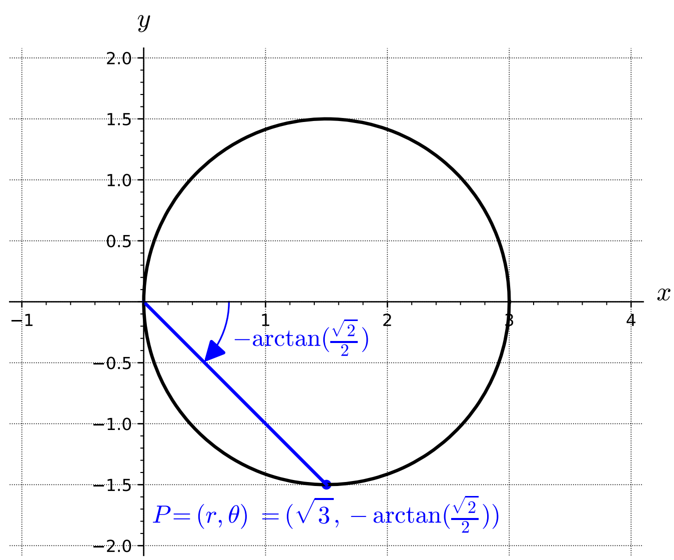


Figure 9.8.20 A plot of $r = 3\cos(\theta)$ with the polar coordinate measurements of part 9.8.3.b

Activity 9.8.4 The following activity will take you through how to create a graph of an equation in r and θ.

(a) Draw a graph of $y = \cos(x)$ for at least $0 \leq x \leq 2\pi$.

(b) For each value of θ listed below, compute the value of r corresponding to $r = \cos(\theta)$.

$$\theta = 0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}, 2\pi$$

(c) 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 $r = \cos(\theta)$.

Activity 9.8.5

(a) Graph the equation $r = 1 + \cos(\theta)$ in the polar plane. You can use

(b) Graph the equation $r = 1 - \cos(\theta)$ in the polar plane.

(c) What are the points of intersections for the graphs of $r = 1 + \cos(\theta)$ and $r = 1 - \cos(\theta)$

Activity 9.8.6 For each of the equations given below, do each of the following:

1. Convert the equation to the other coordinate system (either rectangular to polar or polar to rectangular).

2. State the shape of the graph for the equation and any other information needed to graph

3. Graph the given equation and write a description for how the converted equation (from part a) makes sense in terms of your graph.

(a) $r = 2$

(b) $\theta = \frac{\pi}{3}$

(c) $x = -1$

(d) $y = 2$

(e) $y = -x$

(f) $r = \frac{3}{1-2\cos(\theta)}$

In Figure 9.1.1, you saw how easy it was to read rectangular coordinates of a point when you had a rectangular grid. A polar point grid gives the same tools by plotting an array of constant coordinate graphs. An example of a polar grid is given below.

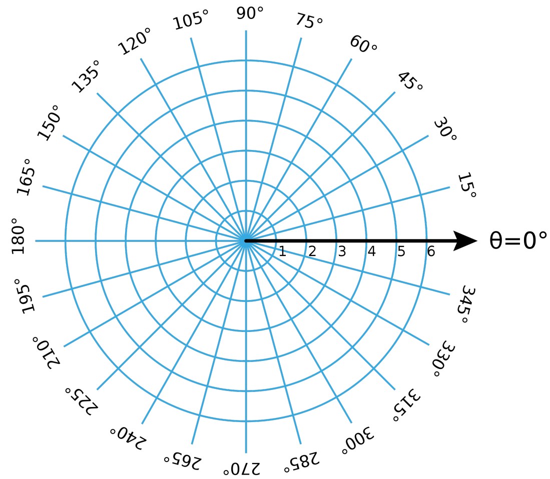


Figure 9.8.21 A polar grid (from https://commons.wikimedia.org/wiki/File:Polar_coordinates_grid.svg)

Activity 9.8.7

(a) Write a few sentences to describe what the graph of $r = c$ looks like, where c is a constant. You should consider what values of c will correspond to different shapes.

(b) Write a few sentences to describe what the graph of $\theta = d$ looks like, 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 9.8.22 The graph of the inequality $-1 \leq x \leq 1$ is the vertical strip of points with horizontal coordinate between negative one and one (inclusive) and has graph given by

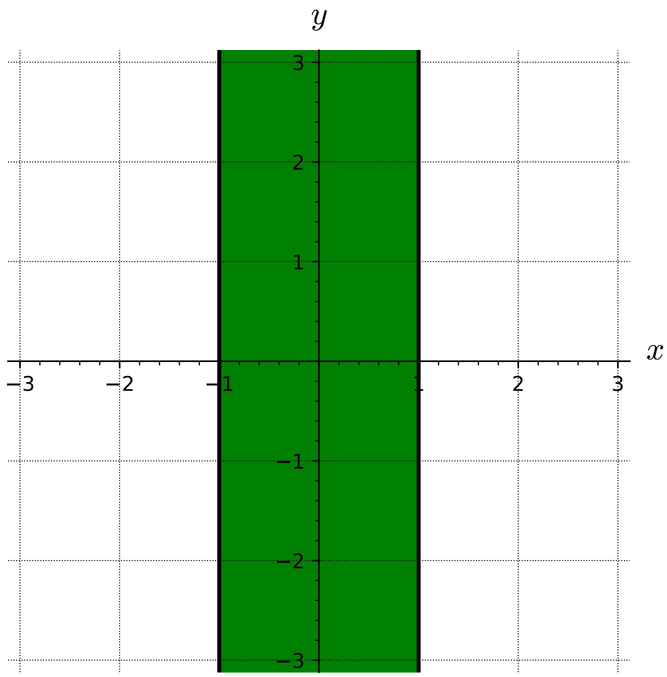


Figure 9.8.23 A plot of $-1 \leq x \leq 1$

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 on the left plot 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 on the right below.

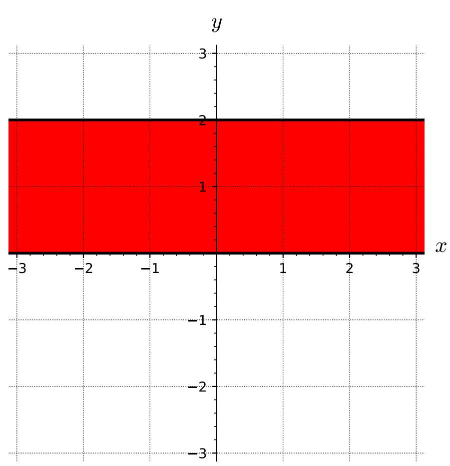


Figure 9.8.24 A plot of $0 \leq y \leq 2$

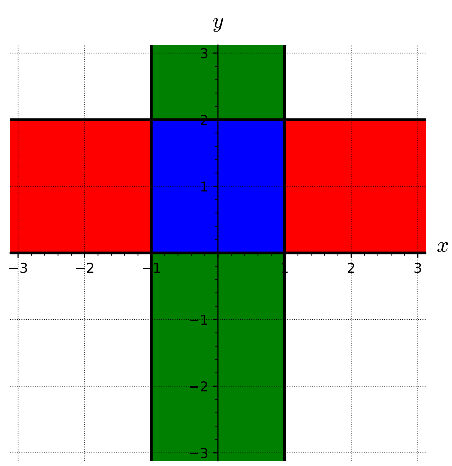


Figure 9.8.25 A plot of $-1 \leq x \leq 1$ and $0 \leq y \leq 2$

□

Activity 9.8.8

(a) Make a graph of the points that satisfy $y = x + 1$.

(b) Make a graph of 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) Make a graph of the points that satisfies the inequality $x^2 + y^2 \leq 4$

(d) Make a graph of the points that satisfies the inequalities $x^2 + y^2 \leq 4$ and $y \geq x + 1$.

Activity 9.8.9

(a) Draw a plot of the region that cooresponds to the polar inequality $1 \leq r \leq 3$. Write a sentence or two to describe this region.

(b) Draw a plot of the region that cooresponds to the polar inequality $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$. Write a sentence or two to describe this region.

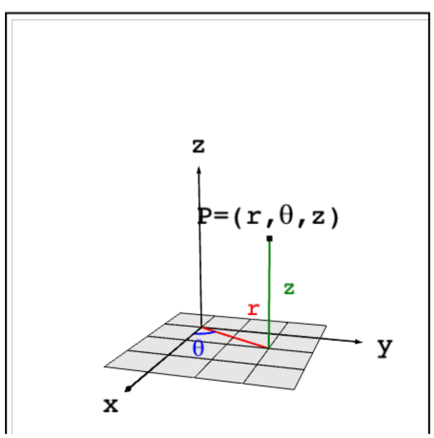
(c) Draw a plot of the region that cooresponds to the polar inequalities $1 \leq r \leq 3$, $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$. Write a sentence or two to describe this region.

(d) Draw a plot of the region that cooresponds to the polar inequality $1 \leq r \leq 3\cos(\theta)$. Write a sentence or two to describe this region.

9.8.2 Cylindrical Coordinates

We have encountered two different coordinate systems in $\mathbb{R}^2$, rectangular and polar coordinates, and seen 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. As we saw earlier, polar coordiantes will be an advantageous coordinate system when considering points/graphs/regions which have rotationally symmetric properties with respect to the origin. Given that we are already familiar with the Cartesian coordinate system for $\mathbb{R}^3$ (from Section 9.1), we will investigate cylindrical and spherical coordinates as ways to use angular measurements as part of our description for location in three dimensional space.

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 3D point onto the xy-plane. To convert between rectangular and cylindrical coordinates, we use the same conversion equations as between rectangular and polar coordinates (in 2D).



Standalone
Embed

Figure 9.8.26 A plot of point P with y, z, and r labeled

Converting between Rectangular and Cylindrical Coordinates.

- *Converting from cylindrical to rectangular.*

  If we are given the cylindrical coordinates (r, θ, z) of a point P, then the rectangular coordinates (x, y, z) of P satisfy

  $$x = r\cos(\theta) \qquad y = r\sin(\theta) \qquad z = z$$

- *Converting from rectangular to cylindrical.*

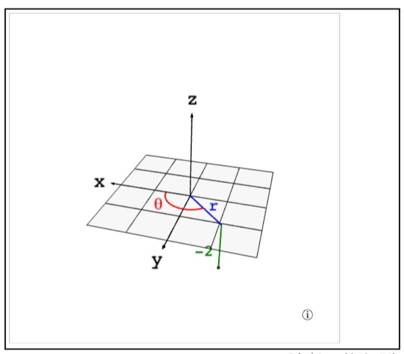
  If we are given the Rectangular coordinates (x, y, z) of a point P, then the Cylindrical coordinates (r, θ, z) of P satisfy

  $$r^2 = x^2 + y^2 \qquad \tan(\theta) = \frac{y}{x} \qquad z = z$$

  assuming $x \neq 0$.

Example 9.8.27

(a) For our first example, we will look at the point with cylindrical coordinates $(r, \theta, z) = (\sqrt{5}, \frac{3\pi}{4}, -2)$. This point is shown in the figure below with the measurements of the cylindrical coordinates.



Standalone
Embed

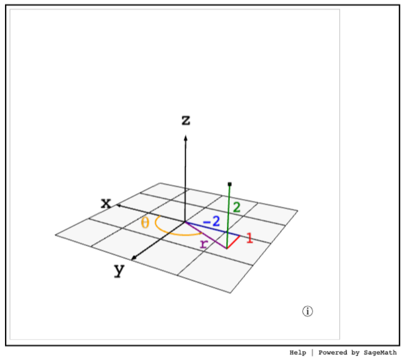
Figure 9.8.28 A plot of $(r, \theta, z) = (\sqrt{5}, \frac{3\pi}{4}, -2)$

To convert our cylindrical coordinates to rectangular, we can use $x = r\cos(\theta)$ and $y = r\sin(\theta)$ to get rectangular coordinates of $(x, y, z) = (-\frac{\sqrt{10}}{2}, \frac{\sqrt{10}}{2}, -2)$.

(b) In this example, we want to find cylindrical coordinates for the point with rectangular coordinates $(x, y, z) = (-2, 1, 2)$. Remember that just as the polar coordinates of a point in 2D are not unique, the cylindrical coordinates of a point in 3D 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 an angle of $\arctan(-\frac{1}{2}) + \pi$.

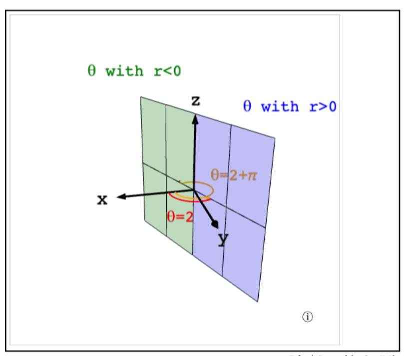
Alterntatively, we can use $\cos(\theta) = \frac{x}{r}$ since arccos will give an output of an angle in the first or second quadrants. You can verify that $\theta = \arctan(-\frac{1}{2}) + \pi = \arccos(-\frac{2}{\sqrt{5}})$. So rectangular coordinates $(x, y, z) = (-2, 1, 2)$ will cooorespond to cylindrical coordinates of $(r, \theta, z) = (\sqrt{5}, \arctan(-\frac{1}{2}) + \pi, 2)$ or $(r, \theta, z) = (\sqrt{5}, \arccos(-\frac{2}{\sqrt{5}}, 2))$.



Standalone
Embed

Figure 9.8.29 A plot of $(x, y, z) = (-2, 1, 2)$

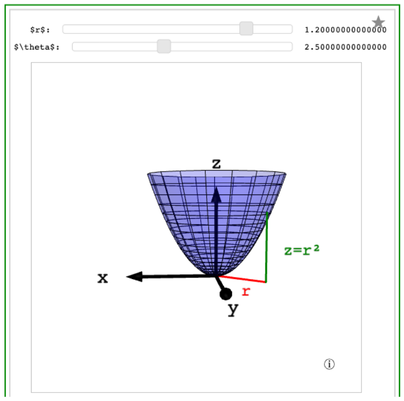
(c) In our earlier work on polar coordinates, we saw how the graph of $\theta = 2$ will be a line with slope $\tan(2)$. If we consider only positive r coordinates then we get a ray with an angle of 2 radians from the positive x-axis. If we wanted to consider the graph of $\theta = 2$ using cylindrical coordinates (in 3D), then we get a plane that contains the z-axis and goes 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
Embed

Figure 9.8.30 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 very simple because we can substitute $r^2 = x^2 + y^2$ to get $z = r^2$. This should make sense for the graph of the elliptic paraboloid because the graph is rotationally symmetric around the z-axis (which comes from not having an explicit θ dependence 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
Embed

Figure 9.8.31 A plot of $z = r^2$ with segments for r in red and $z = r^2$ in green

□

Activity 9.8.10

(a) Convert $(x, y, z) = (1, 2, 3)$ to cylindrical coordinates.

(b) Describe the region given by the 6th octant in terms of inequalities for cylindrical coordinates.

(c) Convert the basic cone, given by $z^2 = x^2 + y^2$ 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 given by $r = 2$. Write a couple of sentences about the shape and properties of this surface.

(e) Draw a plot of the region given by $0 \leq \theta \leq \pi, 0 \leq r \leq 2, 0 \leq z \leq r^2$

These activities are meant not only to give you 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 9.8.11 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 your answer 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?

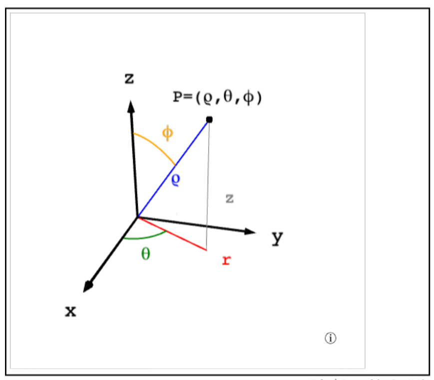
(e) Plot the graph of the cylindrical equation $z = \theta$ for $0 \leq \theta \leq 4\pi$. What does this surface look like?

9.8.3 Spherical Coordinates

Cylindrical coordinates used an angular measurement in the xy-plane as part of the description of location. Spherical coordinates uses 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
- ϕ is the (smallest) angle between the positive z-axis and the vector from the origin to the point

as illustrated in Figure 9.8.32. 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
Embed

Figure 9.8.32 A plot of a point P with spherical coordinate measurement

Coverting between rectangular and spherical coordinates.

- *Converting from rectangular to spherical.*

  If we are given the Cartesian coordinates (x, y, z) of a point P, then the spherical coordinates (ρ, θ, ϕ) of P satisfy

$$\rho = \sqrt{x^2 + y^2 + z^2} \qquad \tan(\theta) = \frac{y}{x} \qquad \cos(\phi) = \frac{z}{\rho}$$

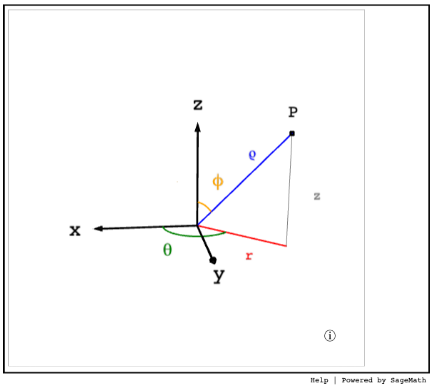
  where in the latter two equations, we require $x \neq 0$ and $\rho \neq 0$.

- *Converting from spherical to rectangular.*

  If we are given the spherical coordinates (ρ, θ, ϕ) of a point P, then the Cartesian coordinates (x, y, z) of P satisfy

$$x = \rho\sin(\phi)\cos(\theta) \qquad y = \rho\sin(\phi)\sin(\theta) \qquad z = \rho\cos(\phi)$$

Example 9.8.33 In the following example, we investigate how to think about the spherical coordinate measurement of a point and how to convert between spherical coordinates and rectangular coordinates.



Standalone
Embed

Figure 9.8.34 A plot of a point $P = (-2, 2, \sqrt{8})$ with spherical coordinate measurements

Consider the point P whose rectangular coordinates are $(-2, 2, \sqrt{8})$. 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. Then, 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 trig 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 will use $\theta = \frac{3\pi}{4} = \pi + \arctan(-1)$.

Based on the illustration in Figure 9.8.34, 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 9.8.35 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).

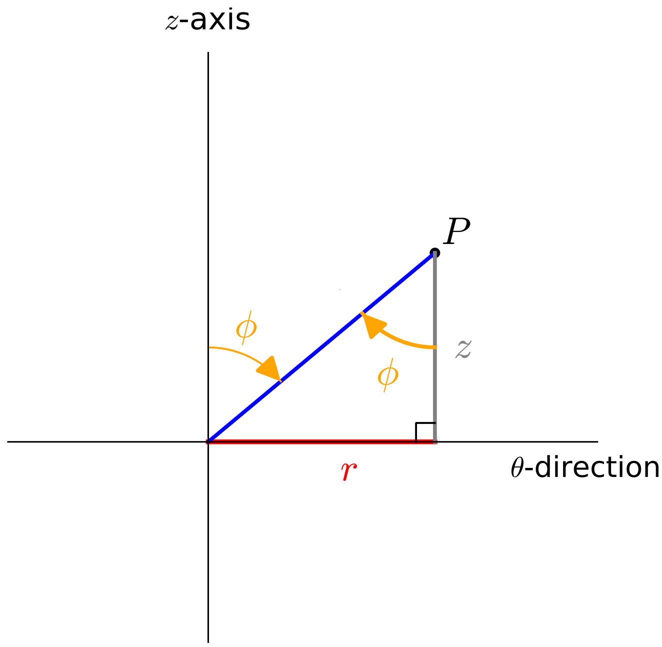


Figure 9.8.35 A 2D plot of a point P with spherical coordinate measurements in the plane of the measurement of θ

So 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 by spherical coordinates of $(\rho, \theta, \phi) = (4, \frac{3\pi}{4}, \frac{\pi}{4})$. □

Activity 9.8.12

(a) Use the conversion equations between rectangular and spherical coordinates to find the rectangular coordinates of the following points:

(a) $(\rho, \theta, \phi) = (3, \dfrac{\pi}{3}, \dfrac{\pi}{2})$

(b) $(\rho, \theta, \phi) = (4, \dfrac{4\pi}{3}, \dfrac{3\pi}{4})$

(c) $(\rho, \theta, \phi) = (\sqrt{7}, 20°, \pi)$

(b) Draw each of the points from part 9.8.12.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 9.8.12.a.

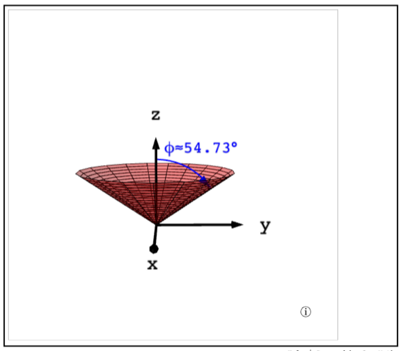
Example 9.8.36

(a) In this example we will look at converting equations and surfaces between rectangular and spherical coordinates. 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 $z^2 = 2x^2 + y^2$ 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 $r^2 = \rho^2\sin(\phi)^2$. Specifically, this means that

$$\begin{aligned} z^2 =& 2x^2 + 2y^2 \\ z^2 =& 2\left(x^2 + y^2\right) \\ \rho^2\cos(\phi)^2 =& 2\left(\rho^2\sin(\phi)^2\right) \\ \tan(\phi)^2 =& 2 \end{aligned}$$

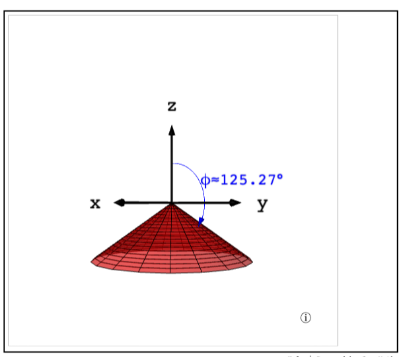
Since we are considering the part of the cone with $z \geq 0$, we can use $\phi = \arctan(\sqrt{2})$. This should not be that suprising that half cones centered on the z-axis coorespond to a constant value of ϕ.



Standalone
Embed

Figure 9.8.37 A plot of the half cone $x^2 + y^2 + z^2 = 4$ with measurement of ϕ shown in orange

If we considered the bottom half of the cone given by $z^2 = 2x^2 + y^2$ where $z \leq 0$, we would need to find the angle ϕ between $\frac{\pi}{2}$ and π that satisfies $\tan(\phi) = \sqrt{2}$. Using complementary angles we see that the bottom half of the cone given by $z^2 = 2x^2 + y^2$ where $z \leq 0$ corresponds to the spherical equation $\phi = \pi - \arctan(\sqrt{2})$.



Standalone
Embed

Figure 9.8.38 A plot of the botom half of the cone $x^2 + y^2 + z^2 = 4$ with measurement of ϕ shown in orange

(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 get the following set of algebraic simplifications:

$$z = x + y$$

$$\begin{aligned}
\rho\cos(\phi) =& \rho\cos(\theta)\sin(\phi) + \rho\sin(\theta)\sin(\phi) \\
\cos(\phi) =& \sin(\phi)\left(\cos(\theta) + \sin(\theta)\right) \\
\tan(\phi) =& \cos(\theta) + \sin(\theta)
\end{aligned}$$

If the graph of $\tan(\phi) = \cos(\theta) + \sin(\theta)$ seems unfamiliar, that is normal. This is not an expression you have (probably) seen before and does not simplify or offer geometric insight. This example shows how easily converting equations between coordinate systems without some geometric reasoning as to why can quickly turn into an exercise of algebra without much meaning.

□

Activity 9.8.13 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:

1. 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.

2. 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, like 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 9.8.14 For each of the surfaces described below,

1. 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

2. draw a plot of surface and label any important measurements in terms of the spherical coordinates

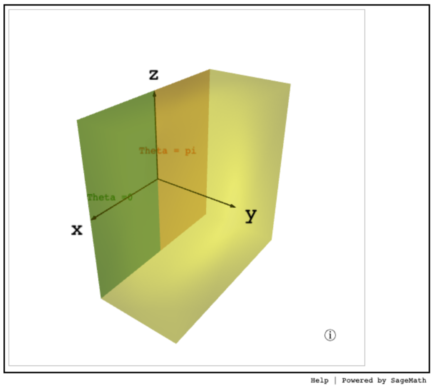
3. write a sentence or two about the corresponding equation in spherical coordinates and any assumptions made in your work.

(a) the xy-plane

(b) the xz-plane

(c) $z = x^2 + y^2$

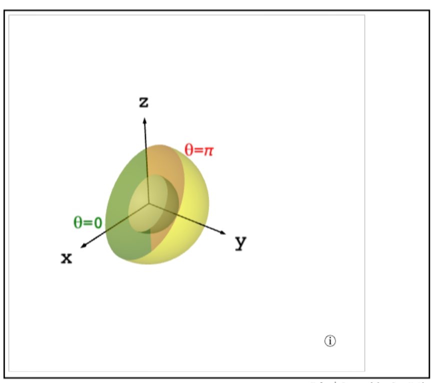
Example 9.8.39 In this example, we want to visualize and go over some tips for thinking about the region given by the spherical coordinates $0 \leq \theta \leq \pi$, $1 \leq \rho \leq 2$, and $\frac{\pi}{4} \leq \phi \leq \pi$. We will look at these inequalities in sequence to figure out what the set of points that satisfies all three looks like. Our first inequality, $0 \le \theta \le \pi$, corresponds to all points in the first, second, fifth, and sixth octants. Many students draw this region with a spherical outer boundary, but 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
Embed

Figure 9.8.40 A plot of the region $0 \le \theta \le \pi$

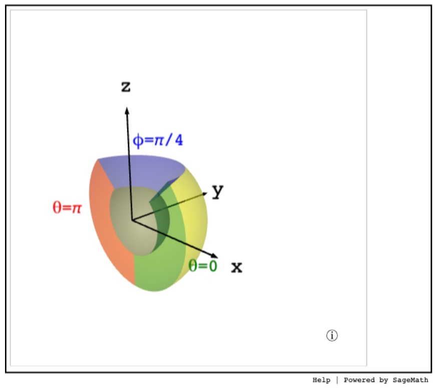
We now want to think about the set of points that satisfy $0 \le \theta \le \pi$ and $1 \le \rho \le 2$. The region given by $1 \le \rho \le 2$ is the set of points between the spheres of radius 1 and radius 2 centered at the origin. So the region given by $0 \le \theta \le \pi$ and $1 \le \rho \le 2$ is the region between the spheres of radius 1 and radius 2 centered at the origin with positive x-coordinates.



Standalone
Embed

Figure 9.8.41 A plot of the region $0 \le \theta \le \pi$ and $1 \le \rho \le 2$

When we consider the additional restriction of $\frac{\pi}{4} \le \phi \le \pi$, then we get the part of the region (shown in Figure 9.8.41) that is also below the cone given by $\phi = \frac{\pi}{4}$. This gives the region shown in Figure 9.8.42. Note that we have shown each of the boundary surfaces in different colors so you can see how each of the upper or lower bounds on each of the coordinates gives rise to a different part of our region.



Standalone
Embed

Figure 9.8.42 A plot of the region $0 \le \theta \le \pi$, $1 \le \rho \le 2$, and $\frac{\pi}{4} \le \phi \le \pi$

□

Activity 9.8.15

(a) Draw the region described by $\rho \le 1$ and $0 \le \phi \le \frac{\pi}{2}$.

(b) Draw the region described by $x^2 + y^2 \le 9$ with $x \le y$ and $-1 \le z \le 3$

(c) Give a set of inequalities in spherical coordinates that describes the points in the seventh octant.

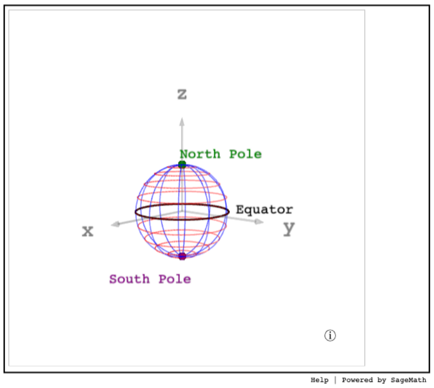
Example 9.8.43 For many people, the idea of lattitude and longitude are the most familiar use of spherical coordinates. In particular, if you look at a globe as a map of the earth, then you will notice a grid of lines that run from the top to bottom or around the globe (horizontally). These lines measure latittude and longitude of different locations on the surface of the Earth. This lattitude and longitude method of describing locations is actually spherical coordinates with a (slightly incorrect) assumption that the surface of the Earth is at a constant value of ρ.

If you look at the lines of longitude and lattitude on a globe, you will see that these are actually angular measurements because each line in the grid is marked with a degree measurement. Lattitude 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 lattitude 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 lattitude 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 lattitude of $90°$ South and a spherical coordinate measurment of $\phi = \pi = 180°$.

You can convert between lattitude and ϕ measurements by

$$\phi = 90° + \text{lattitude},$$

where north lattitudes are considered negative and south lattitudes are considered positive.



Standalone
Embed

Figure 9.8.44 A plot of lines of constant lattitude (in red) and longitude (in blue), with labels for the North and South Poles as well as the equator

The blue curves in Figure 9.8.44 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 lattitude and longitude lines form a rectangular grid, then you get a Mercator Projection map which you may recoginize. 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 9.8.44, the lines with constant lattitude are NOT all the same length, but need to be stretched out considerably near the poles to get a nice rectangular grid, as in Figure 9.8.45. Measuring the distortion of area (or volume) based on these coordinates will be an important idea in the later part of Chapter 12.

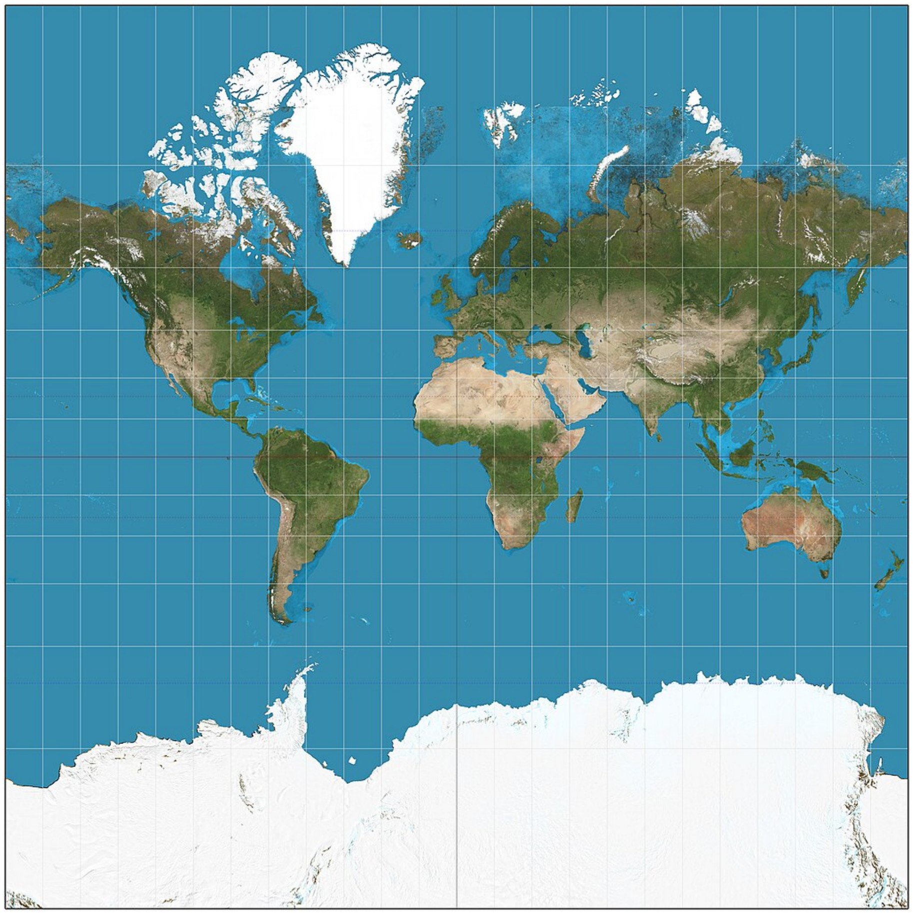
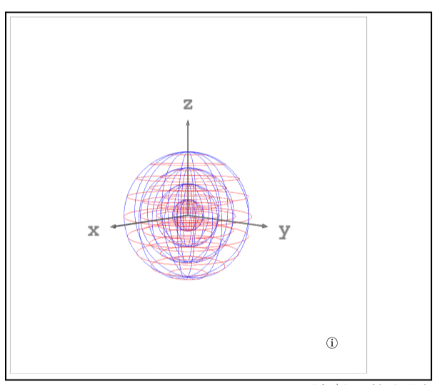


Figure 9.8.45 A Mercator projection of the Earth with lattitude and longitude lines By Strebe - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=17700069

If we extended the grid of latitude and longitude to allow for multiple radii of spheres, we get a spherical coordinate grid as shown below in Figure 9.8.46. You can see how visually cluttered this grid gets and why you will likely not see this idea used in plotting again.



Standalone
Embed

Figure 9.8.46 A plot of a grid of spherical coordinates

□

9.8.4 Summary

- 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 Cartesian 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, ϕ, θ), its rectangular coordinates are

$$x = \rho\sin(\phi)\cos(\theta), \qquad y = \rho\sin(\phi)\sin(\theta), \qquad z = \rho\cos(\phi).$$

9.8.5 Exercises

1. What are the rectangular coordinates of the point whose cylindrical coordinates are
$(r = 9,\ \theta = \frac{5\pi}{5},\ z = 4)$?
x= ____________________
y= ____________________
z= ____________________

2. What are the rectangular coordinates of the point whose spherical coordinates are
$\left(5, -\frac{1}{6}\pi, -\frac{1}{3}\pi\right)$?
$x =$ ____
$y =$ ____
$z =$ ____

3. What are the cylindrical coordinates of the point whose spherical coordinates are
$(1,\ 5,\ \frac{4\pi}{6})$?
$r =$ ____
$\theta =$ ____
$z=$ ____

4. 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: ____

5. 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 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 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?

In our previous work we have seen several examples of curves in space, such as lines in three dimensions and conic sections (in 2D). Recall from The vector form of a line that 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 single vector equation

$$\vec{r}(t) = \vec{r}_0 + t\vec{v}.$$

From this perspective, $\vec{r}(t)$ is a function that depends on the parameter t and outputs a vector. When the output vectors are graphed with initial point at the origin, then the terminal points of this vector output will trace out the line in space, like Figure 9.5.9. We call $\vec{r}(t)$ a vector valued function of one variable (or simply a vector valued function, for now) 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.

Similar to lines, other curves in space are one-dimensional objects (there is only one way to move along this object, which we refer to as forward/backward), and thus we would like to similarly express the coordinates of points on a given curve in terms of a single variable that could describe how we are moving along this path. For instance, the graph of an ellipse, like Figure 9.7.3, cannot be expressed with y as a function of x, but an ellipse is still a one dimensional graph because at any point on the graph, there is only one dimension to move (forward/backward) along the graph.

Vector valued functions (of one dimension) are a perfect vehicle for describing curves in general; we can use vectors based at the origin to identify points in space, and connect the terminal points of these vectors to draw a curve in space. This approach will allow us to draw an incredible variety of graphs in 2- and 3-space, as well as to identify and describe curves in n-space for any n.

This same approach to drawing curves will also allow us to represent traces and cross sections of surfaces in space, which we will discuss more in Chapter 11.

Preview Activity 10.1.1 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 each of these vectors with initial point at the origin on the same set of axes.

- $\langle \cos(0), \sin(0) \rangle$
- $\langle \cos\left(\frac{\pi}{2}\right), \sin\left(\frac{\pi}{2}\right) \rangle$
- $\langle \cos(\pi), \sin(\pi) \rangle$
- $\langle \cos\left(\frac{3\pi}{2}\right), \sin\left(\frac{3\pi}{2}\right) \rangle$

(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.

- $\langle \cos\left(\frac{\pi}{4}\right), \sin\left(\frac{\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{3\pi}{4}\right), \sin\left(\frac{3\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{5\pi}{4}\right), \sin\left(\frac{5\pi}{4}\right) \rangle$
- $\langle \cos\left(\frac{7\pi}{4}\right), \sin\left(\frac{7\pi}{4}\right) \rangle$

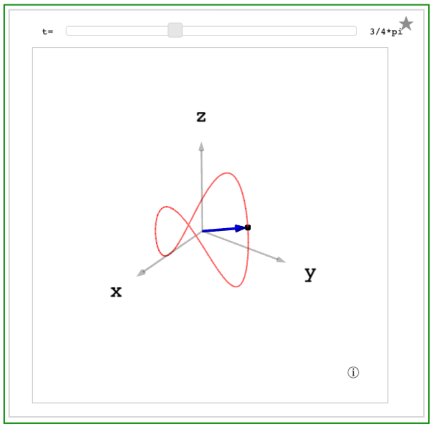
(c) Based on the data you are seeing from the two sets of vector outputs of your software, it appears that the output of your software is the set of vectors of the form $\langle \cos(t), \sin(t) \rangle$, where t assumes values from 0 to 2π. Sketch the set of *terminal* points of $\langle \cos(t), \sin(t) \rangle$, where t assumes values from 0 to 2π and 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 driven be if you drove for t from 0 to 4π?

In the preview activity, we saw how location data can be given by vectors relative to a common initial point. In this section, we will continue to look at the implications of using this as a description of movement along a curve.

10.1.1 Vector-Valued Functions

As in Preview Activity 10.1.1, we can think of a point on the curve shown in Figure 10.1.1 as resulting from a vector from the origin to the point. As the point we are looking at travels along the curve, the vector changes in order to terminate at the desired point. Figure 10.1.1 shows a curve in space with a vector from the origin to a highlighted point on the curve. You can use the slider to change the location on the curve to see how changing the scalar input (parameter) will give different outputs with terminal points on the curve.



Standalone
Embed

Figure 10.1.1 A plot of a curve in space with a position vector shown

Thus, we can think of the curve as a collection of terminal points of vectors emanating from the origin. We therefore view a point traveling along this curve as a function of time t, and define 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 vector outputs of $\vec{r}$ then trace out the curve in space. From this perspective, the x, y, and z coordinates of the point are functions of time, t, say

$$x = x(t), \quad y = y(t), \quad \text{and} \quad z = z(t),$$

and thus we have three coordinate functions that enable us to represent 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 10.1.2 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 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 particularly 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 9.5, 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 10.1.1 has the parametric equations

$$x(t) = \cos(t), \quad y(t) = \sin(t), \quad \text{and} \quad z(t) = \cos(t)\sin(t)$$

which is represented as the vector-valued function $\vec{r}$

$$\vec{r}(t) = \langle \cos(t), \sin(t), \cos(t)\sin(t) \rangle$$

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?" This is a good question because all things being equal, simpler tends to be better. But vectors have a LOT more tools and uses and we will use these vector tools throughout the rest of this course (and a few more) to make calculations easier and to make geometric sense of many measurements. For example, we used a (constant) vector to measure the direction a line moves 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), or how much of a measurement is in a particular direction, or how much is orthogonal to another direction, or...

Activity 10.1.2 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.

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. $\vec{r}(t) = \langle \cos(t^2), \sin(t^2) \rangle$

The examples in Activity 10.1.2 illustrate that a parameterization allows us to look not only at the graph, but at 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 10.2, we will begin to investigate the calculus of vector-valued functions, which 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 allow us to not only indicate the points on the curve itself, but also to describe how motion can occur along the curve.

In the next subsection we will cover some ideas that will be useful in parameterizing curves but first we will address a couple of ideas related to the new category of functions we will be studying. Remember that the big idea behind functions is that for each input there can only be one output assigned to that input. The domain of a function is the set of allowed inputs for the function. Sometimes a restricted set of inputs will be given that limits what inputs should be considered, but unless otherwise stated, 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
- restricted domains from other functions like arcsin or logarithms

The domain of of a vector valued function of one variable will the subset of $\mathbb{R}$ where all of the component functions are defined. In other words, the domain of the 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 output will be defined.

Example 10.1.3 For example, if $\vec{r}(t) = \langle \frac{-1}{t-2}, \sqrt{25-t^2}, \ln(t+1)\rangle$, then the domains of the component functions will be $\{x \in \mathbb{R} | x \neq 2\}$, $[-5, 5]$, and $(-1, \infty)$. For only the parameter values that *all* three component functions are defined will $\vec{r}$ have an output. Thus the domain of $\vec{r}(t)$ will be $(-1, 2) \cap (2, 5]$. □

10.1.2 Parameterizing Curves

In this subsection we will look at some examples related to parameterizing curves in space. A bigger 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.

Using parametric equations to define vector-valued functions in two dimensions is much more versatile than just 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

$$\vec{r}(t) = \langle t, f(t)\rangle,$$

and thus every single-variable function may be described parametrically. In addition, as we saw in Preview Activity 10.1.1 and Activity 10.1.2, 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). (As a side note: vector-valued functions make it easy to plot the inverse of a one-to-one function in two dimensions. To see how, if $y = f(x)$defines a one-to-one function, then we can parameterize this function by $\vec{r}(t) = \langle t, f(t)\rangle$. Since the inverse function just reverses the role of input and output, a parameterization for f^{-1} is $\langle f(t), t\rangle$.)

Example 10.1.4 Parameterizing $y = f(x)$. In this example we will look at parameterizing the parabola given by $(x-2)^2 = 8(y+1)$. From Figure **??**, you can see that the graph of $(x-2)^2 = 8(y+1)$ will pass the vertical line test, which means we can write the y-coordinate as a function of the x-coordinate. Specifically, we get $y = f(x) = \frac{(x-2)^2}{8} - 1$. Using the idea above, we set $x = t$ to be our parameter, which means we can express both coordinates of points on our curve as a function of t. In particular, $\vec{r}(t) = \langle t, f(t)\rangle = \langle t, \frac{(t-2)^2}{8} - 1\rangle$ will parameterize the parabola given by $(x-2)^2 = 8(y+1)$.

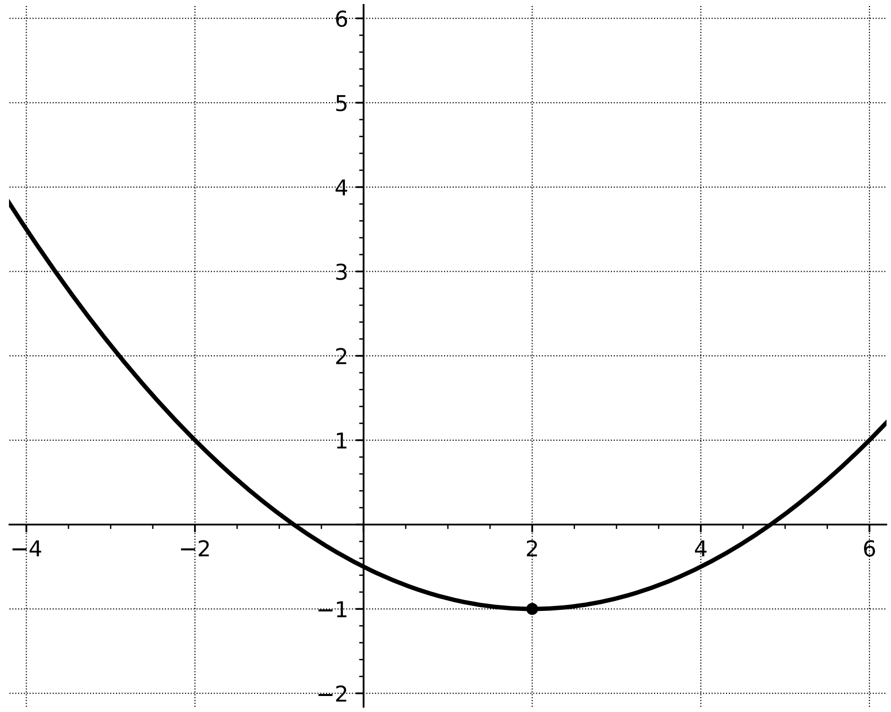


Figure 10.1.5 A plot of the parabola given by $(x-2)^2 = 8(y+1)$

Remember that there is NOT a unique way to parameterize a curve, so the choices above are not the only way to describe motion on the given parabola. You could move along the curve with double the speed (relative to the previous example), which would give $\vec{r}(t) = \langle 2t, \frac{(2t-2)^2}{8} - 1\rangle$.

Other choices for the parameterization might include wanting to center your parameter on the vertex ($t = 0$ corresponds to $(2, -1)$), which could give a parameterization like $\vec{r}(t) = \langle t-2, \frac{t^2-2}{2}\rangle$. Remember that there is more than one way to walk the same path. □

Activity 10.1.3 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 3-space.)

e. 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 intersection of surfaces in space are often curves and a convenient way of describing different conditions. Thus, we may determine parameterizations for these curves defined by satisfying multiple equations. For example, if $z = \cos(x^2 + y^2)$, the intersection with the $y = 1$ fundamental plane is given by setting $y = 1$ in $z = \cos(x^2 + y^2)$. We can then let x be parameterized by the variable t; then, the trace is the curve whose parameterization is $\langle t, 1, \cos(t^2+1)\rangle$.

Activity 10.1.4 For this activity we will consider the paraboloid defined by $z = x^2 + y^2$.

(a) Find a parameterization for the intersection of the paraboloid with the $x = 2$ fundamental plane. What type of curve does this intersection describe?

(b) Find a parameterization for the paraboloid and the $y = -1$ fundamental plane. What type of curve does this intersection describe?

(c) Find a parameterization for the intersection of the paraboloid with the fundamental plane $z = 25$. What type of curve does this intersection describe?

(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: $\sec^2(t) - \tan^2(t) = 1$.

As you saw in Activity 10.1.4, it will be convenient to parameterize conic sections and other curves defined by common relationships between coordinates. In our next activity, we will generalize our work for parameterizing circles earlier to give a general parameterized of a translated ellipse. These ideas are extended in Exercises where hyperbolas and parabolas are parameterized.

Activity 10.1.5 Earlier 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 equation for the 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

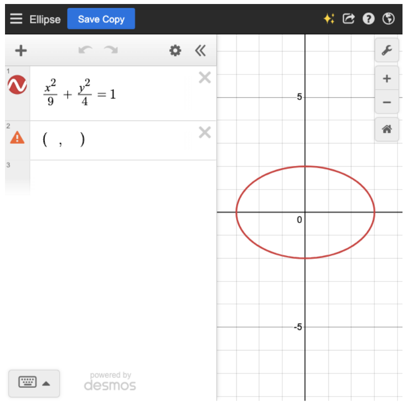
$$x(t) = a\cos(t) \qquad y(t) = b\sin(t).$$

(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

Figure 10.1.6 A Desmos Interactive for Activity **??**

(c) In Figure **??**, 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. 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 the equation in x and y as well as the parametric equations so that the ellipse has a center of $(-2, 1)$.

(f) Since we know that

$$(\sin(t))^2 + (\cos(t))^2 = 1 \quad \text{for every } t \in \mathbb{R}$$

and the form of a translated ellipse is given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1,$$

set

$$\frac{(x-h)^2}{a^2} = (\sin(t))^2$$

and solve for x.

(g) Set up a similar equation involving y and $\cos(t)$ then solve for y to get a general set of parametric equations for the translated ellipse.

10.1.3 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.

10.1.4 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

$$\mathbf{r}(t) = \left\langle \ln(7t), \sqrt{t+15}, \frac{1}{\sqrt{14-t}} \right\rangle$$

using interval notation[1].

Domain: ______________________

2. Find a parametrization of the circle of radius 5 in the xy-plane, centered at the origin, oriented clockwise. The point $(5,0)$ should correspond to $t = 0$. Use t as the parameter for all of your answers.

$x(t) =$ ______________

$y(t) =$ ______________

3. Find a vector parametrization of the circle of radius 6 in the xy-plane, centered at the origin, oriented clockwise so that the point $(6,0)$ corresponds to $t = 0$ and the point $(0,-6)$ corresponds to $t = 1$.

$\vec{r}(t) =$ ______________

4. Find a vector parametric equation $\vec{r}(t)$ for the line through the points $P = (2,-3,3)$ and $Q = (3,-7,5)$ for each of the given conditions on the parameter t.

(a) If $\vec{r}(0) = \langle 2,-3,3\rangle$ and $\vec{r}(7) = \langle 3,-7,5\rangle$, then

$\vec{r}(t) =$ ______________

(b) If $\vec{r}(7) = P$ and $\vec{r}(9) = Q$, then

$\vec{r}(t) =$ ______________

(c) If the points P and Q correspond to the parameter values $t = 0$ and $t = -3$, respectively, then

$\vec{r}(t) =$ ______________

5. Suppose parametric equations for the line segment between $(7,8)$ and $(0,-5)$ have the form:

$$x = a + bt$$
$$y = c + dt$$

If the parametric curve starts at $(7,8)$ when $t = 0$ and ends at $(0,-5)$ at $t = 1$, then find a, b, c, and d.

$a =$ ________,

$b =$ ________,

$c =$ ________,

$d =$ ________.

6. Find a parametrization of the curve $x = -2z^2$ in the xz-plane. Use t as the parameter for all of your answers.

$x(t) =$ ______________

$y(t) =$ ______________

$z(t) =$ ______________

7. Find parametric equations for the quarter-ellipse from $(2,0,8)$ to $(0,-5,8)$ centered at $(0,0,8)$ in the plane $z = 8$. Use the interval $0 \le t \le \pi/2$.

$x(t) =$ ______________

[1] /webwork2_files/helpFiles/IntervalNotation.html

$y(t) =$ ______________________________
$z(t) =$ ______________________________

8. Are the following statements true or false?

(a) The parametric curve $x = (3t+4)^2, y = 5(3t+4)^2 - 9$, for $0 \leq t \leq 3$ is a line segment.

(b) A parametrization of the graph of $y = \ln(x)$ for $x > 0$ is given by $x = e^t, y = t$ for $-\infty < t < \infty$.

(c) The line parametrized by $x = 7, y = 5t, z = 6 + t$ is parallel to the x-axis.

9. Find a vector function that represents the curve of intersection of the paraboloid $z = 3x^2 + 4y^2$ and the cylinder $y = 6x^2$. Use the variable t for the parameter.
$\mathbf{r}(t) = \langle t,$ ____________________, ____________________$\rangle$

10. 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:
$x = 10\theta - 8\sin(\theta)$
$y = 10 - 8\cos(\theta)$
What must we have for R and d?
R= __________
d = __________

11. 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 $[0, \pi]$.

12. 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. Explain why the vector function $\vec{r}$ defined by $\vec{r}(t) = \langle a\cos(t), b\sin(t) \rangle$, $0 \leq t \leq 2\pi$ is one 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 counterclockwise.

c. Find a parameterization of the ellipse $\frac{(x+3)^2}{4} + \frac{(y-2)^2}{9} = 1$.

d. Determine the x-y equation of the ellipse that is parameterized by

$$\vec{r}(t) = \langle 3 + 4\sin(2t), 1 + 3\cos(2t)\rangle.$$

13. Consider the two-variable function $z = f(x, y) = 3x^2 + 4y^2 - 2$.

a. Determine a vector-valued function $\vec{r}$ that parameterizes the curve which is the $x = 2$ trace of $z = f(x, y)$. 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 $z = f(x, y)$. 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 $z = f(x, y)$. 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 $z = f(x, y)$. In addition, write two sentences to describe the characteristics of the surface.

14. 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.

15. 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. $\vec{r}(t) = \langle \cos(st), \sin(st)\rangle$

16. **Parameterizing a Translated Hyperbola.**Another conic section that will be useful to quickly parameterize is the hyperbola. The translated hyperbola is given by

$$\frac{(x-h)^2}{a^2} - \frac{(y-k)^2}{b^2} = 1 \tag{10.1.1}$$

for horizontal orientation and

$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1 \tag{10.1.2}$$

for vertical orientation.

(a) Find a trig identity of the form

$$(\underline{\hspace{2cm}})^2 - (\underline{\hspace{2cm}})^2 = 1$$

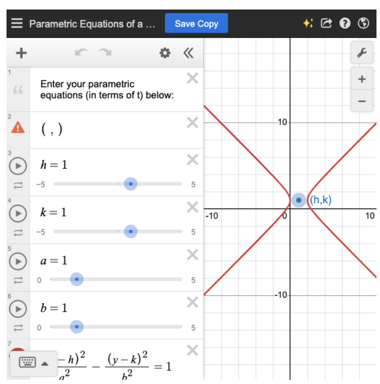
Hint. Look for one involving secant and tangent.

(b) Set the first blank element of your trig identity from the previous part equal to the following expression, and then solve for x.

$$\frac{(x-h)^2}{a^2}$$

(c) Set up a similar equation involving y and the trig function from the second blank of part 10.1.4.16.**??** then solve for y to get a general set of parametric equations for the translated hyperbola.

(d) Subsitute in your parametric equations for the translated hyperbola into the Desmos Interactive below to 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

Figure 10.1.7 A Desmos Interactive for Exercise 10.1.4.16

17. For each of the following equations, give a parameterization and be sure state the bounds of your parameterization.

(a) $\dfrac{(x-2)^2}{9} + \dfrac{(y+1)^2}{16} = 1$

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

(c) $(y-7)^2 = -4(x+1)$

10.2 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 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 compute it?
- How do we describe the motion of a projectile if the only force acting on the object is acceleration due to gravity?

In your first calculus course you looked at what was measured by the new operations of limits, derivatives, and integrals. Each of these ideas and its applications was likely introduced as a way to measure something specific by approximating the measurement and then looking at how to improve this approximation. Limits offered a powerful tool to precisely describe how these approximations would converge to the relevant measurement. This is the classic calculus approach: 1) approximate the measurement, 2) quantify how the approximation changes on a finer scale, and 3) use a limit to show how the approximate calculus converge to the measurement of interest.

In this course, we will use this structure to investigate properties of new kinds of functions, equations, and graphs. In this section, we want to investigate the meaning of limits, derivatives, and integrals on our new class of functions, vector valued functions of one variable. Additionally, we will connect these measurements to properties of curves in space as the graphs of these vector valued functions. As we expand our perspective and scope of study, you will likely see deeper meanings in the definitions of limits, derivative, and integrals from your first calculus course.

Preview Activity 10.2.1 As the only employee of *Steer Clear*, you have decided that you need to understand how the timing of position measurements will change different measurements needed for 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 **??**.

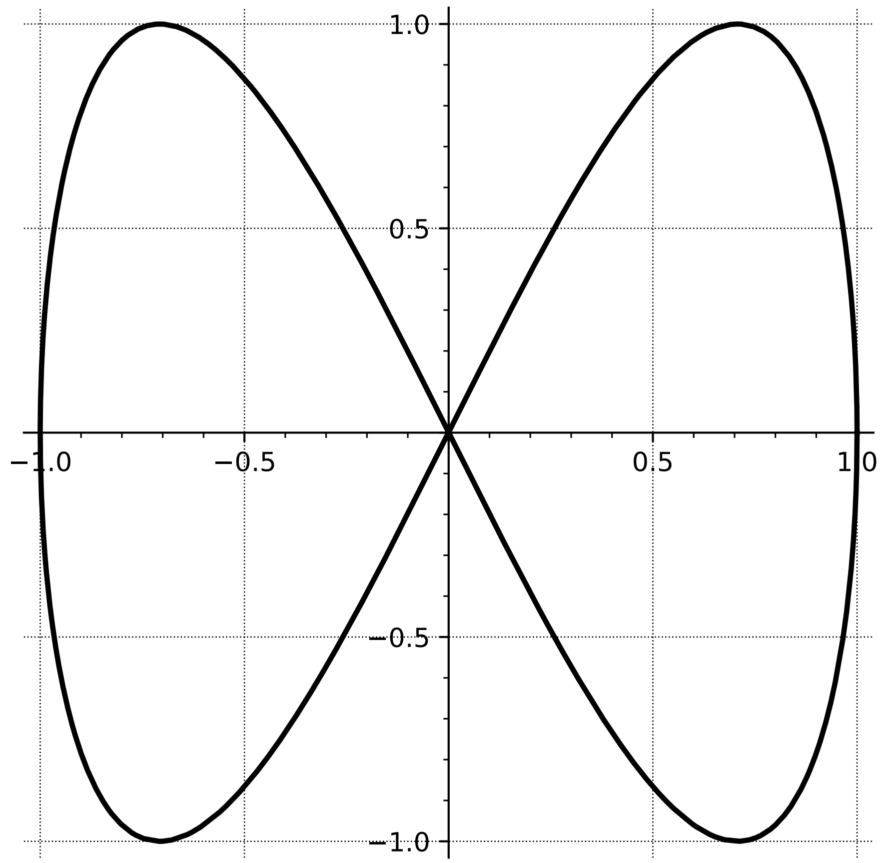


Figure 10.2.1 A plot of the figure eight curve given by $\vec{r}_8(t) = \langle \cos(t), \sin(2t) \rangle$ for $0 \leq t \leq 2\pi$

(a) In order to understand how often to have your software collect location data, you decide to look at your position for a few different times. Calculate the following and draw the output vectors of $\vec{r}_8$ on a plot of the curve. You should state your answer with components rounded to three decimal places.

(a) $\vec{r}_8(3)$

(b) $\vec{r}_8(3.1)$

(c) $\vec{r}_8(3.14)$

These correspond to the locations that would be sampled if wanted to know the location of our car at $\vec{r}_8(\pi)$ but collected data every second, every tenth of a second, and hundredth of a second (respectively).

(b) Write a couple of sentences to describe (geometrically and algebraically) what happens to the output of $\vec{r}_8(t)$ as $t \to \pi$.

(c) Calculate $\vec{r}_8(\pi)$ and $\vec{r}_8(\pi) - \vec{r}_8(3)$. Plot $\vec{r}_8(\pi)$, $\vec{r}_8(3)$, and $\vec{r}_8(\pi) - \vec{r}_8(3)$ on a graph of the figure eight curve.

(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 10.2.1.**??**, 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 input gets 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 very 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 $x_i(t)$ 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.

10.2.1 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 10.2.2 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 10.2.1 have form of the classic calculus approach described above. In particular, both definitions are of the form

$$\text{derivative} = \lim_{stepsize\to 0} \frac{\text{change in output of function}}{\text{change in input of function}}$$

Let $\vec{r}$ be the vector-valued function whose graph is shown in Figure 10.2.2, and let h be a scalar that represents a small change in time. The vector $\vec{r}(t)$ is the blue vector in Figure 10.2.2 and $\vec{r}(t+h)$ is the green vector.

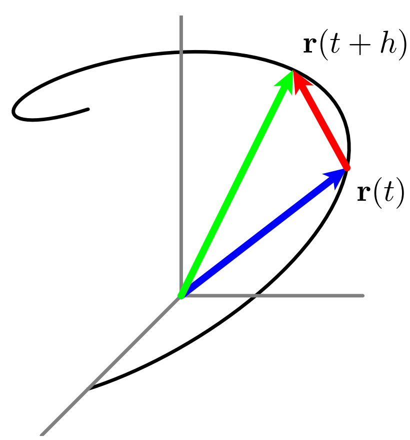


Figure 10.2.3 A single difference quotient of a vector valued function

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 $[t, t+h]$, which is also the **average velocity** of the object on this interval. Thus, 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), so $\vec{r}'(t) = \vec{v}(t)$ is the instantaneous **velocity** of the object at time t. Furthermore, we can interpret the derivative $\vec{r}'(t)$ as the direction vector of the line tangent to the graph of $\vec{r}$ at the value t.

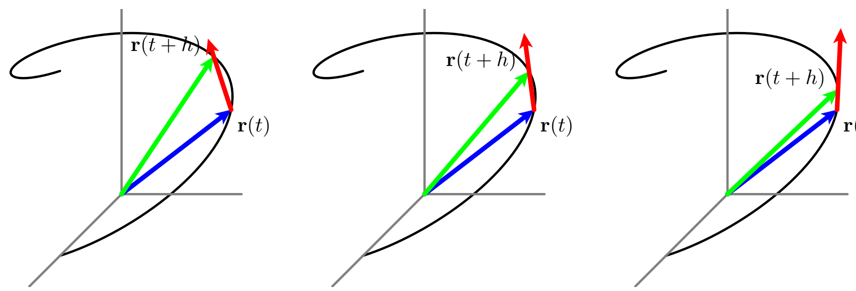


Figure 10.2.4 Snapshots of several difference quotients as h gets smaller

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 limits exists, and thus $\vec{v}'(t) = \vec{a}(t)$ is the **acceleration** of the moving object.

Both the velocity and acceleration are *vector quantities*: they have magnitude and direction. By contrast, the magnitude of the velocity vector, $\|\vec{v}(t)\|$, which is the **speed** of the object at time t, is a scalar quantity.

10.2.2 Computing Derivatives

As we learned in single variable calculus, computing derivatives using the definition is often difficult. Fortunately, properties of the limits and vectors make it straightforward to calculate the derivative of a vector-valued function similar to how we developed shortcut differentiation rules in calculus I. If we breakdown the steps in Definition 10.2.1 we notice that

1. vector subtraction $(\vec{r}(t+h) - \vec{r}(t))$
2. scalar multiplication (multiply by $\frac{1}{h}$)
3. evaluate limit as $h \to 0$

are all evaluated componentwise on our vector valued function. Thus we can apply our derivative rules on each component of $\vec{r}(t)$ to compute the derivative $\frac{d\vec{r}}{dt}$.

Key Idea 10.2.5 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 10.2.1, 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 simply differentiating its components. ■

Activity 10.2.2 For each of the following vector-valued functions, state any values of t for which the limit will *NOT* exist, then find $\vec{r}\,'(t)$.

a. $\vec{r}(t) = \langle \cos(t), t\sin(t), \ln(t) \rangle$.

b. $\vec{r}(t) = \langle t^2 + 3t, e^{-2t}, \frac{t}{t^2+1} \rangle$.

c. $\vec{r}(t) = \langle \tan(t), \cos(t^2), te^{-t} \rangle$.

d. $\vec{r}(t) = \langle \sqrt{t^4+4}, \sin(3t), \cos(4t) \rangle$.

In first-semester calculus, we developed several important differentiation rules, including the constant multiple, product, quotient, and chain rules. For instance, recall that we 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 rules for each of the kinds of multiplication we have defined involving vectors. These rules, which are easily verified, are summarized as follows.

Properties of derivatives of vector-valued functions.

Let f be a differentiable real-valued function of a real variable t and let $\vec{r}$ and $\vec{s}$ be differentiable vector-valued functions of the real variable t. Then

1. $\dfrac{d}{dt}\left[\vec{r}(t) + \vec{s}(t)\right] = \vec{r}\,'(t) + \vec{s}\,'(t)$
2. $\dfrac{d}{dt}[f(t)\vec{r}(t)] = f(t)\vec{r}\,'(t) + f'(t)\vec{r}(t)$
3. $\dfrac{d}{dt}\left[\vec{r}(t) \cdot \vec{s}(t)\right] = \vec{r}\,'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}\,'(t)$
4. $\dfrac{d}{dt}\left[\vec{r}(t) \times \vec{s}(t)\right] = \vec{r}\,'(t) \times \vec{s}(t) + \vec{r}(t) \times \vec{s}\,'(t)$
5. $\frac{d}{dt}\left[\vec{r}(f(t))\right] = f'(t)\ \vec{r}\,'(f(t))$.

When applying these properties, pay attention to whether the quantities involved as 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.

In the next several sections of this chapter, we will examine how to utilize the calculus tools above and vector measurements to calculate different properties of curves in space and the associated velocity and acceleration vectors that come from their parameterizations.

I think the elements of the next activity are done other places in developing the splitting of acceleration. I think the calculation portions can be folded into the tangent line activity.

Activity 10.2.3 The left side of Figure 10.2.4 shows the curve described by the vector-valued function $\vec{r}$ defined by

$$\vec{r}(t) = \left\langle 2t - \frac{1}{2}t^2 + 1, t - 1 \right\rangle.$$

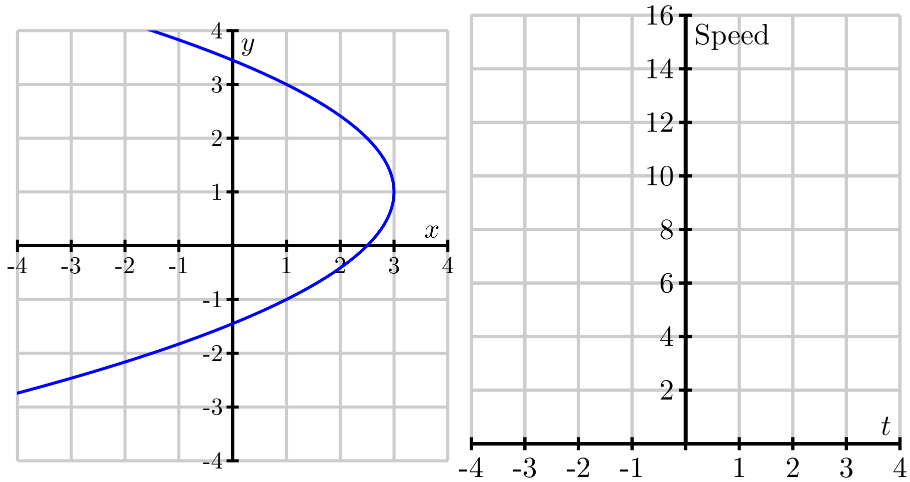


Figure 10.2.6 The curve $\vec{r}(t) = \left\langle 2t - \frac{1}{2}t^2 + 1, t - 1 \right\rangle$ and its speed.

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 10.2.4 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 10.2.4. 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

$$\frac{d}{dt}|\vec{v}(t)| = \text{comp}_{\vec{v}}\vec{a}.$$

10.2.3 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 curve generated by a differentiable function looks very much like a line. Indeed, when we zoom in sufficiently far on a particular point, the curve looks indistinguishable from its tangent line. You can look at Figure 9.5.1 at the beginning of Section 9.5 for an interactive visualization of this idea.

In the same way, we expect that a smooth curve in 3-space will be locally linear. In Activity 10.2.5, we investigate how to find the tangent line to such a curve. Recall from our work in Section 9.5 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

$$\vec{L}(t) = \vec{L}_0 + t\vec{u}.$$

In parametric form, the line $\vec{L}$ is given by

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

Activity 10.2.4

(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 end of $\vec{r}(\pi)$.

(b) 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.

We see that our work in Activity 10.2.5 can be generalized. Given a differentiable vector-valued function $\vec{r}$, the tangent line to the curve at the input value a is given by

$$\vec{L}(t) = \vec{r}(a) + t\vec{r}'(a). \qquad (10.2.1)$$

Here we see that because the tangent line is determined entirely by a given point and direction, the point is provided by the function $\vec{r}$, evaluated at $t = a$, while the direction is provided by the derivative, $\vec{r}'$, again 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 (10.2.1) 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

$$\vec{L}(t) = \vec{r}(a) + (t-a)\vec{r}'(a). \tag{10.2.2}$$

(For example, in Equation (10.2.1), $\vec{L}(0) = \vec{r}(a)$, so the line's parameterization "starts" at $t = 0$. When we write the parameterization in the form of Equation (10.2.2), $\vec{L}(a) = \vec{r}(a)$, so the line's parameterization "starts" at $t = a$.)

As we will learn in Chapter 11, a smooth surface in 3-space is also locally linear. That means that the surface will look like a plane, which we call its *tangent plane*, as we zoom in on the graph. It is possible to use tangent lines to traces of the surface to generate a formula for the tangent plane; see Exercise 10.2.6.15 for more details.

10.2.4 Integrating a Vector-Valued Function

Recall from your calculus of scalar valued functions, that the definite integral of a 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 some point in $[x_{i-1}, x_i]$. The definite integral measure the (signed) area between the graph of $f(x)$ and the x-axis between $x = a$ and $x = b$. The definite integral follows our classic calculus approach to measure the accumulation of the output of f over the interval $[a, b]$. Most students never used this Riemann sum definition to calculate any definite integrals, rather they used the Fundamental Therom of Calculus and the associated antiderivatives to efficiently calculate the value of a definite integral. `https://activecalculus.org/single/sec-4-4-FTC.html#FTC` states that if f is a continuous function on $[a, b]$ and $F(x)$ is an antiderivatve of f ($\frac{dF}{dx} = f(x)$), then

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

The set of functions that are the antiderivative 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 (apply integration componentwise). Thus we

Integrating a vector-valued function.

If $\vec{r}(t) = x(t)\hat{\imath} + y(t)\hat{\jmath} + z(t)\hat{k}$, then

$$\int \vec{r}(t)\ dt = \left(\int x(t)\ dt\right)\hat{\imath} + \left(\int y(t)\ dt\right)\hat{\jmath} + \left(\int z(t)\ dt\right)\hat{k}.$$

Similarly, the **indefinite integral** $\int \vec{r}(t)\ dt$ of a vector-valuedfunction $\vec{r}$ is the general antiderivative of $\vec{r}$ and represents the collection of all antiderivatives of $\vec{r}$ and is denoted with

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

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 area under the curve idea does NOT generalize when we think about the integral of $\vec{r}(t)$. The integral of $\vec{r}$ is in relation to the variable t and there is not a t-axis on the graph of $\vec{r}(t)$. need to think this explanation through more...

The concept that the definite integral measures the accumulation of the output of your function will generalize to the vector valued function realm. Recall that if $v(t)$ was the velocity of an object traveling on the x-axis, then we could describe the position of the object as

$$x(t) = \int_0^t v(u)\ du + x_0$$

This was the result of setting up a Riemann sum that looked at the change in position over a small step of time as the velocity on that step times the size of the time step. In other words,

$$x(t) \approx x_0 + \sum_{i=0} nv(t_i^*)\Delta t_i$$

Applying this accumulation of rates of change or velocity ideas componentwise means that if we are able to separate the velocity of our motion into components, then we can integrate (componentwise!) to get the position in terms of each coordinate 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 us to solve for the particular integration constant (the $+C$).

Example 10.2.7 For this example, we will look at 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 or to military applications like mortar or artillery placements, as well 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. In our example, we will try 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:

- the acceleration is given by $m\vec{a}(t) = -mg\hat{\jmath}$ (the only force is gravity pulling down with a constant force)
- the initial position is $\vec{r}(0) = \langle x_0, y_0\rangle$
- the initial velocity is $\vec{v}(0) = \frac{\vec{r}}{dt}(0) = \langle v_0\cos(\theta), v_0\sin(\theta)\rangle$

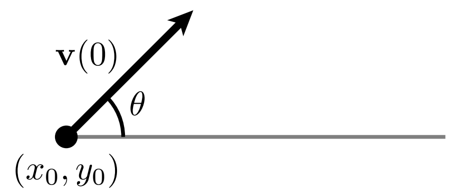


Figure 10.2.8 Projectile motion

We can simplify $m\vec{a}(t) = -mg\hat{\jmath}$ to get $\vec{a}(t) = \langle 0, -g\rangle$, then integrate acceleration to get 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 some constant vector. We need to pick the constant such that the initial velocity is $\vec{v}(0)\langle v_0\cos(\theta), v_0\sin(\theta)\rangle$, so we will have

$$\vec{v}(t) = \langle 0, -gt\rangle + \langle v_0\cos(\theta), v_0\sin(\theta)\rangle = \langle v_0\cos(\theta), v_0\sin(\theta) - gt\rangle$$

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

$$\langle v_0\cos(\theta)t, v_0\sin(\theta)t - \frac{1}{2}gt^2\rangle + \vec{C}$$

Again we need to pick the integration constant, $\vec{C}$, that satisifes our 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 will look at relationships between position, velocity, and acceleration in a situation that is more complicated than the constant acceleration of Example **??**. The process of using integration and differentiation will remain the same but the results will not be nearly as simple.

Activity 10.2.5 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 10.2.6. 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 10.2.6.

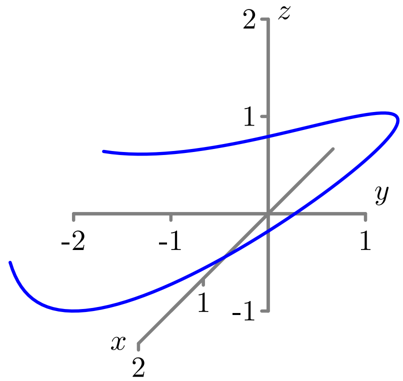


Figure 10.2.9 The position graph for the function in Activity 10.2.6

(d) Give the vector equation for the tangent line, $\vec{L}(t)$, that is tangent to the position curve at $t = 1$.

The situation in the previous activity shows a case where the motion of our object seems to be driven by forces with different behaviors 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 different measurements related to motion along curves like this. Key to these measurements will be the combined use of our new calculus tools for vector valued functions with the vector measurements from the previous chapter.

10.2.5 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.

10.2.6 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. If $\mathbf{r}(t) = \cos(-2t)\mathbf{i} + \sin(-2t)\mathbf{j} + 7t\mathbf{k}$, compute:
A. The velocity vector $\mathbf{v}(t) =$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k}$
B. The acceleration vector $\mathbf{a}(t) =$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k}$
Note: the coefficients in your answers must be entered in the form of expressions in the variable t; e.g. "5 cos(2t)"

2. Given that the acceleration vector is $\mathbf{a}(t) = (-16\cos(-4t))\mathbf{i} + (-16\sin(-4t))\mathbf{j} + (-2t)\mathbf{k}$, 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 $\mathbf{v}(t) =$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k}$
B. The position vector $\mathbf{r}(t) =$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k}$
Note: the coefficients in your answers must be entered in the form of expressions in the variable \emph{t}; e.g. "5 cos(2t)"

3. Evaluate
$\int_0^7 (t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k})dt =$ __________ $\mathbf{i}+$ __________ $\mathbf{j}+$ __________ $\mathbf{k}$.

4. Find parametric equations for line that is tangent to the curve $x = \cos t,\ y = \sin t,\ z = t$ at the point
$(\cos(\frac{3\pi}{6}), \sin(\frac{3\pi}{6}), \frac{3\pi}{6})$.
Parametrize the line so that it passes through the given point at t=0. All three answers are required for credit.
$x(t) =$ ______________________________
$y(t) =$ ______________________________
$z(t) =$ ______________________________

5. If $\mathbf{r}(t) = \cos(4t)\mathbf{i} + \sin(4t)\mathbf{j} - 2t\mathbf{k}$
compute $\mathbf{r}'(t)=$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k}$
and $\int \mathbf{r}(t)\,dt=$ ______________ $\mathbf{i}+$ ______________ $\mathbf{j}+$ ______________ $\mathbf{k} + \mathbf{C}$
with $\mathbf{C}$ a constant vector.

6. For the given position vectors $\mathbf{r}(t)$,
compute the (tangent) velocity vector $\mathbf{r}'(t)$ for the given value of t .
A) Let $\mathbf{r}(t) = (\cos 5t, \sin 5t)$.
Then $\mathbf{r}'(\frac{\pi}{4})=$ (______________ , ______________)?
B) Let $\mathbf{r}(t) = (t^2, t^3)$.
Then $\mathbf{r}'(1)=$ (______________ , ______________)?
C) Let $\mathbf{r}(t) = e^{5t}\mathbf{i} + e^{-t}\mathbf{j} + t\mathbf{k}$.
Then $\mathbf{r}'(3)=$ __________ $\mathbf{i}+$ __________ $\mathbf{j}+$ __________ $\mathbf{k}$?

7. Suppose $\vec{r}(t) = \cos(\pi t)\,\boldsymbol{i} + \sin(\pi t)\,\boldsymbol{j} + 2t\boldsymbol{k}$ 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 8?
$t =$ ____________________
(b) What is the velocity of the particle when its height is 8?
$\vec{v} =$ ____
(c) When the particle has height 8, 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.
$L(t) =$ ____

8. Suppose the displacement of a particle in motion at time t is given by the parametric equations
$x(t) = (3t-1)^2, \quad y(t) = 4, \quad z(t) = 243t^4 - 108t^3.$
(a) Find the speed of the particle when $t = 3$.
Speed = ____
(b) Find t when the particle is stationary.
$t =$ ____________________

9. Find the derivative of the vector function
$\mathbf{r}(t) = t\mathbf{a} \times (\mathbf{b} + t\mathbf{c})$, where
$\mathbf{a} = \langle -2, -5, -5\rangle$, $\mathbf{b} = \langle -1, 5, -4\rangle$, and $\mathbf{c} = \langle -5, 5, 2\rangle$.
$\mathbf{r}'(t) = \langle$ _________________________, _________________________, _________________________ $\rangle$

10. Let $\mathbf{c}_1(t) = (e^{-t}, \sin(2t), -t^3)$, and $\mathbf{c}_2(t) = (e^{-4t}, \cos(5t), 2t^3)$
$\frac{d}{dt}[\mathbf{c}_1(t) \cdot \mathbf{c}_2(t)] =$ ____
$\frac{d}{dt}[\mathbf{c}_1(t) \times \mathbf{c}_2(t)] =$ ____$\mathbf{i}+$
____$\mathbf{j}+$
____$\mathbf{k}$

11. A gun has a muzzle speed of 80 meters per second. What angle of elevation should be used to hit an object 170 meters away? Neglect air resistance and use $g = 9.8\,\mathrm{m/sec}^2$ as the acceleration of gravity.
Answer: ____ radians

12. 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

$$x = 25\cos t, \quad y = 25\sin t, \quad z = 120 - 5t.$$

(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 = ____ $\mathrm{ft/min}^2$

13. 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 10.2.3, 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. $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle \times \langle -\sin(t), \cos(t), 1\rangle$

14. 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.

15. In this exercise, we determine the equation of a plane tangent to the surface defined by $f(x, y) = \sqrt{x^2 + y^2}$ at the point $(3, 4, 5)$.

a. Find a parameterization for the $x = 3$ trace of f. What is a direction vector for the line tangent to this trace at the point $(3, 4, 5)$?

b. Find a parameterization for the $y = 4$ trace of f. What is a direction vector for the line tangent to this trace at the point $(3, 4, 5)$?

c. The direction vectors in parts (a) and (b) form a plane containing the point $(3, 4, 5)$. What is a normal vector for this plane?

d. Use your work in parts (a), (b), and (c) to determine an equation for the tangent plane. Then, use appropriate technology to draw the graph of f and the plane you determined on the same set of axes. What do you observe? (We will discuss tangent planes in more detail in Chapter 10.)

16. 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 $\vec{r}$. Is the resulting quantity a scalar or a vector? What does it measure?

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. $\vec{r}(t) = \left\langle \frac{t}{1+t^2}, te^{t^2}, \frac{1}{1+t^2} \right\rangle$

17. In this exercise, we develop the formula for the position function of a projectile that has been launched at an initial speed of $|\vec{v}_0|$ and a launch angle of θ. Recall that $\vec{a}(t) = \langle 0, -g \rangle$ is the constant acceleration of the projectile at any time t.

a. Find all velocity vectors for the given acceleration vector $\vec{a}$. When you anti-differentiate, remember that there is an arbitrary constant that arises in each component.

b. Use the given information about initial speed and launch angle to find $\vec{v}_0$, the initial velocity of the projectile. You will want to write the vector in terms of its components, which will involve $\sin(\theta)$ and $\cos(\theta)$.

c. Next, find the specific velocity vector function $\vec{v}$ for the projectile. That is, combine your work in (a) and (b) in order to determine expressions in terms of $|\vec{v}_0|$ and θ for the constants that arose when integrating.

d. Find all possible position vectors for the velocity vector $\vec{v}(t)$ you determined in (c).

e. Let $\vec{r}(t)$ denote the position vector function for the given projectile. Use the fact that the object is fired from the position (x_0, y_0) to show it follows that

$$\vec{r}(t) = \left\langle |\vec{v}_0| \cos(\theta)t + x_0, -\frac{g}{2}t^2 + |\vec{v}_0| \sin(\theta)t + y_0 \right\rangle.$$

18. 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 10.2.8.

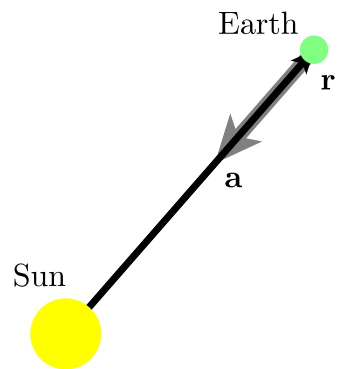


Figure 10.2.10 A central force.

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 $\vec{L}$.

10.3 Arc Length

Motivating Questions

- How can a definite integral be used to measure the length of a curve in 2- or 3-space?
- Why is arc length useful as a parameter?

In the previous sections we have introduced the notation and graphs for vector valued functions, as well as how the calculus operations of limits, derivatives, and integrals will apply to these functions. In this section we will start our work on utilizing these calculus elements alongside our vector tools (like magnitude, direction, or projections) to measure important properties and find ways to efficiently calculate how these properties change along our curve. The first property we will investigate comes up very naturally when looking at curves in space: "What is the length of a given curve in space?" Remember that the distance between two points in space is given by

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

The length of a curve is different than the displacement (change in position) as we will see in the Preview Activity.

Preview Activity 10.3.1 In Preview Activity 10.1.1 and Preview Activity 10.2.1 we saw how the location tracking system's (LTS)output is related to a vector valued function that corresponds to how our 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, you drive your car up a parking ramp and note that the LTS recorded the car had coordinates of $\langle 10, -5, 0\rangle$ at $t = 0$ and $\langle -8, 10, 12\rangle$ when $t = 30$. What is the distance between initial ($t = 0$) and final ($t = 30$)positions of the car?

(b) Do you think the distance the car actually traveled is greater than, less than, or the same as your answer to part **??.??**? 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 $t = 0$
- $\langle 6, 0, 6\rangle$ when $t = 10$
- $\langle -10, 2, 10\rangle$ when $t = 20$
- $\langle -8, 10, 12\rangle$ when $t = 30$

Calculate the distance between successive times and estimate how far your car went in on the drive up the parking ramp.

(d) Do you think your estimate from part **??.??** 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 to output the location data. You decide that getting the location every half a second seems like a good idea. How many data points will that correspond to for your drive up the parking ramp? Write a few 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.

The last task is an example of a vital skill mathematics, scientific, and engineering fields called algorithmic thinking. Algorithmic thinking involves breaking larger problems into different tasks where a precise set of rules or operations can be stated for each task. We have referred to our classic calculus approach in a few scenarios (and will again this section) and we will take a moment here to mention how the classic calculus approach has the structure of this algorithmic thinking. Specifically, the tasks for our classic calculus approach can be summarized as as 1) approximate, 2) quantify how approximation changes with scale or refinement, 3) apply limit to describe how approximation converges to the measurement of interest. The details of how each of these tasks is done will change depending on the measurement, which is why so much time will be spent talking about the details of tasks 1 and 2 when introducing new concepts.

10.3.1 Arc Length

The central question we want to answer in this section is "Given, C, a curve from P to Q, what the length of C?"

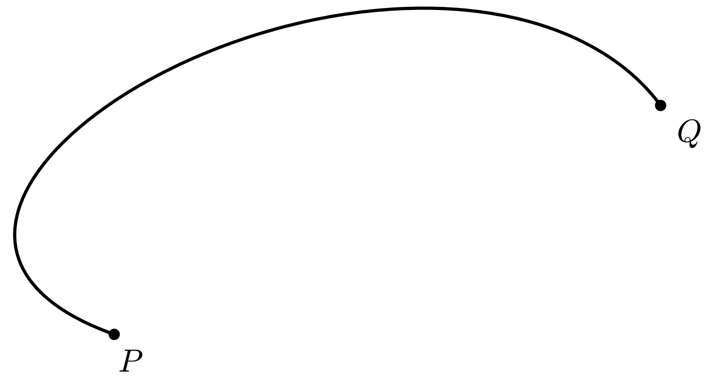


Figure 10.3.1 A curve C oriented from the point P to the point Q

As suggested by our work in the Preview Activity and the classic calculus approach, we will approximate the length of the curve C by computing the length of a collection of line segments between points on C. In particular, we will call these intermediate points $\vec{r}_0, \vec{r}_1, ...\vec{r}_n$. Remember that we are specifying location using vectors to be graphed in standard position. Using vectors to specify position means that the corresponding line segments between successive vr_i positions can be represented by the vector difference $\Delta\vec{r}_i = \vec{r}_i - \vec{r}_{i-1}$.

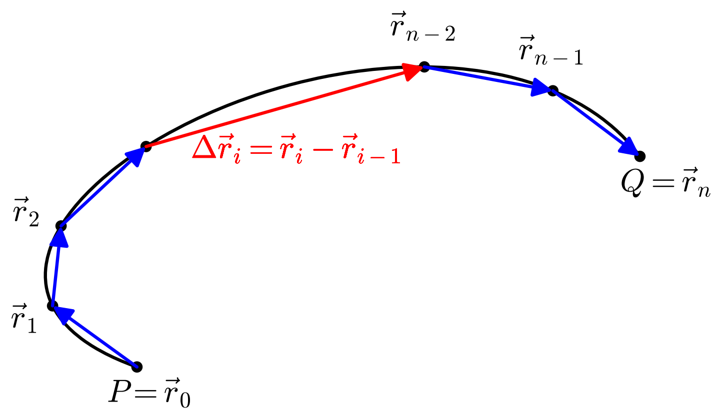


Figure 10.3.2 A curve C oriented from the point P to the point Q split into n parts with endpoints $\vec{r}_0, \vec{r}_1, ...\vec{r}_n$

Notice that the description above does not involve a parameterization of C or a description for how you are traveling along C. In general, it is easiest to generate the positions $\vec{r}_i$ by evaluating a parameterization at a different t-values. In particular, we will use $\vec{r}(t) = \langle x(t), y(t), z(t)$, a parameterization of C to specify the points between P and Q. Specifically, if $\vec{r}(t)$ is a parameterization of C with $P = \vec{r}(a)$ and $Q = \vec{r}(b)$, then we 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 which will not correspond to parts with equal length.

Our approximation of the length of C will be the sum of the lengths of the vectors $\Delta\vec{r}_i = \vec{r}(t_i) - \vec{r}(t_{i-1})$. The length or magnitude of $\Delta\vec{r}_i$ will be given by the distance formula

$$\sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2 + (z(t_i) - z(t_{i-1}))^2}$$

To make our notation a little more compact, we will write $x(t_i) - x(t_{i-1})$ as Δx_i, $y(t_i) - y(t_{i-1})$ as Δy_i, and $z(t_i) - z(t_{i-1})$ as Δz_i. Our approximation of the length of the curve will be calculated by the sum of these lengths which gives

$$\begin{aligned} L &\approx \sum_{i=1}^{n} \|\Delta \vec{r}_i\| \\ &\approx \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2} \end{aligned}$$

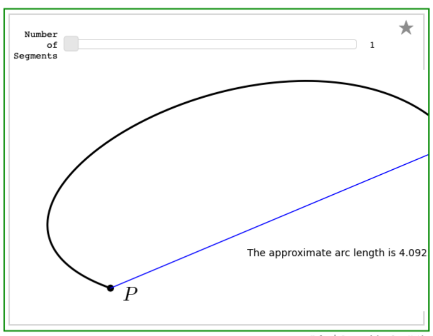
Since this approximation will approach the actual length as we use smaller and smaller steps, we define the arclength as follows:

Definition 10.3.3 The arclength of C, a curve in space is given by

$$\lim_{n\to\infty} \sum_{i=1} n\sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2}$$

◊

You can use the slider on the interact to change the number of segments to be used in the approximation of the arclength 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 (≈ 6.3286).



Standalone
Embed

Figure 10.3.4 A plot of curve in space

Since we will want to shrink the step size Δt to get better approximations we will rewrite our approximation to allow us to see how our approximation will change as we look at smaller scales (smaller Δt).

$$\begin{aligned} L &\approx \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2}\frac{\Delta t}{\Delta t} \\ &\approx \sum_{i=1}^{n} \sqrt{(\frac{\Delta x_i}{\Delta t})^2 + (\frac{\Delta y_i}{\Delta t})^2 + (\frac{\Delta z_i}{\Delta t})^2}\ \Delta t \end{aligned}$$

This accomplishes the second task in our classic calculus approach (quantify how the approximation changes over a refinement) and formats our approximation as a Riemann sum. In particular, our approximation is a Riemann sum of the function $\sqrt{(\frac{\Delta x_i}{\Delta t})^2 + (\frac{\Delta y_i}{\Delta t})^2 + (\frac{\Delta z_i}{\Delta t})^2}$ which will simplify greatly when we look at this function as Δt goes to 0. Looking at how the limit will apply to each term in the square root separately, we can see

$$
\begin{aligned}
x'(t) &= \lim_{\Delta t \to 0} \frac{x(t_i) - x(t_{i-1})}{\Delta t}, \\
y'(t) &= \lim_{\Delta t \to 0} \frac{y(t_i) - y(t_{i-1})}{\Delta t}, \quad \text{and} \\
z'(t) &= \lim_{\Delta t \to 0} \frac{z(t_i) - z(t_{i-1})}{\Delta t},
\end{aligned}
$$

and the function we are evaluating in our Riemann sum becomes

$$\sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}$$

as $\Delta t \to 0$ or $n \to \infty$. Thus the limit of the Riemann sum from our definition of arc length **??** gives a definite integral

$$
\begin{aligned}
\lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2 + (\Delta z_i)^2} &= \sum_{i=1}^{n} \sqrt{(\frac{\Delta x_i}{\Delta t})^2 + (\frac{\Delta y_i}{\Delta t})^2 + (\frac{\Delta z_i}{\Delta t})^2}\ \Delta t \\
&= \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}\ dt
\end{aligned}
$$

The following theorem summarizes how our definition of arc length can be evaluated as a definite integral using a parameterization.

Theorem 10.3.5 The length of a curve. *If $\vec{r}(t)$ defines a smooth curve C on an interval $[a, b]$, then the* length *L of C is given by*

$$L = \int_a^b \|\vec{r}\,'(t)\|\, dt. \tag{10.3.1}$$

Note that formula (10.3.1) 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 (10.3.1) says that we simply integrate the speed of an object traveling over the curve to find the distance traveled by the object, which is the same as the length of the curve, just as in one-variable calculus.

Activity 10.3.2 In this activity, we will use parameterizations to find the length of a couple of common curves.

(a) Parameterize a circle of radius R centered at the origin and be sure to give bounds on your parameter.

(b) Use your parameterization from the previous task in the definite integral of Theorem 10.3.1 to calculate the circumference of a circle of radius R.

(c) Find the exact length of the spiral defined by $\vec{r}(t) = \langle R\cos(t), R\sin(t), t\rangle$ 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.

Exercise 10.3.5.6 will apply Definition **??** and Theorem 10.3.1 to get the length of a graph given by $y = f(x)$, which you may have seen in previous calculus courses.

10.3.2 Traveling with Unit Speed

In Activity 10.3.2, you looked at parameterizations of a circle and a helix that traveled with constant speed. This made the calculation of arc length in each case much easier and since there is not one parameterization for a curve, you may be wondering if it is possible to create a parameterization for each curve that moves with constant speed. The answer is YES but in practice this parameterization might be difficult to write out algebraically.

Before we dig into the details of this constant speed parameterization, we will make our lives even easier by trying find a parameterization that describes traveling along our curve with unit speed; In other words, we want to find a parameterization of our curve, $\vec{r}(t)$, that has speed one for all time, $\|\vec{r}'(t)\| = 1$. If you are always moving with speed one, then the distance traveled IS the time elapsed. 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 (the units are different...) Remember that we do not have to walk in a straight line (constant direction) to keep our speed constant.

Example 10.3.6 In this example, we look a few simple paths and how to write a parameterization that moves with unit speed. For our first path, we will consider the linear path given by starting at the point (x_0, y_0, z_0) when $t = a$ and going to the point (x_1, y_1, z_1) when $t = b$. This line will be described by the parameterization $\vec{r}(t) = \langle x_0 + \frac{t-a}{b-a}(x_1 - x_0), y_0 + \frac{t-a}{b-a}(y_1 - y_0), z_0 + \frac{t-a}{b-a}(z_1 - z_0)\rangle$. This parameterization will have speed

$$\|\vec{r}'(t)\| = \sqrt{(\frac{x_1 - x_0}{b-a})^2 + (\frac{y_1 - y_0}{b-a})^2 + (\frac{z_1 - z_0}{b-a})^2}$$

Note that the speed of this parameterization is constant, so for simplicity we will call this constant speed v_{line}.

In order to change this parameterization to move with speed 1, we will need to travel v_{line} times slower, which can be accomplished by replacing t in our parameterization with $\frac{t}{v_{line}}$. In other words, we adjusted the parameter used to make sure the speed with respect to new parameter will be one. This adjustment was easy to do because the amount we needed to change the parameter did not change at different times.

Algebraically, we can characterize this change by thinking of $\vec{r}(u)$ as the “old” parameterization with parameter u and we are looking to replace u with some function $\alpha(t)$ so that $\|\vec{r}'(t)\| = 1$ for all t. In our linear path case, our choice of $\alpha(t) = \frac{t}{v_{line}}$ meant that applying the chain rule to $\|\vec{r}'(t)\|$ gives

$$\|\vec{r}(\alpha(t))\| = \left\|\frac{d\vec{r}}{du}\right\| \left|\frac{d\alpha}{dt}\right| = v_{line}\frac{1}{v_{line}} = 1$$

□

The left part of the equation at the end of the previous example shows the big idea needed to find our unit speed parameterization; Specifically, we will need to pick our $\alpha(t)$ transformation to satisfy

$$\left\|\frac{d\vec{r}}{du}\right\| \left|\frac{d\alpha}{dt}\right| = 1$$

and this is algebraically ***VERY*** difficult to write out in closed form but will be possible because the arclength traveled as a function of your parameter cannot decrease. If we graph the parameter versus arclength traveled, we get a graph that passes the horizontal line test and is thus invertible. An example of this kind of plot is shown in Figure **??**.

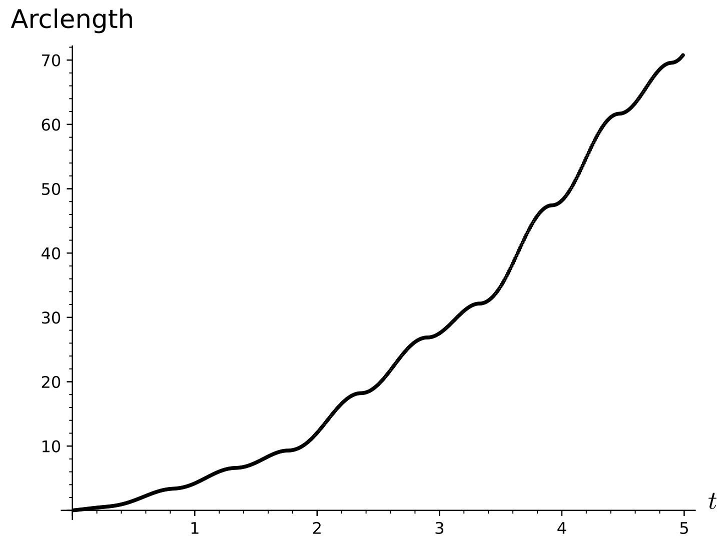


Figure 10.3.7 A plot of the arclength traveled as a function of the parameter value to that point for $\vec{r}(t) = \langle 3t\cos(4t)\sin(3-2t), t-\cos(2+t)\rangle$

In the next example, we will work through the details of finding this transformation of a parameterization with variable speed into one that has unit speed.

Example 10.3.8 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 give a parameterization of C that has unit speed. Our first step is to find the speed of our parameterization, so

$$\begin{aligned}
\|\vec{r}'(t)\| &= \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} \\
&= \sqrt{(2t)^2 + (4t^{1/2})^2 + (4)^2} \\
&= \sqrt{4t^2 + 16t + 16} \\
&= \sqrt{(2t+4)^2} \\
&= 2t + 4
\end{aligned}$$

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 $t > 0$ we only need to consider the positive branch of the square root calculation.

Next, we will calculate $s(x)$, the arclength 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 can use Theorem 10.3.1 to get the following for our arclength calculation:

$$\begin{aligned}
s(a) &= \int_0^a \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}\, dt \\
&= 2\int_0^a t + 2\, dt \\
&= \left(t^2 + 4t\right)\Big|_0^a
\end{aligned}$$

$$= a^2 + 4a$$

So the arclength traveled from $t = 0$ to $t = a$ along the curve C is $a^2 + 4a$. It is important to remember at this point that we want to transform our parameterization to have unit speed, which will mean that the arclength and the parameter will be the same value as we move along the curve. This means that we want a function that takes an input of arclength along the curve C and outputs the time at which our original parameter reaches that location. In other words, we want the inverse function of $s(a) = a^2 + 4a$ because $s(a)$ gives the arclength that corresponds to the parameter value a. Since $t \geq 0$, we can solve the equation $s = a^2 + 4a$ (or $a^2 + 4a - s = 0$) for a to obtain $a = \frac{-4+\sqrt{16+4s}}{2} = -2 + \sqrt{4+s}$. Note again that by restricting to positive parameter values, we only need to consider one branch of our solutions.

Let's slow down and make sure we understand our two functions and their composition. The original parameterization $\vec{r}(t)$ has inputs of parameter values and outputs corresponding to locations on our curve C. Our function $a(s)$ has inputs that are arclengths (specifically, the arclength traveled from $t = 0$ to $t = a$ along the curve C) and outputs corresponding to the parameter values (of the original parameterization).

We can now do a composition of 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 idea that the "time elapsed = arclength traveled", which is exactly $\vec{r}(a(s))$ because the parameter being used is the amount of arclength being traveled. The composition $\vec{r}(a(s))$ will have an input of arclength traveled as it's parameter and the "inner" function, $a(s)$, will output the parameter value (of the original parameterization) where that arclength is achieved. Thus $\vec{r}(a(s))$ will output the location on C where an arclength of s will be achieved.

You may be a bit overwhelmed by the previous paragraphs conceptually, so let's go back to the algebra of our problem. For our example, the composition $\vec{r}(a(s))$ will give the following as the unit speed parameterization.

$$\vec{r}_{new}(s) = \left\langle \left(-2+\sqrt{4+s}\right)^2, \frac{8}{3}\left(-2+\sqrt{4+s}\right)^{3/2}, 4\left(-2+\sqrt{4+s}\right) \right\rangle$$

If we are diligent and persistent with our calculations, we can show that $\|\vec{r}'_{new}\| = 1$.

$$\begin{aligned}
\|\vec{r}'_{new}\| =& \| \left\langle 2(-2+\sqrt{4+s})\frac{1/2}{\sqrt{4+s}} \right. \\
& \left. , \frac{8}{3}\left(\frac{3}{2}\left(-2+\sqrt{4+s}\right)^{1/2}\left(\frac{1/2}{\sqrt{4+s}}\right)\right), 4\frac{1/2}{\sqrt{4+s}} \right\rangle \\
=& \left\| \frac{1}{\sqrt{4+s}} \left\langle -2+\sqrt{4+s}, 2\sqrt{-2+\sqrt{4+s}}, 2 \right\rangle \right\| \\
=& \sqrt{\frac{1}{\sqrt{4+s}}\left(\left(-2+\sqrt{4+s}\right)^2 + \left(2\sqrt{-2+\sqrt{4+s}}\right)^2 + (2)^2\right)} \\
=& \sqrt{\frac{1}{\sqrt{4+s}}\left((-2+\sqrt{4+s})+2\right)^2} \\
=& \sqrt{\frac{4+s}{4+s}} = 1
\end{aligned}$$

The crucial step in our work (algebraically and conceptually) was the inverse function of $s(a)$ which was relatively simple for this example. In general, this can be very difficult to write out explicitly. □

The example above shows how to convert a parameterization to one that has unit speed. Key to this was understanding that distance along the curve is not the same as the parameter being used (in general). If we consider the curve defined by the parabola $y = x^2/2$ with $x \in [0, 2]$, then we can parameterize this curve by $\vec{r}(t) = \langle t, t^2/2 \rangle$ for $t \in [0, 2]$. We can plot equally spaced points in our parameter and equally spaced points in arclength to see th

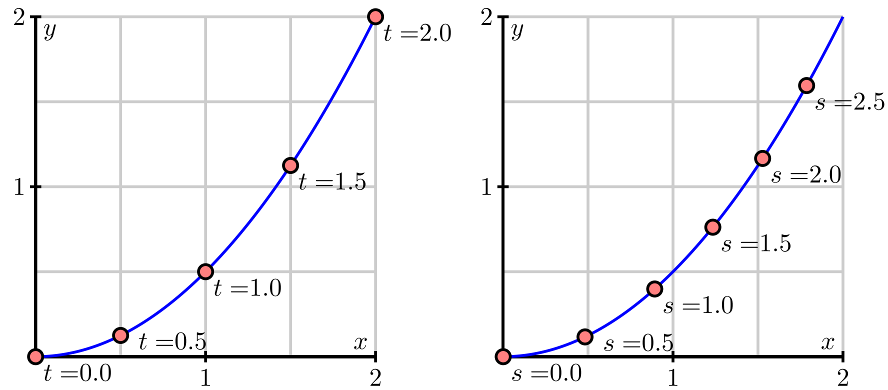


Figure 10.3.9 A plot of $y = x^2/2$ with $x \in [0, 2]$ with points equally spaced in the parameter value (on left) and in arclength (on right)

In many ways, a parameterization with unit speed is a more 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, thus creating a function $\vec{r}$. 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 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.

These examples illustrate a general method. Of course, evaluating an arc length integral and finding a formula for the inverse of a function can be difficult, so while this process is theoretically possible, it is not always 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.

Many other textbooks will refer to the unit speed parameterization of a curve as the arclength parameterization. Conceptually, we derived the value of this parameterization from wanting to move with constant speed of 1 which is why we have opted to use the "unit speed parameterization" name here. After seeing the motivation for wanting to move with unit speed, you can see the correspondence between the arclength and the parameter necessary to give this new description.

10.3.3 The Driver or The Road?

In the next activity, we will look at several measurements as we consider different drivers completing one lap around the track seen in Figure 10.3.5.

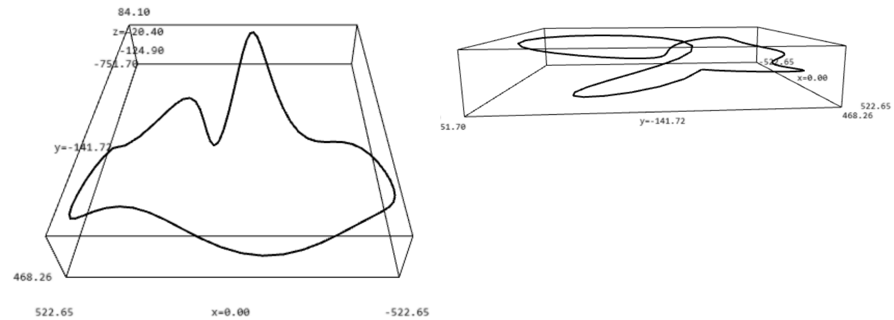


Figure 10.3.10 A plot of the racetrack with scale in meters

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 10.3.3 Is it a property of the driver or the road? In this activity, we want to determine if the different measurements that are described are a property of the driver or the road. 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) 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) Now that we are warmed up, let's look at some more interesting measurements.

Speedometer Reading: 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 1. The vector calculus quantities we have covered are time, position, velocity, acceleration, speed, arc length, unit tangent, unit normal, binormal, curvature, tangential acceleration component, normal acceleration component, and osculating circle.

Hint 2. This is an easy one, I hope.

(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.

10.3.4 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 our parameter is the same as the arclength traveled to that point.

10.3.5 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

$$x = 3 + 2t, \quad y = 2 - 2t, z = 4 - 4t,$$

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. Consider the curve $\mathbf{r} = (e^{4t}\cos(-3t), e^{4t}\sin(-3t), e^{4t})$.

Compute the arclength function $s(t)$: (with initial point $t = 0$).

3. Find the length of the given curve:

$$\mathbf{r}\,(t) = (5t, 5\sin t, 5\cos t)$$

where $-3 \leq t \leq 4$.

4. Consider the path $\mathbf{r}(t) = (4t, 2t^2, 2\ln t)$ defined for $t > 0$.

Find the length of the curve between the points $(4, 2, 0)$ and $(16, 32, 2\ln(4))$.

5. Starting from the point $(-3, -3, -1)$, reparametrize the curve

$\mathbf{x}(t) = (-3 - 2t, -3 + 2t, -1 - 2t)$ in terms of arclength.

$\mathbf{y}(s) = ($ ____________, ____________, ____________ $)$

6. 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 (10.3.1) to show that the length of the curve defined by

f on an interval $[a, b]$ is

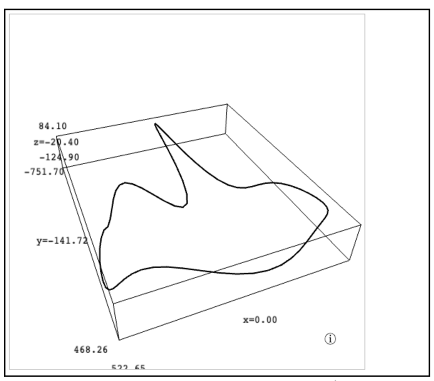
$$\int_a^b \sqrt{1 + [f'(t)]^2}\, dt.$$

10.4 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?

In Section 10.1, 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
Embed

Figure 10.4.1 A plot of our racetrack

For example, consider the curve in space that we want a race track as shown in Figure 10.4.1. Not every driver will travel along this racetrack in the same way, but every driver has to go through these same points in space (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 times. More generally, some properties you 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 10.4.1 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 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 10.4.2 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.

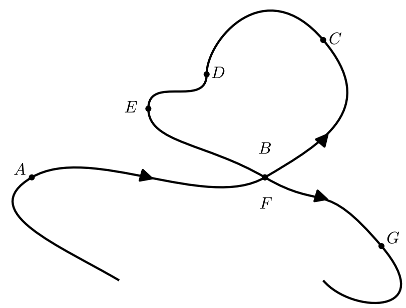


Figure 10.4.2 The path taken on a country road to collect data with your LTS

(a) At each of the labeled points on the curve, draw a vector in the direction of travel. Write a sentence about how you 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.

This preview activity 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 give precise ways of measuring the direction of travel and the direction of turning at a particular point on the curve.

10.4.1 Direction of travel

Based on our work in Subsection 10.2.1, 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 a unit vector), we define the **direction of travel** as follows.

Definition 10.4.3 The **unit tangent vector** to a curve is denoted by $\vec{T}(t)$ is computed by

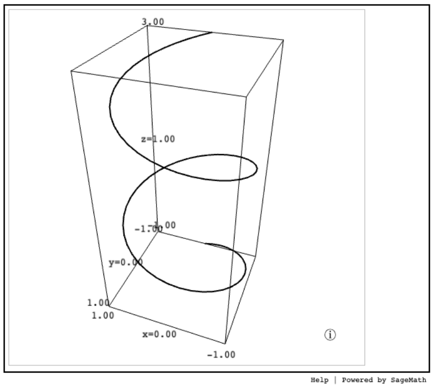
$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|}$$

and measures the direction of travel along a parameterized curve. ◊

From the definition, the unit tangent 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 ($\vec{v}$ does not exist), the direction of travel at that instant does not make sense because there is not a consistent way to define motion.

The definition of $\vec{T}$ allows us to separate the velocity into its magnitude (speed) and direction (unit tangent). Remember that $\vec{T}$ changes along the curve and is calculated as a function of the parameter.

Example 10.4.4 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 $\vec{r}(t) = \langle \cos(t), \sin(t), \frac{t}{\pi} \rangle$.



Standalone
Embed

Figure 10.4.5 A plot of our racetrack

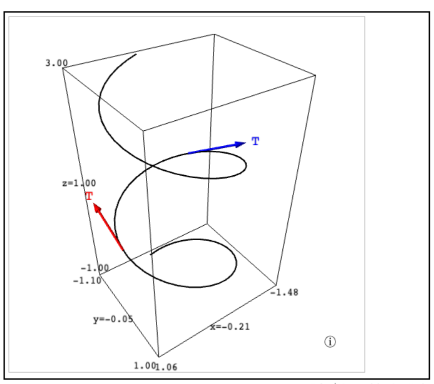
Based on our parameterization, we have velocity vector given by $\vec{v}(t) = \langle -\sin(t), \cos(t), \frac{1}{\pi} \rangle$, which has speed given by

$$\text{speed}(t) = \|\vec{v}(t)\| = \sqrt{(-\sin(t))^2 + (\cos(t))^2 + (\frac{1}{\pi})^2} = \sqrt{1 + \frac{1}{\pi^2}}$$

One of the useful and unique ideas for this example is that we have a constant speed. No matter what parameter value we look at, the speed is always $\sqrt{1 + \frac{1}{\pi^2}}$. Using Definition 10.4.3, we can see that

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1 + \frac{1}{\pi^2}}} \langle -\sin(t), \cos(t), \frac{1}{\pi} \rangle$$

If we look at $t = \frac{\pi}{2}$, then $\vec{T}(\frac{\pi}{2}) = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle -1, 0, \frac{1}{\pi} \rangle$. If $t = \frac{5\pi}{4}$, then $\vec{T}(\frac{5\pi}{4}) = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{\pi} \rangle$. In Figure 10.4.6, $\vec{T}(\frac{\pi}{2})$ 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
Embed

Figure 10.4.6 A plot of our racetrack

Note that we can also find $\vec{T}(\frac{\pi}{2})$ by taking $\vec{v}(\frac{\pi}{2}) = \langle -1, 0, \frac{1}{\pi} \rangle$ and shrinking $\vec{v}$ to be a unit vector. Because our definitions and theorems work for all t-values, we can apply them at a specific t-value, which has the added advantage that computations tend to be manageable to be done by hand. This will only useful for finding information at a fixed t-value because you are not concerned with how the measurement is done for other parameter values. □

Activity 10.4.2

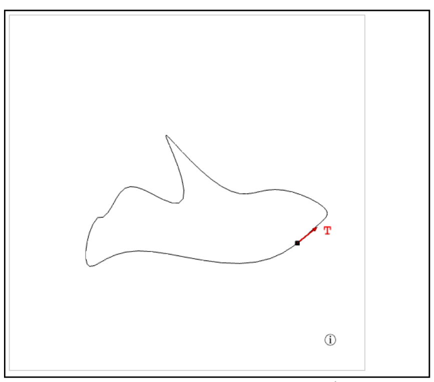
(a) Find $\vec{v}$, speed, and $\vec{T}$ for a line given by the parameterization $\vec{r}(t) = \langle 3t-1, 2-2t, 5+t \rangle$. Write a few sentences about why your results make sense and do not change for different parameter values.

(b) Find $\vec{v}$, speed, and $\vec{T}$ for the curve given by the parameterization $\vec{r}(t) = \langle t, t^2, t^3 \rangle$.

(c) Find $\vec{T}$ at the point P if the tangent line to the curve through P is given by $\langle 2-7t, 3t+1, -4t-1 \rangle$.

10.4.2 The Direction of Turning

Recalling our analogy of our parameterization describing how someone drives along a racetrack, $\vec{T}$ measures the direction that that a car is heading at any given instant. If we look at $t = 2$, we see that $\vec{T}(2) \approx \langle -0.8813, -0.4691, 0.0568 \rangle$. This is shown in Figure 10.4.7. 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 relative direction to the car. So how can we measure the direction that the car is turning at this instant, in terms of the locations given by the parameterization in a xyz-coordinate system?



Standalone
Embed

Figure 10.4.7 A plot of our racetrack with $\vec{T}(2)$ plotted in red

Let's start looking at $\vec{T}' = \frac{d\vec{T}}{dt}$, which will measure how quickly $\vec{T}$ is changing in terms of the parameter. The vector $\vec{T}' = \frac{d\vec{T}}{dt}$ will measure the change in the direction of travel.

Remember that vectors have magnitude and direction, so the derivative of a vector valued function of 1 variable will measure the change in the magnitude *and* change in the direction. Since $\vec{T}$ will always have length 1 (when $\vec{T}$ exists), *all* of the change in $\vec{T}$ will correspond to a change in direction. So the direction of $\vec{T}' = \frac{d\vec{T}}{dt}$ will be exactly what we are looking for; namely, $\vec{T}' = \frac{d\vec{T}}{dt}$ will be in the direction of turning. We don't care (for now) about how fast the car is turning, just what direction that turning is in. So we define the unit normal vector, denoted $\vec{N}$, as the unit vector in the direction of $\vec{T}' = \frac{d\vec{T}}{dt}$.

Definition 10.4.8 The unit normal vector to a curve is denoted by $\vec{N}(t)$ is computed by

$$\vec{N}(t) = \frac{\vec{T}'(t)}{\|\vec{T}'(t)\|}$$

The unit normal vector measures the direction of turning along a parameterized curve. ♢

Activity 10.4.3 Based on the definition of the unit normal, there are a few different characteristics a curve can have that would cause $\vec{N}$ to not exist. We will look at three different cases:

1. $\vec{N}$ will not exist if $\vec{T}$ does not exist
2. $\vec{N}$ will not exist if $\vec{T}' = \frac{d\vec{T}}{dt}$ does not exist
3. $\vec{N}$ will not exist if $\vec{T}' = \frac{d\vec{T}}{dt} = \vec{0}$

(a) Draw an example of a curve such that $\vec{T}$ does not exist at a point on your curve. Explain why $\vec{T}$ does not exist based on your plot and explain why $\vec{N}$ does not exist at the same point.

(b) Draw an example of a curve such that $\vec{T}' = \frac{d\vec{T}}{dt}$ does not exist at a point on your curve. Explain why $\vec{T}' = \frac{d\vec{T}}{dt}$ does not exist based on your plot and explain why $\vec{N}$ does not exist at the same point.

(c) Draw an example of a curve such that $\vec{T}' = \frac{d\vec{T}}{dt} = \vec{0}$ at a point on your curve. Explain why $\vec{T}' = \frac{d\vec{T}}{dt} = \vec{0}$ based on your plot and explain why $\vec{N}$ does not exist at the same point.

Example 10.4.9 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 10.4.4 and was plotted in Figure 10.4.5.

From our work in Example 10.4.4, we calculated

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \left\langle -\sin(t), \cos(t), \frac{1}{\pi} \right\rangle$$

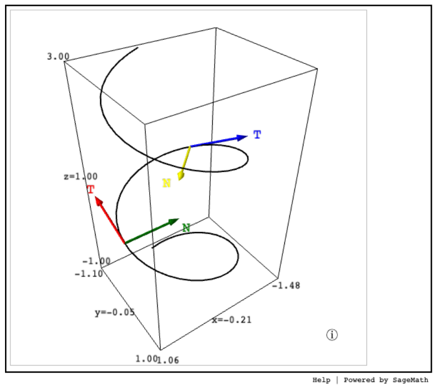
Using the definition of $\vec{N}$, we need to compute the derivative of $\vec{T}$ with respect to t.

$$\vec{T}'(t) = \frac{d\vec{T}}{dt} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle -\cos(t), -\sin(t), 0 \rangle$$

Note that the derivative of $\vec{T}$ is easily calculated for all time but this will rarely be the case. Further, making a unit vector in the direction of $\vec{T}(t)' = \frac{d\vec{T}}{dt}$ will be fairly easy and give us the following calculation:

$$\begin{aligned}
\vec{N}(t) &= \frac{\vec{T}'(t)}{\|\vec{T}'(t)\|} \\
&= \frac{\frac{1}{\sqrt{1+\frac{1}{\pi^2}}} \langle -\cos(t), -\sin(t), 0 \rangle}{\frac{1}{\sqrt{1+\frac{1}{\pi^2}}}} \\
&= \langle -\cos(t), -\sin(t), 0 \rangle
\end{aligned}$$

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 , we can see that the direction of turning (as measured by $\vec{N}$) is pointing toward the center of the helix. This should make sense the the direction you are turning is toward the inside of the circular trace. Note that the Unit Tangent vector has a vertical component because 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. The z component of $\vec{T}$ is unchanging so the vertical component of $\vec{T}'$ will be zero.



Standalone
Embed

Figure 10.4.10 A plot of the helix with $\vec{T}$ and $\vec{N}$ plotted for a couple points

This example 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 10.4.4 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 $\vec{r}(t) = \langle t, t^2, t^2\rangle$.

(a) Given the parameterization, $\vec{r}(t) = \langle t, t^2, t^2\rangle$, calculate $\vec{v}(t)$ and speed(t).

(b) Now calculate $\vec{T}(t)$ for the $\vec{r}(t) = \langle t, t^2, t^2\rangle$ path.

(c) We are going to take it slow and calculate the first component of $\vec{T}' = \frac{d\vec{T}}{dt}$. Remember that the derivative with respect to t is computed component-wise.

(d) Now calculate the second component of $\vec{T}' = \frac{d\vec{T}}{dt}$.

(e) Now calculate the third component of $\vec{T}' = \frac{d\vec{T}}{dt}$.

(f) Put together your work for the three components of $\vec{T}' = \frac{d\vec{T}}{dt}$ and compute $\left\|\frac{d\vec{T}}{dt}\right\|$.

(g) You likely gave up on the previous task or ran out of paper to write out $\frac{d\vec{T}}{dt}$ and $\left\|\frac{d\vec{T}}{dt}\right\|$. Instead of more punishing algebra, describe how you would calculate $\vec{N}$ if you had $\frac{d\vec{T}}{dt}$ and $\left\|\frac{d\vec{T}}{dt}\right\|$.

As talked about in Subsection 10.4.1, 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 10.4.10, we saw that the unit tangent and unit normal vectors were orthogonal to each other (when measured at the same point).

Question 10.4.11 Can $\vec{N}$ have any part parallel to $\vec{T}$ or must $\vec{T}$ and $\vec{N}$ be orthogonal? □

We will answer this question here with a conceptual, geometric argument. By Definition 10.4.8, 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 will include change in the magnitude and in the direction of $\vec{T}$. Since the length of $\vec{T}$ does not change, then all of $\vec{T}' = \frac{d\vec{T}}{dt}$ will be a change in the direction of $\vec{T}$ and not at all in the direction of $\vec{T}$. This is an intuitive argument for why $\vec{T}$ and $\vec{N}$ are orthogonal.

We can also give a more rigorous algebraic argument, but this argument does not provide insight into the measurements or relationship between $\vec{T}$ and $\vec{N}$. For this argument, we will look at the derivative of $\|\vec{T}\|^2$. Because the length of $\vec{T}$ is constant, then $\frac{d}{dt}\left(\|\vec{T}\|^2\right) = 0$. To simplify $\frac{d}{dt}\left(\|\vec{T}\|^2\right)$, we will use properties of the dot product related to the magnitude, the derivative product rule (across the dot product), then the commutative property of the dot product, respectively.

$$\begin{aligned}\frac{d}{dt}\left(\|\vec{T}\|^2\right) &= \frac{d}{dt}\left(\vec{T}\cdot\vec{T}\right)\\ &= \vec{T}'\cdot\vec{T} + \vec{T}\cdot\vec{T}'\\ &= 2\vec{T}\cdot\vec{T}'\end{aligned}$$

If $2\vec{T}\cdot\vec{T}' = 0$, then $\vec{T}$ and $\vec{T}'$ are orthogonal. Since $\vec{N}$ is parallel to $\vec{T}'$, then $\vec{T}$ and $\vec{N}$ are orthogonal.

10.4.3 The Binormal Vector

Definition 10.4.12 The **Binormal vector**, $\vec{B}$, 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. The Binormal vector will be a unit vector because $\|\vec{T}\times\vec{N}\| = \|\vec{T}\|\|\vec{N}\|\sin(\theta)$ 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, then $\|\vec{T}\|\|\vec{N}\|\sin(\theta) = 1$. The binormal vector, $\vec{B}$, will give the axis of rotation for the motion along the curve. Note that 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.

This TNB frame is sometimes called the Frenet frame and is very useful in environments like aviation and space travel because so much of the important flight information needs be in terms of relative measurements to the craft and it's motion.

Example 10.4.13 In this example, we will complete our calculation of the TNB frame for the helix given by $\vec{r}(t) = \langle \cos(t), \sin(t)\frac{t}{\pi} \rangle$. From Example 10.4.4 and Example 10.4.9 we have the following:

$$\vec{T}(t) = \frac{\vec{v}(t)}{\|\vec{v}(t)\|} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}}\langle -\sin(t), \cos(t), \frac{1}{\pi}\rangle$$

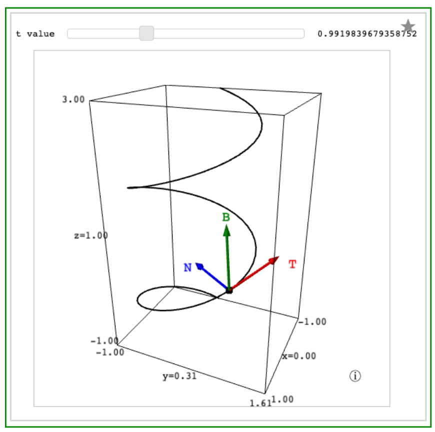
$$\vec{N} = \langle -\cos(t), -\sin(t), 0\rangle$$

Using the definition for $\vec{B}$, we get

$$\vec{B} = \vec{T}\times\vec{N} = \frac{1}{\sqrt{1+\frac{1}{\pi^2}}}\langle -\frac{1}{\pi}\sin(t), -\frac{1}{\pi}\cos(t), \sin(t)^2+\cos(t)^2\rangle$$

Factoring out $\frac{1}{\pi}$ from each component and simplifying the coefficient gives

$$\vec{B} = \frac{1}{\sqrt{\pi^2+1}}\langle -\sin(t), -\cos(t), \pi\rangle$$



Standalone
Embed

Figure 10.4.14 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 with the relative motion of the object. □

10.4.4 The Driver or the Road: TNB Frame

In this subsection, we will talk about whether the measurements of $\vec{T}$, $\vec{N}$, and $\vec{B}$ will be properties of the driver or the road. A parameterized curve also 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 the different measurements at the same location on the curve. 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 parametertized 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 10.3.5.

Activity 10.4.5

(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) *Direction of Turning*: The direction of turning is measured by checking to see if the steering wheel is turned left or 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 a property of the driver or the road? Be sure to explain your answer.

(c) Is $\vec{B}$ a property of the driver or the road? Justify your ideas.

10.4.5 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

10.4.6 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 unit tangent vector at the indicated point of the vector function

$$\mathbf{r}(t) = e^{7t}\cos t\,\mathbf{i} + e^{7t}\sin t\,\mathbf{j} + e^{7t}\,\mathbf{k}$$

$\mathbf{T}(\pi/2) = \langle$ ____________, ____________, ____________ $\rangle$

2. Consider the vector function

$\mathbf{r}(t) = \langle t, t^5, t^6\rangle$

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 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.

4. Consider the standard helix parameterized by $\vec{r}(t) = \cos(t)\hat{\imath} + \sin(t)\hat{\jmath} + t\hat{k}$.

a. Recall that the unit tangent vector, $\vec{T}(t)$, is the vector tangent to the curve at time t that points in the direction of motion and has length 1. Find $\vec{T}(t)$.

b. Explain why the fact that $|\vec{T}(t)| = 1$ implies that $\vec{T}$ and $\vec{T}'$ are orthogonal vectors for every value of t. (Hint: note that $\vec{T} \cdot \vec{T} = |\vec{T}|^2 = 1$, and compute $\frac{d}{dt}[\vec{T} \cdot \vec{T}]$.)

c. For the given function $\vec{r}$ with unit tangent vector $\vec{T}(t)$ (from (a)), determine $\vec{N}(t) = \frac{1}{|\vec{T}'(t)|}\vec{T}'(t)$.

d. What geometric properties does $\vec{N}(t)$ have? That is, how long is this vector, and how is it situated in comparison to $\vec{T}(t)$?

e. Let $\vec{B}(t) = \vec{T}(t) \times \vec{N}(t)$, and compute $\vec{B}(t)$ in terms of your results in (a) and (c).

f. What geometric properties does $\vec{B}(t)$ have? That is, how long is this vector, and how is it situated in comparison to $\vec{T}(t)$ and $\vec{N}(t)$?

g. Sketch a plot of the given helix, and compute and sketch $\vec{T}(\pi/2)$, $\vec{N}(\pi/2)$, and $\vec{B}(\pi/2)$.

10.5 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?

In Section 10.4 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 the "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.

Note here that given a parameterization of a curve in space, the speed will give us an idea about how fast we are moving along the curve, but we don't have a way to measure how much turning is happening at a given location. In this section, we will measure how quickly a curve is turning. This is a different from measuring how fast an object is turning as it moves along a path. We will return to this idea a bit later.

Preview Activity 10.5.1 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 company, 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 describe the paths your car should be taking). Since everyone on the road drives differently, we would like to measure how quickly a stretch of road is turning in different places. This measurement should not depend on the particular driver but should be a property of the path the road takes. Remember that your car cannot look forward in time to see how it will need to turn, but the car can look forward along the road to see how the path changes.

(a) Figure **??** 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. Draw a vector in the direction of travel at P_0 and P_1 and use these vectors to describe how the direction of travel is changing on the path from P_0 to P_1.

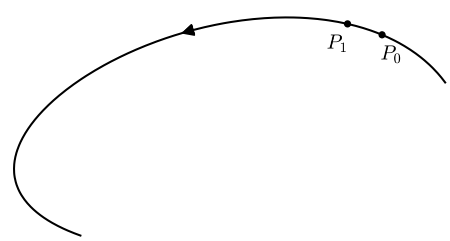


Figure 10.5.1 A plot of the section of road with points P_0 and P_1 separated by 10 meters

(b) Figure **??** 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. Draw a vector in the direction of travel at Q_0 and Q_1 and use these vectors to describe how the direction of travel is changing on the path from Q_0 to Q_1.

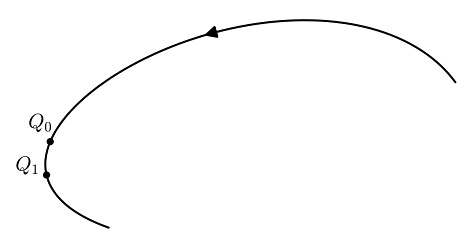


Figure 10.5.2 A plot of the section of road with points Q_0 and Q_1 separated by 10 meters

(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.

(d) If we let P_2 be the point on the path that is 5 meters ahead of P_0 and Q_2 be the point on the path that is 5 meters ahead of Q_0. Would your answer to the previous question change if you compared the path from P_0 to P_2 to the path from Q_0 to Q_2? Explain your reasoning.

Our preview activity showed how looking at the change in $\vec{T}$, the direction of travel, for small steps (in terms of arclength) along our curve will allow us to measure the rate at which a curve is turning. Note that this will be different than measuring how quickly an object that is traveling along the path needs to turn in order to stay on the path. The arguments in the preview activity do not depend on how an object is traveling along the path.

10.5.1 Curvature

We claim that $\vec{r}_1(t)$, the unit speed parameterization of a curve is unique; If you and I drive the same path and both maintain a speed of exactly 1 meter per second, then for any point on the path, we will both reach that point in the same amount of time. For instance,it will take us both exactly 15 seconds to reach the point that is 15 meters down the path (remember that the parameter value for $\vec{r}_1$ is the same as the distance traveled along the curve.) Conceptually, you can see how using the unit speed parameterization will allow us to measure properties of the curve and avoid questions about how different ways of moving along the path will affect our measurements.

We will measure how quickly a path turns as the rate of change in the direction of travel in terms of the unit speed parameterization. If $\vec{r}_1(t)$ is the *unit* speed parameterization of our curve C and we look a section of our path from $\vec{r}(t_0)$ to $\vec{r}(t_0 + \Delta h)$, then rate of change of $\vec{T}$ will be given by

$$\left\| \frac{\vec{T}(t_0 + \Delta h) - \vec{T}(t_0)}{\Delta h} \right\|$$

Note here that this rate of change is a scalar (not a vector) and that Δh measures how long the section we are looking at is in terms of arclength.

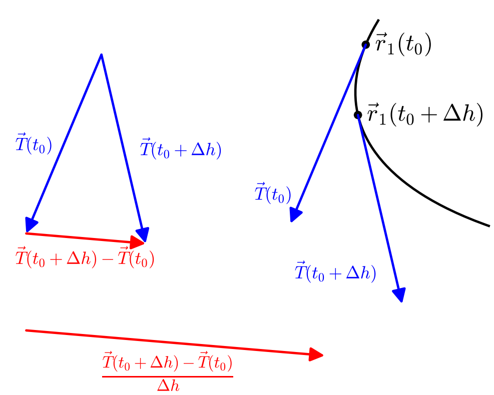


Figure 10.5.3 A plot of a curve with $\vec{T}$ shown for two points

Figure **??** shows a geometric representation of our measurement of how quickly our curve turns over the interval from $\vec{r}_1(t_0)$ to $\vec{r}_1(t_0 + \Delta h)$. If we take the limit as $\Delta h \to 0$, then we have measure of how quickly a curve is turning at the point $P = \vec{r}_1(t_0)$, which can be calculated by

$$\lim_{\Delta h \to 0} \left\| \frac{\vec{T}(t_0 + \Delta h) - \vec{T}(t_0)}{\Delta h} \right\| = \left\| \frac{d\vec{T}}{ds} \right\|$$

Remember that the parameterization used in every part of the above argument is the unit speed parameterization which is why are able to take the derivative of $\vec{T}$ with respect to the arclength.

Definition 10.5.4 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 for the unit speed parameterization C will measure that arclength traveled so far along the curve. Note that curvature is denoted withκ, the Greek lowercase letter "kappa". $\diamondsuit$

Activity 10.5.2 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 **??** we found that the unit speed parameterization of a line through the point (x_0, y_0) with direction vector $\langle a, b \rangle$ can be given by

$$\vec{r}(s) = \langle x(s), y(s) \rangle = \langle x_0 + \frac{a}{\sqrt{a^2 + b^2}} s, y_0 + \frac{b}{\sqrt{a^2 + b^2}} s \rangle$$

Compute the speed of this parameterization to verify that this has unit speed. This will also show that $\vec{r}'(s) = \vec{v}(s) = \vec{T}(s)$ (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 the unit speed parameterization of a line.

(c) Write a few sentences to explain why your result for the calculation of $\kappa = \left\| \frac{d\vec{T}}{ds} \right\|$ of a line makes sense.

(d) Recall that in Activity 10.3.2 we found that the unit speed parameterization of circle in 2-space of radius a centered at the origin can be given by

$$\vec{r}(s) = \left\langle a \cos\left(\frac{s}{a}\right), a \sin\left(\frac{s}{a}\right) \right\rangle$$

Compute the speed of this parameterization to verify that this has unit speed. This will also show that $\vec{r}'(s) = \vec{v}(s) = \vec{T}(s)$ (which is not true in general.)

(e) 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 centered at the origin.

(f) 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.

The definition of curvature relies on our ability to parameterize curves with unit speed. Since finding unit speed parametrization is very difficult to compute directly, we would like to be able to calculate the curvature of a curve in terms of *any* parametrization $\vec{r}(t)$.

Theorem 10.5.5 *If $\vec{r}$ is a parameterization of a curve C, and if $\vec{r}'(t)$ is not zero and if $\vec{r}''(t)$ exists, then the curvature κ of C can be calculated as*

$$\kappa = \kappa(t) = \frac{\left\|\vec{T}'(t)\right\|}{|\vec{r}'(t)|}$$

Further, we can efficiently calculate curvature in terms of the velocity and acceleration functions by

$$\kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3} \tag{10.5.1}$$

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 our curve C. Remember that in order to calculate the curvature, $\vec{T}$ will need to have an input corresponding to the parameter of the unit speed parameterization, which means the input of the $\vec{T}$ will need to measure the arclength of C to the point we at. We will use a function $s(t)$ with input being the parameter of $\vec{r}(t)$ and the output of $s(t)$ will be the arclength on C to the point $\vec{r}(t)$. Note that $s(t)$ will be computed using Theorem 10.3.1 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 will utilize 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 our idea above from Theorem 10.3.1, we can substitute $\frac{ds}{dt} = \|\vec{r}'(t)\|$ and take the magnitude of each side to get the following

$$\left\|\frac{\vec{T}'}{\|\vec{r}'(t)\|}\right\| = \left\|\frac{d\vec{T}}{ds}\right\| = \kappa \tag{10.5.2}$$

If this chain rule argument seems a little technical to you, there is a nice conceptual interpretation of the equation above. Specifically, if we want to convert a measurement like curvature or the derivative of $\vec{T}$ between the unit speed parameterization and any other parameterization, we literally need to adjust by the speed. (**??**) 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 variable speed.

We omit the proof of (**??**) because the argument involves a lot of technical algebraic ideas that do not offer a lot of insight into what is measured by curvature. Exercise 10.5.3.6 will guide you through this calculation if you are interested in these ideas. ■

Before we do our example and activity that calculate curvature for a few other examples, we will take a moment to use (**??**) to make sense of when curvature will not be defined or make sense. From algebraic perspective, the curvature will not be defined when the speed is zero. This should make sense by the fact that if you are not moving along the curve (speed is zero), then it does not make sense to measure how quickly the curve is turning. In Section 10.5, we will make sense of why $\|\vec{v}(t) \times \vec{a}(t)\|$ being zero means that there will be zero curvature (which is very different than the curvature being undefined.)

Example 10.5.6 In this example, we will calculate the curvature of the parabola given by $y = x^2$ for $-2 < x < 2$. We will parameterize this curve in three dimensions so that we can use (**??**) (cross product is defined in 3D only). Let $\vec{r}(t) = \langle t, t^2, 0\rangle$ with $t \in [-2, 2]$.

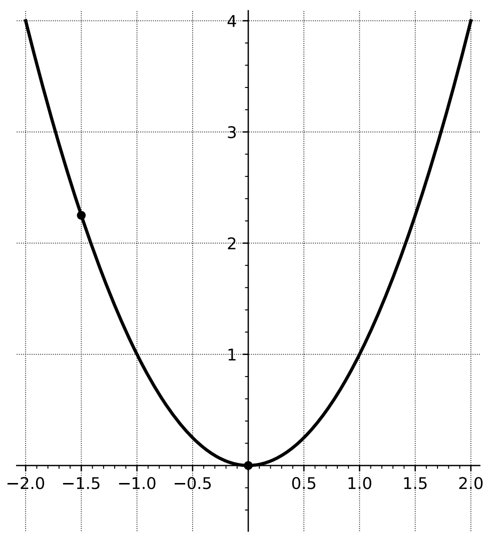


Figure 10.5.7 A plot of $y = x^2$ with points $(0, 0)$ and $(-\frac{3}{2}, \frac{9}{4})$

Using this parameterization, we get the following set of results:

$$\begin{aligned} \vec{r}'(t) &= \vec{v}(t) = \langle 1, 2t, 0\rangle \\ \vec{r}''(t) &= \vec{a}(t) = \langle 0, 2, 0\rangle \\ \vec{v}(t) \times \vec{a}(t) &= \langle 0, 0, 2\rangle \\ \|\vec{v}(t) \times \vec{a}(t)\| &= 2 \\ \|\vec{v}(t)\| &= \sqrt{1 = 4t^2} \end{aligned}$$

Using (**??**), we get the curvature to be

$$\kappa(t) = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3} = \frac{2}{(1 + 4t^2)^{3/2}}$$

Comparing the points $(0, 0)$ and $(-\frac{3}{2}, \frac{9}{4})$ on Figure **??**, you can see that the parabola will have a larger curvature at $(0, 0)$ because the path is bending much more at $(0, 0)$ than at $(-\frac{3}{2}, \frac{9}{4})$. Using our result above, we can see that the curvature at $(0, 0) \Rightarrow t = 0$ is $\kappa(0) = 2$ and the curvature at $(-\frac{3}{2}, \frac{9}{4}) \Rightarrow t = -\frac{3}{2}$ is $\kappa(\frac{3}{2}) = \frac{2}{(5)^{3/2}} \approx 0.1789$. This calculation of the curvature will match up with our geometric meaning.

We can also think of the curvature as a function of the parameter we used (in our case $t = x$.) A plot of $\kappa(t) = \frac{2}{(1+4t^2)^{3/2}}$ as a function of x is shown in Figure **??**. You can see that the curvature is highest near the vertex of the parabola and the curvature gets closer to zero as you look farther up the parabola. This should make sense because the parabola is “flatter” away from the origin.

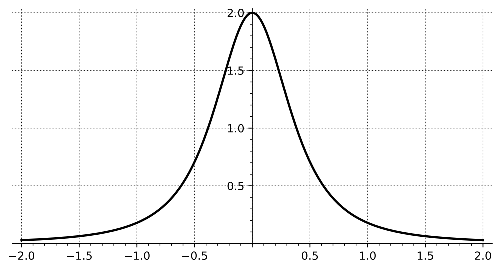
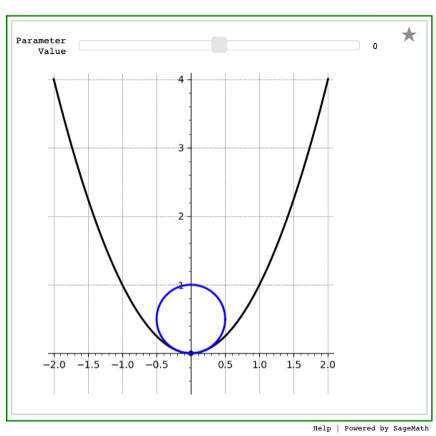


Figure 10.5.8 A plot of the curvature of $y = x^2$ as a function of x

Suppose you were driving along the path given by $y = x^2$ at unit speed and your steering wheel locked up and prevented you from changing how you were turning at the instant $t = 0$. Because the steering wheel is held fixed, you will no longer stay on the path given by $y = x^2$ but instead will drive in a circle. What do you think the radius of the circle will be?

It turns out that the radius of this circle will be $\frac{1}{\kappa(0)} = \frac{1}{2}$! This shouldn't be too surprising because in Activity 10.6.2, we saw that the curvature of a circle of radius R is $\frac{1}{R}$. In general, the **radius of curvature** is $\frac{1}{\kappa}$ and the circle that is tangent to the curve and has 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 tangent point since the osculating circle goes through the tangent point and has the same curvature.

In Figure **??**, you can see a plot of the path with the osculating circle drawn at a particular point. You can use the slider to change the parameter value at which the osculating circle is drawn. You should notice how the radius of the circle becomes much larger as you move farther away from the vertex in either direction. Remember that curvature and the radius of curvature have an reciprical relationship; large curvature means small radius and small curvature means a large radius.



Standalone
Embed

Figure 10.5.9 A plot of $y = x^2$ with the osculating circle drawn tangent to the curve

□

Activity 10.5.3 In this activity, we will calculate the curvature for an ellipse and a helix, then make sense of these results in terms of their graphs.

(a) The ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ has parameterization

$$\vec{r}(t) = \langle a\cos(t), b\sin(t)\rangle.$$

Find the curvature of the ellipse using this parameterization.

(b) Plot the ellipse given by $\frac{x^2}{4} + \frac{y^2}{1} = 1$. Use your result from the previous task (with $a = 2$ and $b = 1$) to compare the curvature at the extreme points on this ellipse. Write a few sentences that describes which points on the ellipse with have the largest curvature (in terms of both your plot and your calculation). Additionally, describe which points on the ellipse with have the smallest curvature (in terms of both your plot and your calculation).

(c) The standard helix has parameterization $\vec{r}(t) = \cos(t)\hat{\imath} + \sin(t)\hat{\jmath} + t\hat{k}$. Find the curvature of the helix using this parameterization.

(d) 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.

Activity 10.5.4 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 10.3.5. 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). The explanations for the next tasks may be difficult for you to write but will be very helpful in ensuring you understand the vector calculus concepts of this chapter.

(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.

10.5.2 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:

$$\kappa = \frac{\|\vec{r}'(t) \times \vec{r}''(t)\|}{\|\vec{r}'(t)\|^3} = \frac{\|\vec{v}(t) \times \vec{a}(t)\|}{\|\vec{v}(t)\|^3}$$

10.5.3 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 $\langle x(t), y(t), z(t)\rangle$. 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 $\mathbf{r}(t) = (1 \sin t)\,\mathbf{i} + (1 \sin t)\,\mathbf{j} + (4 \cos t)\,\mathbf{k}$

3. A factory has a machine which bends wire at a rate of 6 unit(s) of curvature per second. How long does it take to bend a straight wire into a circle of radius 8?

_______________ seconds

4. 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

$$\kappa = \frac{|f''(x)|}{[1 + (f'(x))^2]^{3/2}}.$$

5. 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 10.5.3.4 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?

6. 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 Exercise 10.4.6.4 we showed that $|\vec{T}(t)| = 1$ implies that $\vec{T}(t)$ is orthogonal to $\vec{T}'(t)$ for every value of t. Explain what this tells us about $|\vec{T}(t) \times \vec{T}'(t)|$ and conclude that

$$|\vec{r}'(t) \times \vec{r}''(t)| = \left(\frac{ds}{dt}\right)^2 |\vec{T}'(t)|.$$

e. Finally, use the fact that $\kappa = \frac{|\vec{T}'(t)|}{|\vec{r}'(t)|}$ to verify that

$$\kappa = \frac{|\vec{r}'(t) \times \vec{r}''(t)|}{|\vec{r}'(t)|^3}.$$

7. 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 10.4.3 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 10.6.4 for an illustration using an arbitrary function f.

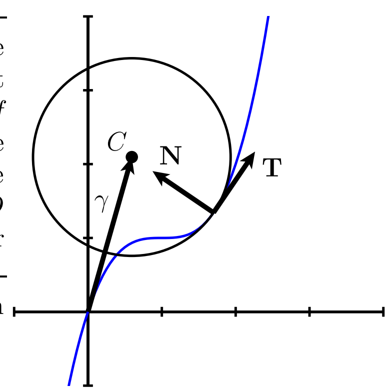


Figure 10.5.10 An osculating circle.

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.

8. Exercise about how v cross a will mean that there is no turning for the path at that point. Reference this in the section text.

10.6 Splitting the Acceleration Vector

Motivating Questions

- How can we split acceleration into the direction of travel and the direction of turning.

Velocity and acceleration are vital measurements in the study of physics and mechanics. Velocity is the derivative of position/location (the central measurement of kinematics, the study of motion). To better understand velocity as a vector valued function of time, we looked at how the magnitude and direction of position change. In particular, we separated velocity into its critical parts, namely speed (magnitude of velocity) and unit tangent vector (direction of travel). Speed is a scalar measurement for how fast the position of an object is changing as a function of time and $\vec{T}$ is a unit vector that measures in what direction position is changing.

Newton's Second Law of Motion relates the forces acting on an object to its acceleration through the formula $\vec{F}_{net} = m\vec{a}$, which is summarized as the sum of the forces acting on an object is the same as the mass times the acceleration of the object. We would like to understand 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. Further, we will make sense of the physical meaning of these different measurements.

Preview Activity 10.6.1 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 new infusion of cash to buy a gyroscopic accelerometer which will measure the acceleration as a vector with magnitude and direction. You will need to understand how your accelerometer readings will relate to the operations needed to drive the car.

(a) Figure 10.5.1 shows the path you drove on a test drive with the point P and the direction of travel P labeled. Draw and label vectors in the directions of $\vec{T}$ and $\vec{N}$ at the point P.

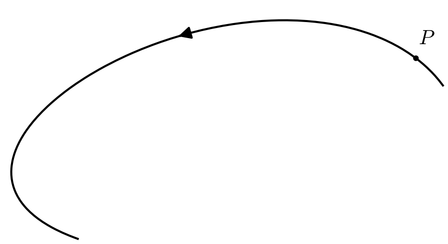


Figure 10.6.1 The curve C with direction of travel and point P labeled

Figure 10.5.2 shows $\vec{a}$, the acceleration vector of your test car at the point P (drawn in red). Draw the projection of $\vec{a}$ on $\vec{T}$ and the projection of $\vec{a}$ on $\vec{N}$ at the point P.

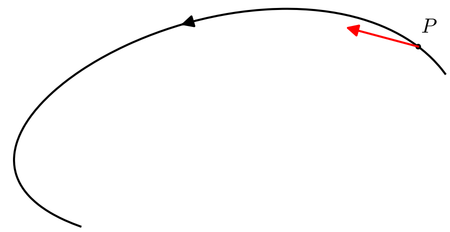


Figure 10.6.2 A 2D curve with acceleration vector shown in red at the point P

Based on your answers to part .b, Is $\vec{a} \cdot \vec{T}$ is positive, negative, or zero at P? Write a sentence or two to justify your answer.Based on your answers to part .b, Is $\vec{a} \cdot \vec{N}$ is positive, negative, or zero at P? Write a sentence or two to justify your answer.

(b) Figure 10.5.1 shows the path you drove on a test drive with the point Q and the direction of travel P labeled. Draw and label vectors in the directions of $\vec{T}$ and $\vec{N}$ at the point Q.

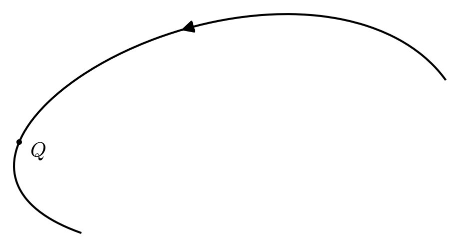


Figure 10.6.3 The curve C with direction of travel and point Q labeled

Figure 10.5.4 shows $\vec{a}$, the acceleration vector at the point Q (drawn in red). Draw the projection of $\vec{a}$ on $\vec{T}$ and the projection of $\vec{a}$ on $\vec{N}$ at the point Q.

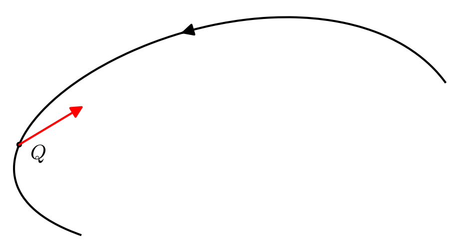


Figure 10.6.4 A 2D curve with acceleration vector shown in red at the point Q

Based on your answers to part .f, Is $\vec{a} \cdot \vec{T}$ is positive, negative, or zero at Q? Write a sentence or two to justify your answer.Based on your answers to part .f, Is $\vec{a} \cdot \vec{N}$ is positive, negative, or zero at Q? Write a sentence or two to justify your answer.

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 and efficient tools for calculating the unit normal vector to a curve, $\vec{N}$.

10.6.1 Splitting Acceleration

Recall that by the definition of $\vec{T}$, $\vec{v}$ is parallel to $\vec{T}$ (when $\vec{v}$ is not the zero vector or DNE) and that $\vec{T} \cdot \vec{N} = 0$. In other words, all of the velocity is in a parallel direction to the unit tangent vector and the unit normal is *always* orthogonal to the unit tangent 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.

Definition 10.6.5 The **tangential component of acceleration** is defined as $a_{\vec{T}} = \vec{a} \cdot \vec{T}$ and the **normal component of acceleration** is defined as $a_{\vec{N}} = \vec{a} \cdot \vec{N}$. Because $\vec{T}$ and $\vec{N}$ are orthogonal to each other and unit vectors, $a_{\vec{T}}$ and $a_{\vec{N}}$ will give the amount of acceleration in the direction of travel and the direction of turning, respectively. ◊

Given the definitions of $a_{\vec{T}}$ and $a_{\vec{N}}$, can the acceleration vector of parameterized curve have any additional parts? Can $\vec{a}$ have any component in another direction (like $\vec{B}$)?

It is not obvious, but the answer is ***No!*** The acceleration vector will only have components in the $\vec{T}$ and $\vec{N}$ directions. 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 neat product rule argument based on how we split the velocity vector into speed (magnitude) and $\vec{T}$ (direction):

$$\begin{aligned}
\vec{v}(t) =& \text{speed}(t)\ \vec{T}(t)\\
\frac{d\vec{v}}{dt} =& \frac{d}{dt}(\text{speed})\ \vec{T} + (\text{speed})\ \frac{d}{dt}(\vec{T})\\
\frac{d\vec{v}}{dt} =& \frac{d}{dt}(\text{speed})\ \vec{T} + (\text{speed})\ \left\|\frac{d\vec{T}}{dt}\right\| \vec{N}
\end{aligned}$$

The last line of this algebra follows from the definition of the unit normal vector, $\vec{N} = \frac{\vec{T}'}{\|\vec{T}'\|} \Rightarrow \vec{T}' = \|\vec{T}'\|\vec{N}$.

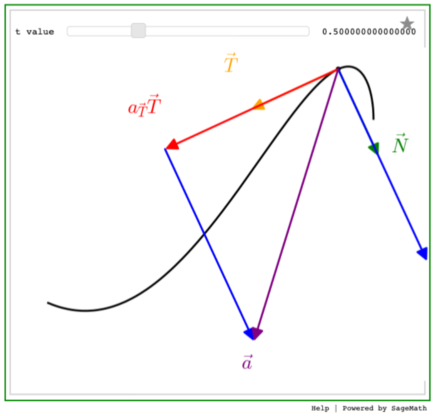
This algebraic splitting of acceleration gives a nice way to interpret $a_{\vec{T}}$. Specifically, $a_{\vec{T}}$ is the rate of change of speed!

The scalar $a_{\vec{N}}$ is harder to interpret right now because our algebraic splitting above gives $a_{\vec{N}} = (\text{speed})\|\frac{d\vec{T}}{dt}\|$ and $\frac{d\vec{T}}{dt}$ (and $\|\frac{d\vec{T}}{dt}\|$) are just as hard to calculate as $\vec{N}$ directly (which is awful as shown in Activity 10.4.4). We will address 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.

Definition 10.6.6 When we refer to the **splitting of the acceleration (vector)**, we mean that we want to explicitly describe how much of the acceleration is in the unit tangent direction and how much is in the unit normal direction (relative to the motion on the curve).

In particular, we want to find $a_{\vec{T}}$ and $a_{\vec{N}}$ such that $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$. Figure 10.5.7 shows graphically what this splitting looks like. You can use the slider at the change the location on the curve where the splitting of the acceleration is shown. You should look at 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 it's 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
Embed

Figure 10.6.7 A 2D curve with the splitting of the acceleration vector shown graphically

◊

Activity 10.6.2 In this activity we will make sense for whether $a_{\vec{T}}$ or $a_{\vec{N}}$ can be zero or negative.

(a) The tangential part of acceleration is defined as $a_{\vec{T}} = \vec{a} \cdot \vec{T}$ and our algebraic interpretation also gives us that $a_{\vec{T}} = \frac{d(\text{speed})}{dt}$. Using these formulas as a basis for your explanations, state whether $a_{\vec{T}}$ can be zero or not. Either explain why this is not possible or what it would mean for $a_{\vec{T}}$ to be zero.

(b) The tangential part of acceleration is defined as $a_{\vec{T}} = \vec{a} \cdot \vec{T}$ and our algebraic interpretation also gives us that $a_{\vec{T}} = \frac{d(\text{speed})}{dt}$. Using these formulas as a basis for your explanations, state whether $a_{\vec{T}}$ can be negative or not. Either explain why this is not possible or what it would mean for $a_{\vec{T}}$ to be negative.

(c) The normal part of acceleration is defined as $a_{\vec{N}} = \vec{a} \cdot \vec{N}$ and our algebraic interpretation also gives us that $a_{\vec{N}} = (\text{speed})\|\frac{d\vec{T}}{dt}\|$. Using these formulas as a basis for your explanations, state whether $a_{\vec{N}}$ can be negative or not. Either explain why this is not possible or what it would mean for $a_{\vec{N}}$ to be negative.

(d) The normal part of acceleration is defined as $a_{\vec{N}} = \vec{a} \cdot \vec{N}$ and our algebraic interpretation also gives us that $a_{\vec{N}} = (\text{speed})\|\frac{d\vec{T}}{dt}\|$. Using these formulas as a basis for your explanations, state whether $a_{\vec{N}}$ can be zero or not. Either explain why this is not possible or what it would mean for $a_{\vec{N}}$ to be zero.

Exercise 10.6.4.3 and Exercise 10.6.4.4 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) will affect these vector calculus quantities. In Exercise 10.6.4.5, we will look at how to visually identify the different measurements involved the splitting of the acceleration vector.

Example 10.6.8

(a) In this example we will look at the splitting of acceleration for an algebraically simple curve traced by $\vec{r}(t) = \langle t, t^2, t^3 \rangle$ as t goes from 0 to 2. We will first look at calculating the splitting of acceleration for all t-values in our interval. In particular, we will calculate $\vec{v}$, $\vec{a}$, $\vec{T}$, $a_{\vec{T}}$, $a_{\vec{N}}$, and $\vec{N}$ as functions of t.

$$\begin{aligned}
\vec{r}(t) &= \langle t, t^2, t^3 \rangle \\
\vec{v}(t) &= \langle 1, 2t, 3t^2 \rangle \\
\vec{a}(t) &= \langle 0, 2, 6t \rangle
\end{aligned}$$

The calculation of the velocity and acceleration are generally speaking the most familiar and as this example will show, it is possible to state each of our other calculations in terms of velocity and acceleration.

The next quantity to calculate is $\vec{T}$. Note that speed$(t) = \|\vec{v}(t)\| = \sqrt{1 + 4t^2 + 9t^4}$ for our example.

$$\vec{T}(t) = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\langle 1, 2t, 3t^2 \rangle}{\sqrt{1 + 4t^2 + 9t^4}}$$

This will allow us to calculate $a_{\vec{T}}$ next from the definition $a_{\vec{T}} = \vec{a} \cdot \vec{T}$

$$\begin{aligned}
a_{\vec{T}} = \vec{a} \cdot \vec{T} &= \langle 0, 2, 6t \rangle \cdot \frac{\langle 1, 2t, 3t^2 \rangle}{\sqrt{1 + 4t^2 + 9t^4}} \\
&= \frac{1}{\sqrt{1 + 4t^2 + 9t^4}} \left(1(0) + 2(2t) + 6t(3t^2) \right) \\
&= \frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}}
\end{aligned}$$

So far in this example, we have used the definitions to calculate both $\vec{T}$ and $a_{\vec{T}}$, but $\vec{N}$ is quite difficult to calculate using the definition. Our next couple of steps will use some vector calculations related to Definition 10.5.6. Because the unit tangent and unit normal vectors are both length one and are orthogonal to each other, then $\|\vec{a}\|^2 = (a_{\vec{T}})^2 + (a_{\vec{N}})^2$. In particular, we can solve for $a_{\vec{N}} = \sqrt{\|\vec{a}\|^2 - (a_{\vec{T}})^2}$ which will not require an interpretation of positive/negative because $a_{\vec{N}}$ will always be non-negative. For our example,

$$\begin{aligned}
a_{\vec{N}} &= \sqrt{\|\vec{a}\|^2 - (a_{\vec{T}})^2} \\
&= \sqrt{(4 + 36t^2) - \left(\frac{4t + 18t^3}{\sqrt{1 + 4t^2 + 9t^4}} \right)^2} \\
&= \sqrt{(4 + 36t^2) - \frac{(4t + 18t^3)^2}{1 + 4t^2 + 9t^4}} \\
&= \sqrt{\frac{(1 + 4t^2 + 9t^4)(4 + 36t^2) - (4t + 18t^3)^2}{1 + 4t^2 + 9t^4}} \\
&= \sqrt{\frac{4(9t^4 + 9t^2 + 1)}{1 + 4t^2 + 9t^4}}
\end{aligned}$$

Note how we were able to find the acceleration in the normal direction ($a_{\vec{N}}\vec{N}$) by subtracting out the acceleration in the tangential direction.

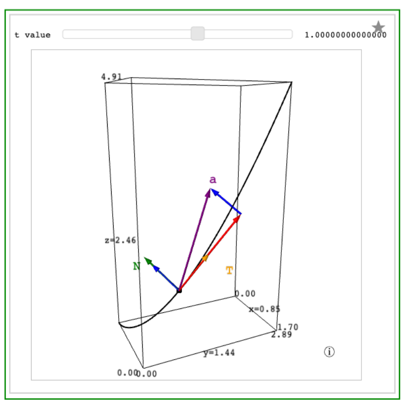
Our calculation above utilizes that $a_{\vec{N}}$ is the length of this acceleration in the normal direction. The other side to this is that we can calculate $\vec{N}$ as a function of t without having to go through the difficulties involved in $\frac{d\vec{T}}{dt}$. Specifically, $a_{\vec{N}}\vec{N} = \vec{a} - a_{\vec{T}}\vec{T}$ implies that $\vec{N} = \frac{\vec{a}-a_{\vec{T}}\vec{T}}{a_{\vec{N}}}$. This is hugely useful because we were able to calculate everything on the right side of this equation in terms of $\vec{v}$ and $\vec{a}$. For our example,

$$\begin{aligned} a_{\vec{N}}\vec{N} &= \vec{a} - a_{\vec{T}}\vec{T} \\ &= \langle 0,2,6t\rangle - \frac{4t+18t^3}{\sqrt{1+4t^2+9t^4}}\frac{\langle 1,2t,3t^2\rangle}{\sqrt{1+4t^2+9t^4}} \end{aligned}$$

Thus 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 may seem a bit algebraically tedious, these steps are much less difficult than finding $\vec{N}$ directly (as in Activity 10.4.4) AND we get insight into how the acceleration is split in the direction of travel and direction of turning. Graphically, we can represent this splitting with the following figure.



Standalone
Embed

Figure 10.6.9 The curve given by $\vec{r}(t) = \langle t, t^2, t^3\rangle$ with the splitting of acceleration shown graphically

(b) In many circumstances, we need to compute the splitting of acceleration for only one time. For our example, we will look at how much cleaner the computations above will be if we want to look only at $t = 1$. The first steps for calculating $\vec{v}$ and $\vec{a}$ will be the same as before.

$$\begin{aligned} \vec{r}(t) &= \langle t, t^2, t^3\rangle \\ \vec{v}(t) &= \langle 1, 2t, 3t^2\rangle \\ \vec{a}(t) &= \langle 0, 2, 6t\rangle \end{aligned}$$

You may want to go back and note that this was the only calculus done in the example above; all of our other steps were vector calculations and properties of orthogonal or unit vectors. We will follow the same method as above, but instead of applying these ideas for all t-values, we will look only at $t = 1$. Remember if the formula we are using works for all t, then the formula will work for a specific t.

$$\begin{aligned}
\vec{r}(1) &= \langle 1, 1, 1\rangle \\
\vec{v}(1) &= \langle 1, 2, 4\rangle \\
\vec{a}(1) &= \langle 0, 2, 6\rangle
\end{aligned}$$

The next quantity to calculate is $\vec{T}$. Note that speed$(1) = \|\vec{v}(1)\| = \sqrt{1+4+9} = \sqrt{14}$ for our example.

$$\vec{T}(1) = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\langle 1, 2, 3\rangle}{\sqrt{14}}$$

This will allow us to calculate $a_{\vec{T}}(1)$ next from the definition $a_{\vec{T}}(1) = \vec{a}(1) \cdot \vec{T}(1)$

$$\begin{aligned}
a_{\vec{T}}(1) = \vec{a}(1) \cdot \vec{T}(1) = \langle 0, 2, 6\rangle \cdot \frac{\langle 1, 2, 3\rangle}{\sqrt{14}} \\
= \frac{1}{\sqrt{14}}(1(0) + 2(2) + 3(6)) = \frac{22}{\sqrt{14}}
\end{aligned}$$

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.

$$\begin{aligned}
a_{\vec{N}}(1) &= \sqrt{\|\vec{a}(1)\|^2 - (a_{\vec{T}}(1))^2} \\
&= \sqrt{(40) - \left(\frac{22}{\sqrt{14}}\right)^2} \\
&= \sqrt{\frac{560}{14} - \frac{22^2}{14}} \\
&= \sqrt{\frac{76}{14}}
\end{aligned}$$

We can now compute how much acceleration there is due to turning and use that explicitly compute $\vec{N}(1)$.

$$\begin{aligned}
a_{\vec{N}}(1)\vec{N}(1) &= \vec{a}(1) - a_{\vec{T}}(1)\vec{T}(1) \\
&= \langle 0, 2, 6\rangle - \frac{22}{\sqrt{14}}\frac{\langle 1, 2, 3\rangle}{\sqrt{14}} \\
&= \langle -\frac{11}{7}, -\frac{16}{7}, \frac{9}{7}\rangle
\end{aligned}$$

You can show that the length of $\langle -\frac{11}{7}, -\frac{16}{7}, \frac{9}{7}\rangle$ is the same as our calculation for $a_{\vec{N}}$above. In particular, this means that

$$\vec{N} = \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}} = \sqrt{\frac{7}{38}}\langle -\frac{11}{7}, -\frac{16}{7}, \frac{9}{7}\rangle$$

With some diligent computations, you should be able to verify that this result is the same as applying $t = 1$ to our expression for $\vec{N}$ from the first task.

$\square$

The method used in the previous example can be applied broadly and shows how to calculate all parts of Definition 10.5.6 in terms of $\vec{v}$ and $\vec{a}$ (for both an interval or a single value of t).

Efficient Calculation of $a_{\vec{T}}, a_{\vec{N}}, \vec{T}, \vec{N}$.

$$\vec{T} = \frac{\vec{v}}{\|\vec{v}\|}$$
$$a_{\vec{T}} = \frac{\vec{a}\cdot\vec{v}}{\|\vec{v}\|} \Rightarrow a_{\vec{T}}\vec{T} = \frac{\vec{a}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$$
$$a_{\vec{N}} = \|\vec{a} - a_{\vec{T}}\vec{T}\| = \sqrt{\|\vec{a}\| - (a_{\vec{T}})^2} = \sqrt{(\vec{a}\cdot\vec{a}) - \frac{(\vec{a}\cdot\vec{v})^2}{\|\vec{v}\|}}$$
$$a_{\vec{N}}\vec{N} = \vec{a} - a_{\vec{T}}\vec{T}$$
$$a_{\vec{N}} = \frac{\vec{a} - a_{\vec{T}}\vec{T}}{a_{\vec{N}}}$$

Activity 10.6.3 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 $\vec{N}$, then verify $\vec{a} = a_{\vec{T}}\vec{T} + a_{\vec{N}}\vec{N}$ and that $\vec{T}$ is orthogonal to $\vec{N}$.

10.6.2 The Driver or the Road?

Exercise 10.6.4.3, Exercise 10.6.4.4, and Exercise 10.6.4.5 offer great conceptual and physical meaning to the different measurements involved in the splitting of acceleration, but 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 I said the same thing twice, but $a_{\vec{N}}$ and κ are not the same. Theorem **??** 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 $\left\|\frac{d\vec{T}}{dt}\right\| = \kappa(speed)$.

We saw earlier in this section that

$$a_{\vec{N}} = (speed)\|\frac{d\vec{T}}{dt}\|$$

Thus, we can write $a_{\vec{N}}$ in terms of speed and curvature as

$$a_{\vec{N}} = \kappa(\text{speed})^2 \tag{10.6.1}$$

(**??**) should help to demonstrate the difference between $a_{\vec{N}}$ and κ. Specifically, κ measures how fast the path is curving and the normal component of acceleration, $a_{\vec{N}}$, measures the acceleration felt due to turning. (**??**) shows that if you and I are moving along the same path but you are going twice as fast as I am, then you will feel *four times* as much acceleration due to turning (even though we are going along the exact same path). Curvature is how fast the path is turning and $a_{\vec{N}}$ is how fast you feel like you are turning (based on your speed and the parameterization).

Activity 10.6.4 In this activity, we will look at the measurements related to the splitting of acceleration and determine if these measurements are properties of the driver or properties of the road. This is a continuation of Activity 10.3.5. 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). The explanations for the next tasks may be difficult for you to write but will be very helpful in ensuring you understand the vector calculus concepts of this chapter.

(a) Pedal Usage: We measure the pedal usage as either +, 1, or zero in the following way:

- + if the gas pedal is being used
- - if the brake pedal is being used
- zero if no pedal is being used

Since our race car drivers are safety minded, they do not use more than one pedal at a time. Note that this measurement will just be a sign and not a numerical value. The pedal usage will give us the sign of which vector calculus quantity? Be sure to explain your answer.

(b) Describe how the value of $a_{\vec{N}}$ would be felt by the racecar driver in our analogy.

(c) In terms of our race car situation, explain why $a_{\vec{N}}$ can't be negative.

(d) 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) $\vec{N}(b)$

(e) 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) $\vec{N}(c)$

(f) 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) $a_{\vec{N}}(c)$

(g) Is Jane or Nick turning the steering wheel farther? Explain how you know this based on your calculations above.

(h) 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.

(i) 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.

(j) Nick and Jane talk to their friend Jeremy 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.

10.6.3 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.

10.6.4 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.

(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 $\vec{N}$.

2. 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 $\vec{N}(11)$

3. 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.

4. 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 $a_{\vec{N}}$?

5. 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 10.5.10. Match each color with proper vector.

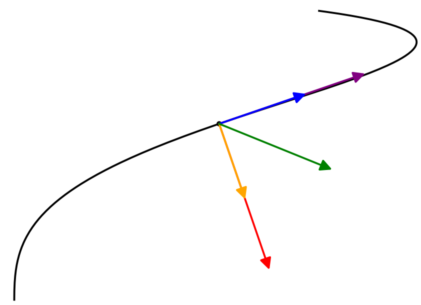


Figure 10.6.10 A 2D curve with splitting of acceleration shown using randomized colors

(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 10.5.11. Match each color with proper vector.

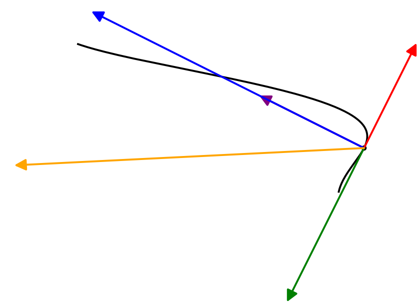


Figure 10.6.11 A 2D curve with splitting of acceleration shown using randomized colors

(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 10.5.12. Match each color with proper vector.

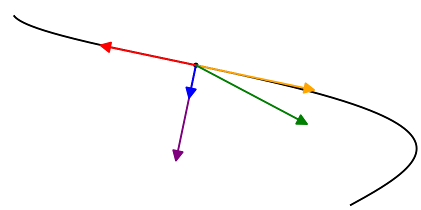


Figure 10.6.12 A 2D curve with splitting of acceleration shown using randomized colors

Chapter 11

Derivatives of Multivariable Functions

11.1 Functions of Several Variables

Motivating Questions

- What is a function of several variables? What do we mean by the domain of a function of several variables?
- How do we find the distance between two points in $\mathbb{R}^3$? What is the equation of a sphere in $\mathbb{R}^3$?
- What is a trace of a function of two variables? What does a trace tell us about a function?
- What is a level curve of a function of two variables? What does a level curve tell us about a function?

Much of algebra, precalclus, and calculus has been focused on working with functions that have a scalar input and a scalar output. In previous chapter, we expanded this to address functions with a scalar input and a vector output. The calculus of these functions was not very interesting since we applied limits, derivatives, and integrals to these functions componontwise.

Many problems involve a much 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 (measured by a single number) that will vary over three dimensions (north/south, east/west, and over 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. Wind is measured with a vector (magnitude and direction) and also varies with location in a three dimensional space. So wind would be given by a function with a location in 3D as an input and a vector as an output. We will look at the calculus of multivariable functions with scalar outputs (like temperature and pressure) in the next couple of chapters and will look at studying functions with multivariable inputs and outputs in the last chapter.

Throughout our mathematical careers we have studied functions of a single variable. We define a function of one variable as a rule that assigns exactly one output to each input. We analyze these functions by looking at their graphs, calculating limits, differentiating, integrating, and more. Functions of several variables will be the main focus of Chapters 10 and 11, where we will analyze these functions by looking at their graphs, calculating limits, differentiating, integrating, and more. We will see that many of the ideas from single variable calculus translate well to functions of several variables, but we will have to make some adjustments as well. In this chapter we introduce functions of several variables and then discuss some of the tools (vectors and vector-valued functions) that will help us understand and analyze functions of several variables.

Preview Activity 11.1.1 Suppose you invest money in an account that pays 5% interest compounded continuously. If you invest P dollars in the account, the amount A 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 $A(5000, 8)$. Explain in words 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 Table 11.1.1. Round payments to the nearest penny.

Table 11.1.1 Amount of money in an account with an initial investment of 1000 dollars.

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 Table 11.1.2. Round payments to the nearest penny.

Table 11.1.2 Amount of money in an account after 10 years.

Initial investment (dollars)	500	1000	5000	7500	10000
Amount (dollars)					

e. Describe as best you can the combinations of initial investments and time that result in an account containing \$10,000.

11.1.1 Functions of Several Variables

Up to this point we have been concerned with functions of a single variable. What defined such a function is that every input in the domain produced a unique output in the range. We saw similar behavior in Preview Activity 11.1.1, where each pair (P, t) of inputs produces a unique output $A(P, t)$. Additionally, the two variables P and t had no real relation to each other. That is, we could choose any value of P without considering 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 11.1.3 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, of course, 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.

As with functions of a single variable, it is important to understand the set of inputs for which the function is defined.

Definition 11.1.4 The **domain** of a function f is the set of all inputs at which the function is defined. ♢

Activity 11.1.2 Identify the domain of each of the following functions. Draw a picture of 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. $s(x, y) = \dfrac{1}{\sqrt{1 - xy^2}}$

11.1.2 Representing Functions of Two Variables

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 will have to allow us to keep track of both input variables. We can do this with a 2-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, using a golf club, a cannon, or some other device, from ground level. Under ideal conditions (ignoring wind resistance, spin, or any other forces except the force of gravity) the horizontal distance the object will travel 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. (Note that g is constant, 32 feet per second squared. We will derive this equation in a later section.) 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 then displayed in the location where the x row intersects the y column, as shown in Table 11.1.5 (where we measure x in feet per second and y in radians).

Table 11.1.5 Values of $f(x, y) = \frac{x^2 \sin(2y)}{g}$**.**

$x \backslash y$	0.2	0.4	0.6	0.8	1.0	1.2	1.4
25	7.6	14.0	18.2	19.5	17.8	13.2	6.5
50	30.4	56.0	72.8	78.1	71.0		26.2
75	68.4		163.8	175.7	159.8	118.7	58.9
100	121.7	224.2	291.3	312.4	284.2	211.1	104.7
125	190.1	350.3	455.1		444.0	329.8	163.6
150	273.8	504.4	655.3	702.8	639.3	474.9	235.5
175	372.7	686.5	892.0	956.6	870.2	646.4	
200	486.8	896.7	1165.0	1249.5	1136.6	844.3	418.7
225	616.2	1134.9	1474.5	1581.4	1438.5	1068.6	530.0
250	760.6	1401.1		1952.3	1776.0	1319.3	654.3

Activity 11.1.3 Complete Table 11.1.5 by filling in the missing values of the function f. Round entries to the nearest tenth.

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 11.1.5 identifies points of the form $(x, y, f(x, y))$, and we define the graph of f to be the set of these points.

Definition 11.1.6 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 in the form $(x, y, f(x, y))$ are in three dimensions, so plotting these points takes a bit more work than graphs of functions in two dimensions. To plot these three-dimensional points, we use a coordinate system with three mutually perpendicular axes — the x-axis, the y-axis, and the z-axis (called the *coordinate axes*). As shown in Figure 9.1.8, we will use a right handed coordinate system for our plots. We can draw a graph of the distance function defined by $f(x, y) = \frac{x^2 \sin(2y)}{g}$. Note that the function f is continuous in both variables, so when we plot these points in the right hand coordinate system, we can connect them all to form a surface in 3-space. The graph of the distance function f is shown in Figure 11.1.7.

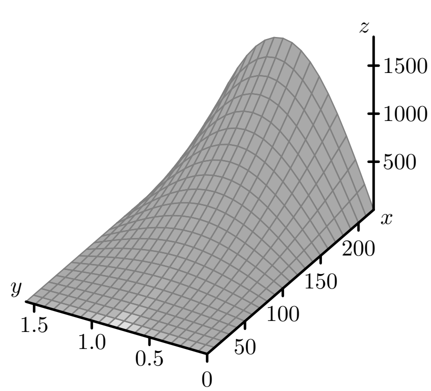


Figure 11.1.7 The distance surface.

There are many graphing tools available for drawing three-dimensional surfaces as indicated in the Preface (see Links to interactive graphics in Features of the Text). Since we will be able to visualize graphs of functions of two independent variables, but not functions of more than two variables, we will primarily deal with functions of two variables in this text. It is important to note, however, that the techniques we develop apply to functions of any number of variables.

Notation: 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).

11.1.3 Some Standard Equations in Three-Space

Activity 9.1.3 showed that the equations where one independent variable is constant lead to planes parallel to ones that result from a pair of the coordinate axes. When we make the constant 0, we get the *coordinate planes*. The xy-plane satisfies $z = 0$, the xz-plane satisfies $y = 0$, and the yz-plane satisfies $x = 0$ (see Figure 11.1.8).

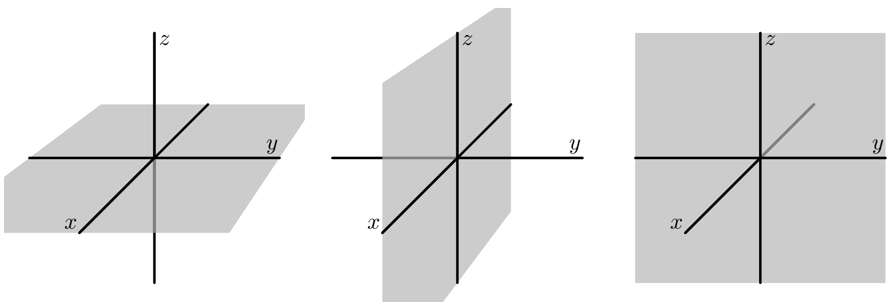


Figure 11.1.8 The coordinate planes.

11.1.4 Traces

When we study functions of several variables we are often interested in how each individual variable affects the function in and of itself. In Preview Activity 11.1.1, 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 set the duration of the investment constant, 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.

As another example, consider again the 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 consider f a function of the initial velocity alone. In this case we have

$$f(x) = \frac{x^2}{32} \sin(2 \cdot 0.6).$$

We can plot this curve on the surface by tracing out the points on the surface when $y = 0.6$, as shown at left in Figure 11.1.9. The formula clearly 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.

Similarly, if we fix the initial velocity at 150 feet per second, we can consider the distance as a function of the launch angle only. In this case we have

$$f(y) = \frac{150^2 \sin(2y)}{32}.$$

We can again plot this curve on the surface by tracing out the points on the surface when $x = 150$, as shown at right in Figure 11.1.9. The formula clearly show 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.

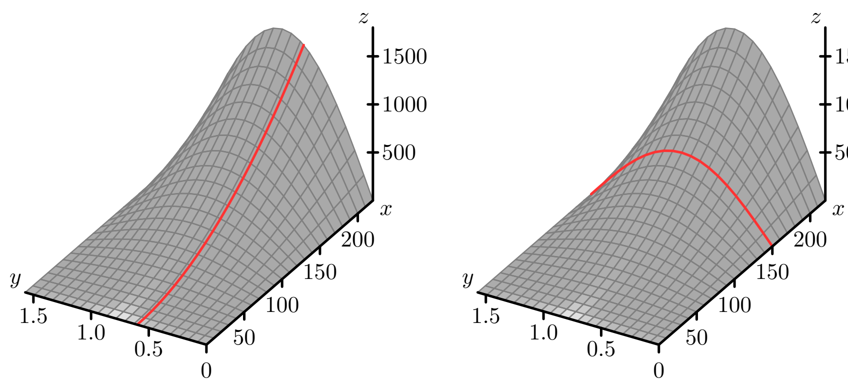


Figure 11.1.9 Left: The trace with $y = 0.6$. Right: The trace with $x = 150$.

The curves we define when we fix one of the independent variables in our two variable function are called *traces*.

Definition 11.1.10 A **trace** of a function f of two independent variables x and y in the x direction is a curve of the form $z = f(x, c)$, where c is a constant. Similarly, a **trace** of a function f of two independent variables x and y in the y direction is a curve of the form $z = f(c, y)$, where c is a constant. $\diamondsuit$

Understanding trends in the behavior of functions of two variables can be challenging, as can sketching their graphs; traces help us with each of these tasks.

Activity 11.1.4 In the following questions, 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 11.1.5. 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 11.1.5. Write a sentence to describe the behavior of the function along this trace.

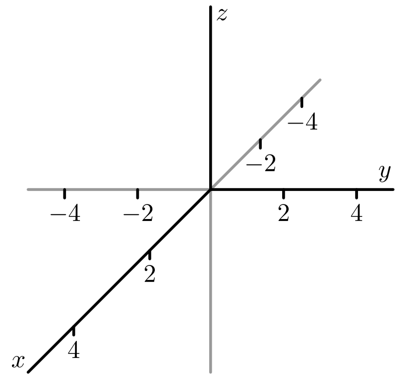


Figure 11.1.11 Coordinate axes to sketch traces.

c. For the function g defined by $g(x, y) = x^2 + y^2 + 1$, explain the type of function that each trace in the x direction will be (keeping y constant). Plot the $y = -4$, $y = -2$, $y = 0$, $y = 2$, and $y = 4$ traces in 3-dimensional coordinate system provided in Figure 11.1.11.

d. For the function g defined by $g(x, y) = x^2 + y^2 + 1$, explain the type of function that each trace in the y direction will be (keeping x constant). Plot the $x = -4$, $x = -2$, $x = 0$, $x = 2$, and $x = 4$ traces in 3-dimensional coordinate system in Figure 11.1.11.

e. Describe the surface generated by the function g.

11.1.5 Contour Maps and Level Curves

We have all seen topographic maps such as the one of the Porcupine Mountains in the upper peninsula of Michigan shown in Figure 11.1.12.[1] The curves on these maps show the regions of constant altitude. The contours also depict changes in altitude: contours that are close together signify steep ascents or descents, while contours that are far apart indicate only slight changes in elevation. Thus, contour maps tell us a lot about three-dimensional surfaces. Mathematically, if $f(x, y)$ represents the altitude at the point (x, y), then each contour is the graph of an equation of the form $f(x, y) = k$, for some constant k.

[1]Map source: Michigan Department of Natural Resources, with permission of the Michigan DNR and Bob Wild.

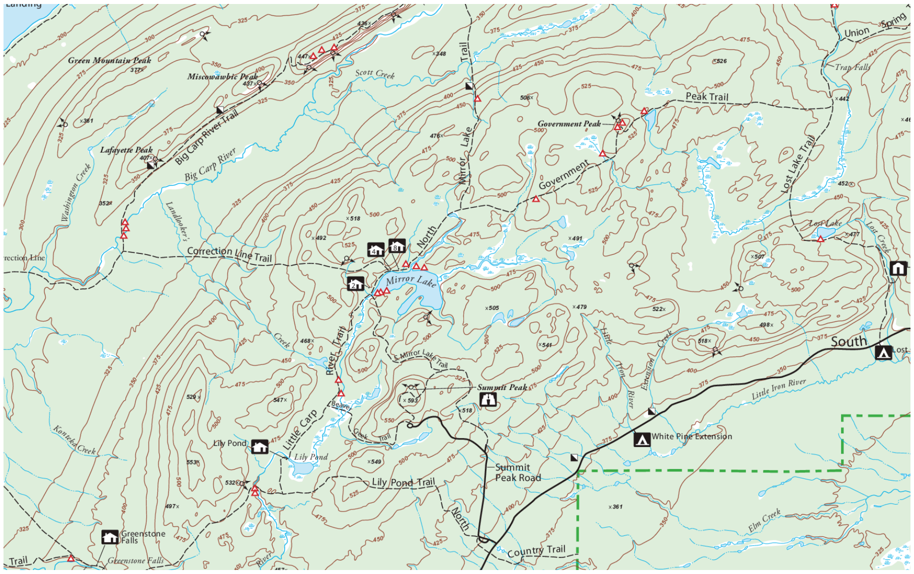


Figure 11.1.12 Contour map of the Porcupine Mountains.

Activity 11.1.5 On the topographical map of the Porcupine Mountains in Figure 11.1.12,

a. identify the highest and lowest points you can find;

b. from a point of your choice, determine a path of steepest ascent that leads to the highest point;

c. from that same initial point, determine the least steep path that leads to the highest point.

Curves on a surface that describe points at the same height or level are called *level curves*.

Definition 11.1.13 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 at left in Figure 11.1.14. 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 11.1.14.

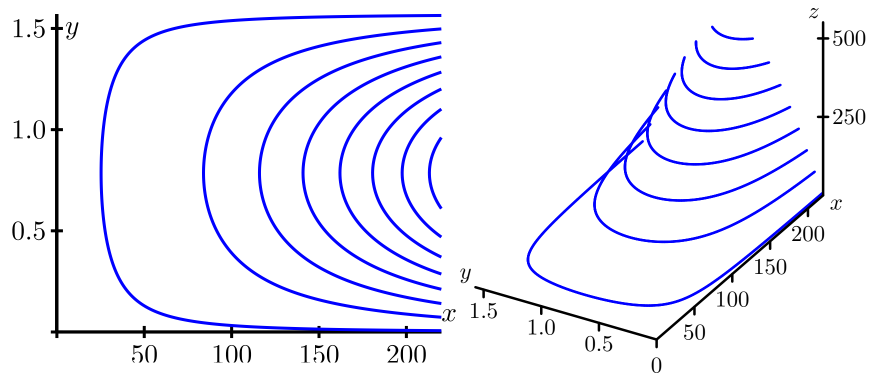


Figure 11.1.14 Left: Level curves. Right: Level curves at appropriate heights.

The use of level curves and traces can help us construct the graph of a function of two variables.

Activity 11.1.6

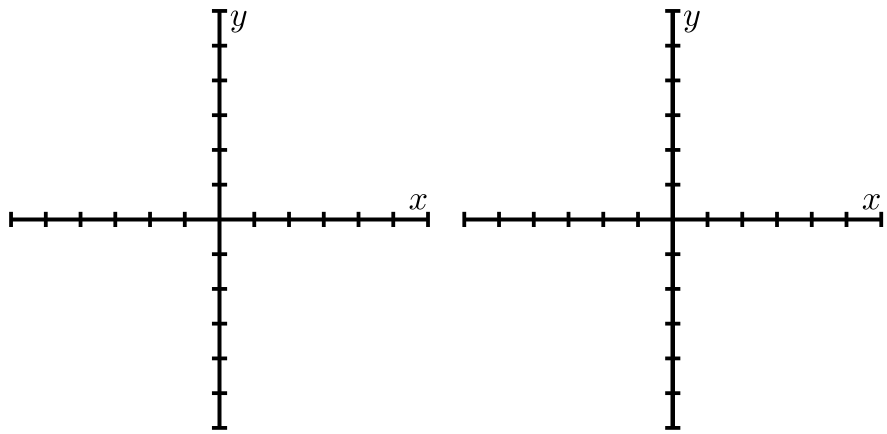


Figure 11.1.15 Left: Level curves for $f(x, y) = x^2 + y^2$. Right: Level curves for $g(x, y) = \sqrt{x^2 + y^2}$.

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 left set of axes given in Figure 11.1.15. (You decide on 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 right set of axes given in Figure 11.1.15. (You decide on the scale of the axes.) 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. In order to understand these traces and level curves better, we will first spend some time learning about vectors and vector-valued functions in the next few sections and return to our study of functions of several variables once we have those more mathematical tools to support their study.

11.1.6 A gallery of functions

We end this section by considering a collection of functions and illustrating their graphs and some level curves.

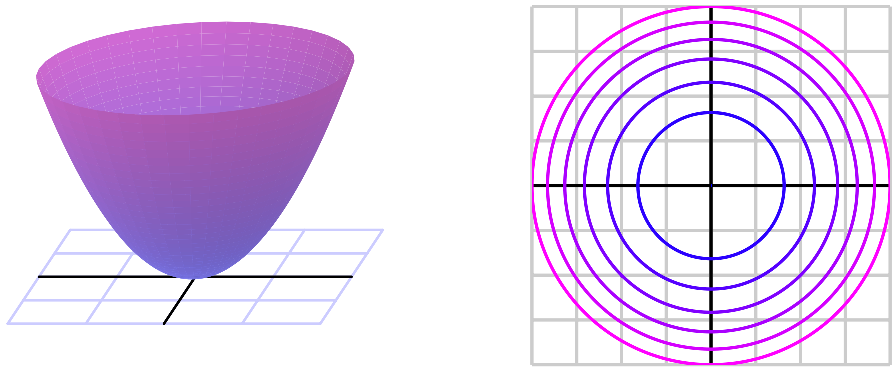


Figure 11.1.16 $z = x^2 + y^2$

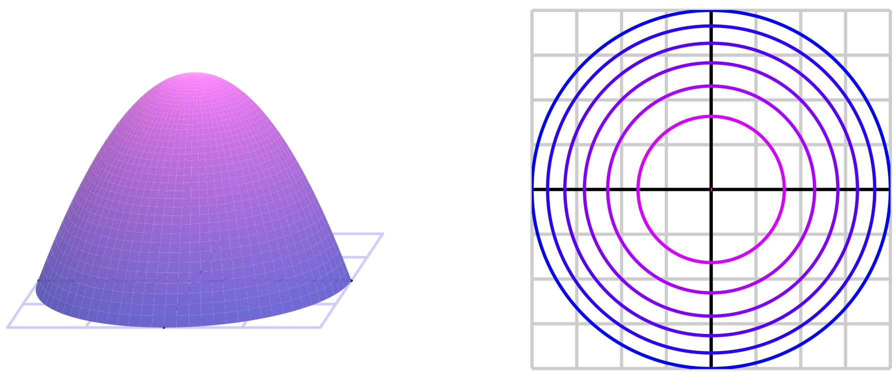


Figure 11.1.17 $z = 4 - (x^2 + y^2)$

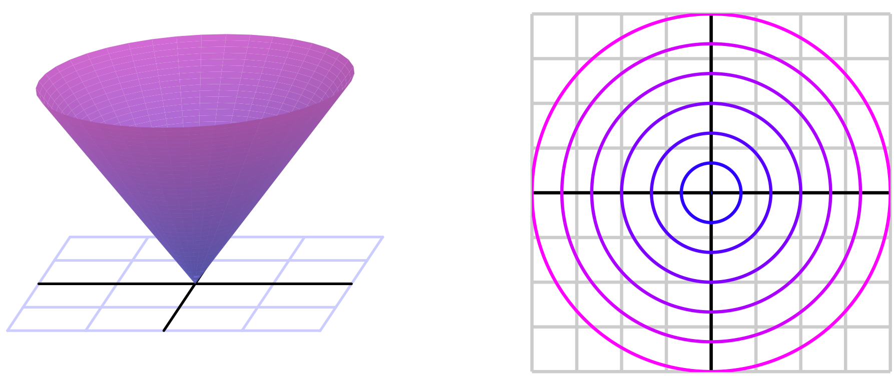


Figure 11.1.18 $z = \sqrt{x^2 + y^2}$

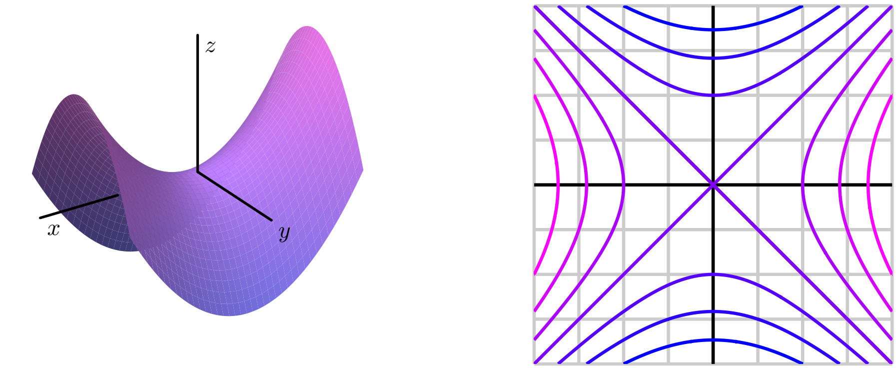


Figure 11.1.19 $z = x^2 - y^2$

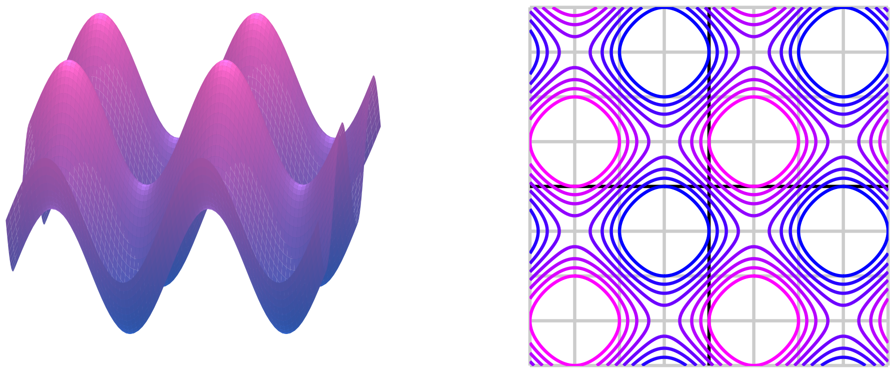


Figure 11.1.20 $z = \sin(x) + \sin(y)$

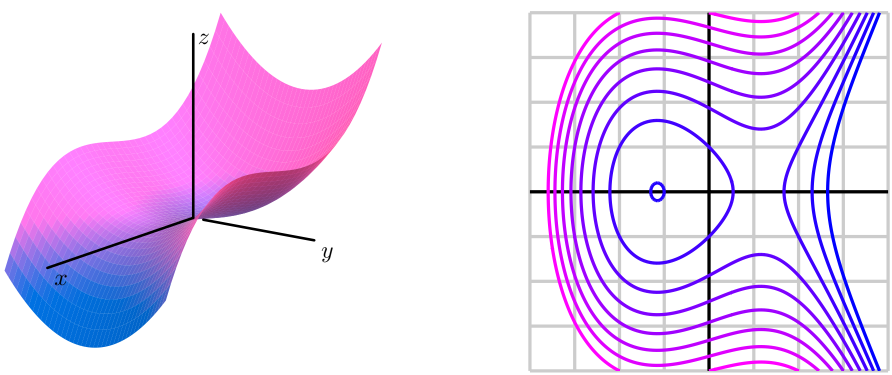


Figure 11.1.21 $z = y^2 - x^3 + x$

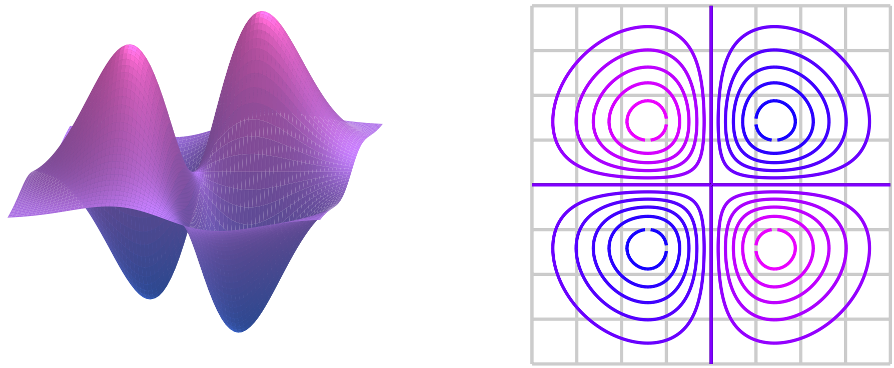


Figure 11.1.22 $z = xye^{-x^2-y^2}$

11.1.7 Summary

- A function f of several variables is a rule that assigns a unique 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.
- 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

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

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

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

- A trace of a function f of two independent variables x and y is a curve of the form $z = f(c, y)$ or $z = f(x, c)$, 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.

11.1.8 Exercises

1. Evaluate the function at the specified points.
$f(x, y) = x + yx^5, (4, 3), (-4, -2), (1, 4)$
At $(4, 3)$: ____________________
At $(-4, -2)$: ____________________
At $(1, 4)$: ____________________

2. Sketch a contour diagram of each function. Then, decide whether its contours are predominantly lines, parabolas, ellipses, or hyperbolas.

(a) $z = y - 5x^2$

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

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

(d) $z = -5x^2$

3. 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) $z = \sqrt{(25 - x^2 - y^2)}$

(b) $z = xy$

(c) $z = 2x + 3y$

(d) $z = x^2 + y^2$

(e) $z = \dfrac{1}{x-1}$

(f) $z = \sqrt{(x^2 + y^2)}$

(g) $z = 2x^2 + 3y^2$

A. a collection of unequally spaced parallel lines

B. a collection of equally spaced parallel lines

C. two straight lines and a collection of hyperbolas

D. a collection of concentric ellipses

E. a collection of equally spaced concentric circles

F. a collection of unequally spaced concentric circles

4. The domain of the function $f(x, y) = \sqrt{x} + \sqrt{y}$ is ____________________

5. Find the equation of the sphere centered at $(-4, 6, 1)$ with radius 5. Normalize your equations so that the coefficient of x^2 is 1.

____ $= 0$.

Give an equation which describes the intersection of this sphere with the plane $z = 2$.

____ $= 0$.

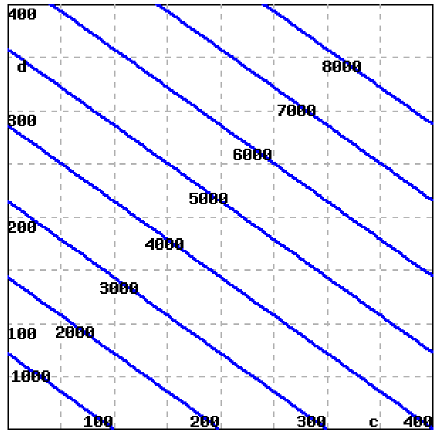
6. A car rental company charges a one-time application fee of 30 dollars, 50 dollars per day, and 15 cents per mile for its cars.

(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 $f(3, 840) =$ ____________________

7. 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

8. 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 $0 \le x \le 5$ mg and $t \ge 0$ hours, we have

$$C = f(x, t) = 26te^{-(5-x)t}$$

$f(3, 5) =$ ______________________________________

Give a practical interpretation of your answer: $f(3, 5)$ is

⊙ the concentration of a 5 mg dose in the blood 3 hours after injection.

⊙ the concentration of a 3 mg dose in the blood 5 hours after injection.

⊙ the change in concentration of a 5 mg dose in the blood 3 hours after injection.

⊙ the amount of a 3 mg dose in the blood 5 hours after injection.

⊙ the change in concentration of a 3 mg dose in the blood 5 hours after injection.

⊙ the amount of a 5 mg dose in the blood 3 hours after injection.

9. A manufacturer sells aardvark masks at a price of \$180 per mask and butterfly masks at a price of \$440 per mask. A quantity of a aardvark masks and b butterfly masks is sold at a total cost of \$550 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

⊙ parabolas

⊙ circles

⊙ lines

⊙ ellipses

10. 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 $0 \leq x \leq 4$ and $t \geq 0$, 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) $f(3,t)$
(b) $f(x,0.5)$
Using your graph in ***(a)***, where is $f(3,t)$
a maximum? $t =$ _______________
a minimum? $t =$ _______________
Using your graph in ***(b)***, where is $f(x,0.5)$
a maximum? $x =$ _______________
a minimum? $x =$ _______________

11. By setting one variable constant, find a plane that intersects the graph of $z = 3y^2 - 9x^2 + 4$ 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.)

12. 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 ($R = 8.314\ \frac{\mathrm{J}}{\mathrm{mol\ K}}$ 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.

13. 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.

14. 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?

11.2 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)?

In this section, we will study limits of functions of several variables, with a focus on limits of functions of two variables. In single variable calculus, we studied the notion of limit, which turned out to be a critical concept that formed the basis for the derivative and the definite integral. In this section we will begin to understand how the concept of limit for functions of two variables is similar to what we encountered for functions of a single variable. The limit will again be the fundamental idea in multivariable calculus, and we will use this notion of the limit of a function of several variables to define the important concept of differentiability later in this chapter. We have already seen its use in the derivatives of vector-valued functions in Section 10.2.

Let's begin by reviewing what we mean by the 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

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

Preview Activity 11.2.1 We investigate the limits of several different functions by working with tables and graphs.

a. Consider the function f defined by

$$f(x) = 3 - x.$$

Complete Table 11.2.1.

Table 11.2.1 Values of $f(x) = 3 - x$.

x	-0.2	-0.1	0.0	0.1	0.2
$f(x)$					

What does the table suggest regarding $\lim_{x \to 0} f(x)$?

b. Explain how your results in (a) are reflected in Figure 11.2.2.

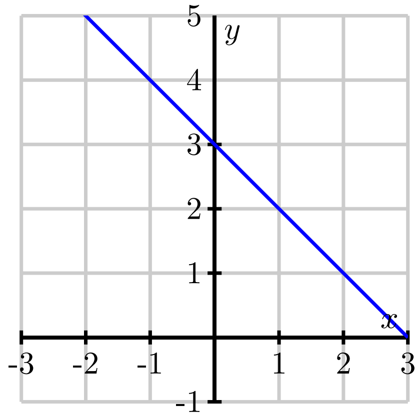


Figure 11.2.2 The graph of $f(x) = 3 - x$.

c. Next, consider

$$g(x) = \frac{x}{|x|}.$$

Complete Table 11.2.3 with values near $x = 0$, the point at which g is not defined.

Table 11.2.3 Values of $g(x) = \frac{x}{|x|}$.

x	-0.1	-0.01	-0.001	0.001	0.01	0.1
$g(x)$						

What does this suggest about $\lim_{x\to 0} g(x)$?

d. Explain how your results in (c) are reflected in Figure 11.2.4.

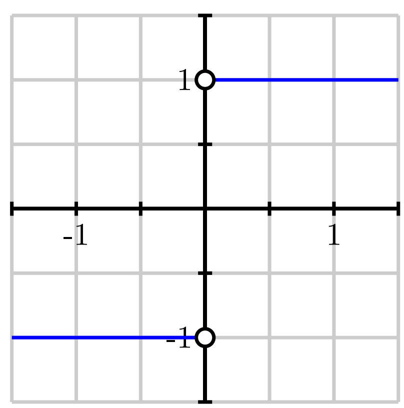


Figure 11.2.4 The graph of $g(x) = \frac{x}{|x|}$.

e. Now, let's examine a function of two variables. Let

$$f(x, y) = 3 - x - 2y.$$

Complete Table 11.2.5.

Table 11.2.5 Values of $f(x,y) = 3 - x - 2y$.

$x\backslash y$	-1.0	-0.1	0.0	0.1	1.0
-1.0		4.2			
-0.1					1.1
0.0				2.8	
0.1	4.9				
1.0			2.0		

What does the table suggest about $\lim_{(x,y)\to(0,0)} f(x,y)$?

f. Explain how your results in (e) are reflected in Figure 11.2.6. Compare this limit to the limit in part (a). How are the limits similar and how are they different?

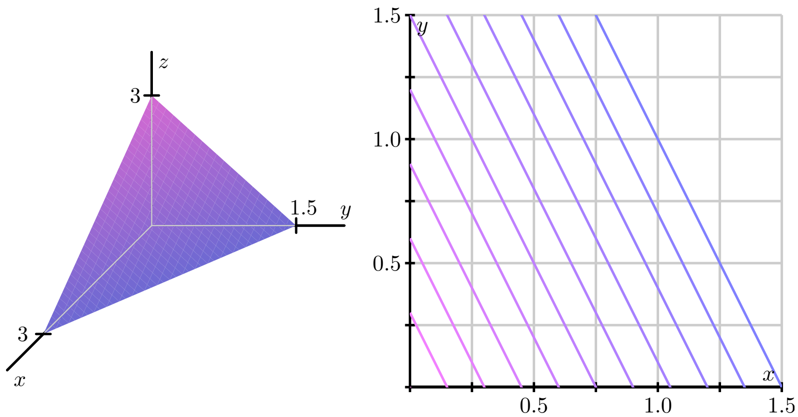


Figure 11.2.6 Left: The graph of $f(x,y) = 3 - x - 2y$. Right: A contour plot.

g. Finally, consider

$$g(x,y) = \frac{2xy}{x^2 + y^2},$$

which is not defined at $(0,0)$. Complete Table 11.2.7. Round to three decimal places.

Table 11.2.7 Values of $g(x,y) = \frac{2xy}{x^2+y^2}$.

$x\backslash y$	-1.0	-0.1	0.0	0.1	1.0
-1.0		0.198			
-0.1					-0.198
0.0			—	0.000	
0.1	-0.198				
1.0			0.000		

What does this suggest about $\lim_{(x,y)\to(0,0)} g(x,y)$?

h. Explain how your results are reflected in Figure 11.2.8. Compare this limit to the limit in part (c). How are the results similar and how are they different?

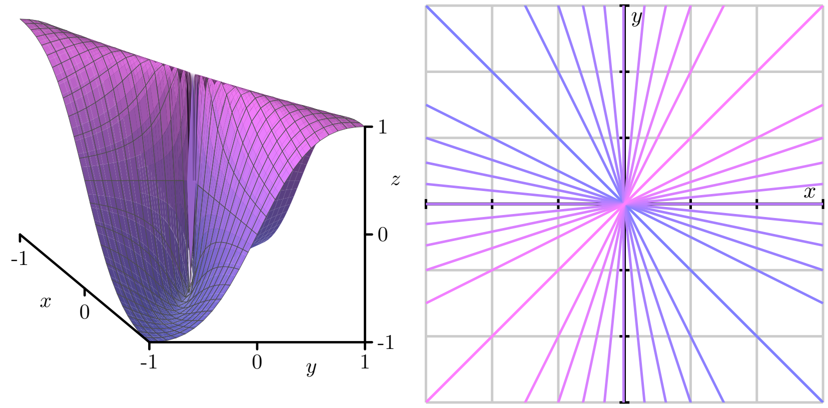


Figure 11.2.8 Left: The graph of $g(x, y) = \frac{2xy}{x^2+y^2}$. Right: A contour plot.

11.2.1 Limits of Functions of Two Variables

In Preview Activity 11.2.1, we recalled the notion of limit from single variable calculus and saw that a similar concept applies to functions of two variables. Though we will focus on functions of two variables, for the sake of discussion, all the ideas we establish here are valid for functions of any number of variables. In a natural followup to our work in Preview Activity 11.2.1, we now formally define what it means for a function of two variables to have a limit at a point.

Definition 11.2.9 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

$$\lim_{(x,y)\to(a,b)} f(x, y) = L.$$

$\diamond$

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) from various directions. This situation is more complicated because there are infinitely many ways in which (x, y) may approach (a, b). 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 11.2.2 Consider the function f, defined by

$$f(x, y) = \frac{y}{\sqrt{x^2 + y^2}},$$

whose graph is shown below in Figure 11.2.10

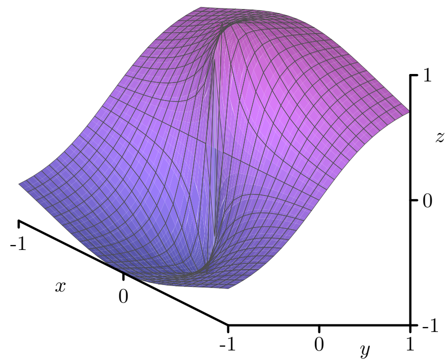


Figure 11.2.10 The graph of $f(x,y) = \frac{y}{\sqrt{x^2+y^2}}$.

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 in Table 11.2.11.

Table 11.2.11 Values of a function f.

$x\backslash y$	-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 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 the 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 $y = 0$. 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 (thus considering $f(x,x)$ as $x \to 0^+$), what value do we find as the limit?

e. In general, if $\lim_{(x,y)\to(0,0)} f(x,y) = L$, then $f(x,y)$ approaches L as (x,y) approaches $(0,0)$, regardless of the path we take in letting $(x,y) \to (0,0)$. Explain what the last two parts of this activity imply about the existence of $\lim_{(x,y)\to(0,0)} f(x,y)$.

f. Shown below in Figure 11.2.12 is a set of contour lines of the function f. What is the behavior of $f(x,y)$ as (x,y) approaches $(0,0)$ along any straight line? How does this observation reinforce your conclusion about the existence of $\lim_{(x,y)\to(0,0)} f(x,y)$ from the previous part of this activity? (Hint: Use the fact that a non-vertical line has equation $y = mx$ for some constant m.)

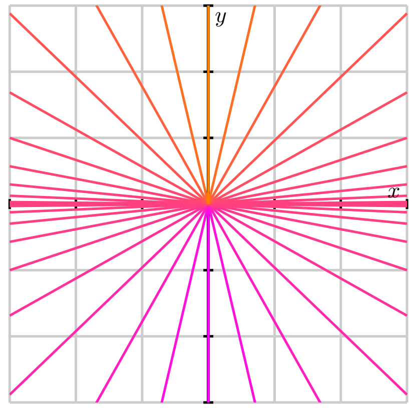


Figure 11.2.12 Contour lines of $f(x,y) = \frac{y}{\sqrt{x^2+y^2}}$.

As we have seen in Activity 11.2.2, if (x,y) approaches (a,b) along two different paths and we find that $f(x,y)$ has two different limits, we can conclude that $\lim_{(x,y)\to(a,b)} f(x,y)$ does not exist. This is similar to the one-variable example $g(x) = x/|x|$ as shown in Figure 11.2.13; $\lim_{x\to 0} g(x)$ does not exist because we see different limits as x approaches 0 from the left and the right.

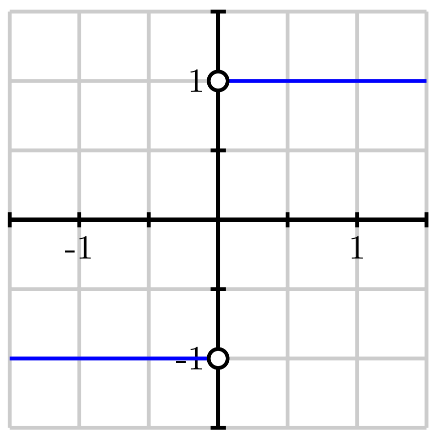


Figure 11.2.13 The graph of $g(x) = \frac{x}{|x|}$.

As a general rule, we have

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 11.2.3 Let's consider the function g defined by

$$g(x,y) = \frac{x^2y}{x^4+y^2}$$

and 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 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 may conclude that the limit does not exist. However, we can never conclude that the limit of a function exists only by considering its behavior along different paths.

Generally speaking, concluding that a limit $\lim_{(x,y)\to(a,b)} f(x, y)$ exists requires a more careful argument.

Example 11.2.14 Consider the function f defined by

$$f(x, y) = \frac{x^2y^2}{x^2 + y^2}.$$

We want to know whether $\lim_{(x,y)\to(0,0)} f(x, y)$ exists.

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 what follows, 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

$$0 \le f(x, y) = \frac{x^2y^2}{x^2 + y^2} = x^2\left(\frac{y^2}{x^2 + y^2}\right) \le x^2.$$

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. $\displaystyle\lim_{(x,y)\to(a,b)} x = a$ and $\displaystyle\lim_{(x,y)\to(a,b)} y = b$

2. $\displaystyle\lim_{(x,y)\to(a,b)} cf(x, y) = c\left(\lim_{(x,y)\to(a,b)} f(x, y)\right)$ for any scalar c

3. $\displaystyle\lim_{(x,y)\to(a,b)} [f(x, y) \pm g(x, y)] = \lim_{(x,y)\to(a,b)} f(x, y) \pm \lim_{(x,y)\to(a,b)} g(x, y)$

4. $\displaystyle\lim_{(x,y)\to(a,b)} [f(x, y)g(x, y)] = \left(\lim_{(x,y)\to(a,b)} f(x, y)\right)\left(\lim_{(x,y)\to(a,b)} g(x, y)\right)$

5. $\displaystyle\lim_{(x,y)\to(a,b)} \frac{f(x, y)}{g(x, y)} = \frac{\lim_{(x,y)\to(a,b)} f(x, y)}{\lim_{(x,y)\to(a,b)} g(x, y)}$ if $\displaystyle\lim_{(x,y)\to(a,b)} g(x, y) \neq 0.$

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.

11.2.2 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 11.2.15 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. $\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$.

$\Diamond$

For instance, we have seen that the function f defined by $f(x,y) = 3x^2y^3 + 2xy^2 - 3x + 1$ is continous 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.

11.2.3 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

$$f(a,b) = \lim_{(x,y)\to(a,b)} f(x,y).$$

11.2.4 Exercises

1. Find the limits, if they exist, or type *DNE* for any which do not exist.

$$\lim_{(x,y)\to(0,0)} \frac{1x^2}{2x^2+3y^2}$$

1) Along the x-axis: ____________________
2) Along the y-axis: ____________________
3) Along the line $y = mx$: ______________________________________
4) The limit is: ____________________

2. **Determining the limit of a funtion.** In this problem we show that the function

$$f(x, y) = \frac{2x^2 - y^2}{x^2 + y^2}$$

does not have a limit as $(x, y) \to (0, 0)$.

(a) Suppose that we consider $(x, y) \to (0, 0)$ along the curve $y = 3x$. Find the limit in this case:

$\lim\limits_{(x,3x)\to(0,0)} \frac{2x^2-y^2}{x^2+y^2} =$ ________________________________

(b) Now consider $(x, y) \to (0, 0)$ along the curve $y = 4x$. Find the limit in this case:

$\lim\limits_{(x,4x)\to(0,0)} \frac{2x^2-y^2}{x^2+y^2} =$ ________________________________

(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:

$\lim\limits_{(x,mx)\to(0,0)} \frac{2x^2-y^2}{x^2+y^2} =$ ________________________________

(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.

$$\lim_{(x,y)\to(0,0)} \frac{3x^2}{2x^2 + 4y^2} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

5. Find the limit, if it exists, or type N if it does not exist.

$$\lim_{(x,y)\to(0,0)} \frac{(x + 13y)^2}{x^2 + 13^2 y^2} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

6. Find the limit, if it exists, or type 'DNE' if it does not exist.

$$\lim_{(x,y)\to(1,1)} e^{\sqrt{3x^2+1y^2}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

7. Find the limit, if it exists, or type N if it does not exist.

$$\lim_{(x,y,z)\to(0,0,0)} \frac{4xy + 2yz + 5xz}{16x^2 + 4y^2 + 25z^2} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

8. Find the limit, if it exists, or type N if it does not exist.

$$\lim_{(x,y,z)\to(5,5,2)} \frac{4ze^{x^2+y^2}}{5x^2 + 5y^2 + 2z^2} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

9. Find the limit (enter 'DNE' if the limit does not exist)

Hint: rationalize the denominator.

$$\lim_{(x,y)\to(0,0)} \frac{(9x^2+4y^2)}{\sqrt{(9x^2+4y^2+1)}-1}$$

____

10. The largest set on which the function $f(x,y) = 1/(3-x^2-y^2)$ is continuous is ____________________

A. All of the xy-plane

B. All of the xy-plane except the circle $x^2+y^2=3$

C. The interior of the circle $x^2+y^2=3$, plus the circle

D. The interior of the circle $x^2+y^2=3$

E. The exterior of the circle $x^2+y^2=3$

11. 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).

12. 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).

13. 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 throught 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 $\lim_{(x,y)\to(0,0)} h(x,y)$.

14. 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.

15. 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

$$k(x,y) = \begin{cases} \frac{x^2y^4}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

11.3 First-Order Partial Derivatives

Motivating Questions

- How are the first-order partial derivatives of a function f of the independent variables x and y defined?
- Given a function f of the independent variables x and y, what do the first-order partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ tell us about f?

The derivative plays a central role in first semester calculus because it provides important information about a function. Thinking graphically, for instance, the derivative at a point tells us the slope of the tangent line to the graph at that point. In addition, the derivative at a point also provides the instantaneous rate of change of the function with respect to changes in the independent variable.

Now that we are investigating functions of two or more variables, we can still ask how fast the function is changing, though 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. Alternatively, we may want to know how fast a function's output changes in response to a change in one of the inputs. Over the next few sections, we will develop tools for addressing issues such as these. Preview Activity 11.3.1 explores some issues with what we will come to call *partial derivatives*.

Preview Activity 11.3.1 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. What is the monthly payment if the interest rate is 3% so that $r = 0.03$, and we pay the loan off in $t = 4$ years?

b. Suppose the interest rate 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 left of Figure 11.3.1. Explain the meaning of the function f.

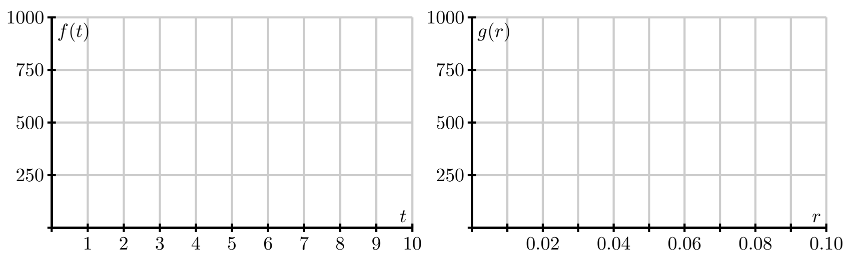


Figure 11.3.1 Left: Graphs $f(t) = M(0.03, t)$. Right: Graph $g(r) = M(r, 4)$.

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. Express M as a function of r alone, using a fixed time of $t = 4$. That is, let $g(r) = M(r, 4)$. Sketch the graph of g on the right of Figure 11.3.1. Explain the meaning of the function g.

e. Find the instantaneous rate of change $g'(0.03)$ and state the units on this quantity. What information does $g'(0.03)$ tell us about our car loan? What information does $g'(0.03)$ tell us about the graph you sketched in (d)?

11.3.1 First-Order Partial Derivatives

In Section 11.1, we studied the behavior of a function of two or more variables by considering the *traces* of the function. Recall that in one example, we considered the function f defined by

$$f(x, y) = \frac{x^2 \sin(2y)}{32},$$

which measures the range, or horizontal distance, in feet, traveled by a projectile launched with an initial speed of x feet per second at an angle y radians to the horizontal. The graph of this function is given again on the left in Figure 11.3.2. Moreover, if we fix the angle $y = 0.6$, we may view the trace $f(x, 0.6)$ as a function of x alone, as seen at right in Figure 11.3.2.

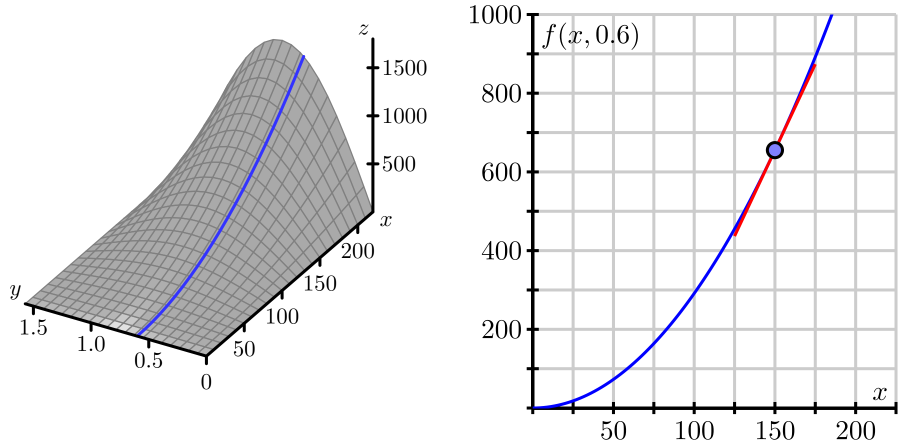


Figure 11.3.2 Left: The trace of $z = \frac{x^2 \sin(2y)}{32}$ with $y = 0.6$.

Since the trace is a one-variable function, we may consider its derivative just as we did in the first semester of calculus. With $y = 0.6$, we have

$$f(x, 0.6) = \frac{\sin(1.2)}{32} x^2,$$

and therefore

$$\frac{d}{dx}[f(x, 0.6)] = \frac{\sin(1.2)}{16} x.$$

When $x = 150$, this gives

$$\frac{d}{dx}[f(x, 0.6)]|_{x=150} = \frac{\sin(1.2)}{16} 150 \approx 8.74 \text{ feet per feet per second},$$

which gives the slope of the tangent line shown on the right of Figure 11.3.2. Thinking of this derivative as an instantaneous rate of change implies that if we increase the initial speed of the projectile by one foot per second, we expect the horizontal distance traveled to increase by approximately 8.74 feet if we hold 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

$$f_x(150, 0.6) = \frac{d}{dx} f(x, 0.6)|_{x=150} = \lim_{h \to 0} \frac{f(150 + h, 0.6) - f(150, 0.6)}{h}.$$

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 11.3.3.

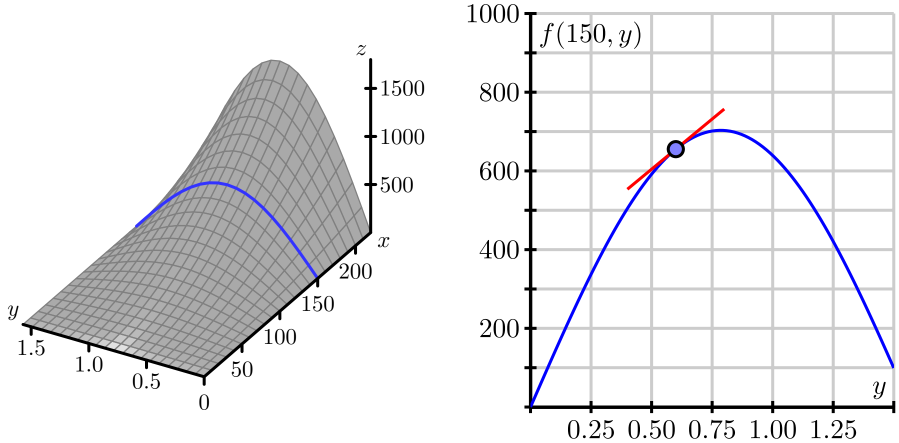


Figure 11.3.3 The trace of $z = \frac{x^2 \sin(2y)}{32}$ with $x = 150$.

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

$$\frac{d}{dy}[f(150, y)]|_{y=0.6} = \frac{150^2}{16} \cos(1.2) \approx 509.5 \text{ feet per radian.}$$

Once again, the derivative gives the slope of the tangent line shown on the right in Figure 11.3.3. Thinking of the derivative as an instantaneous rate of change, we expect that the range of the projectile increases by 509.5 feet for every radian we increase the launch angle y if we keep the initial speed of the projectile constant at 150 feet per second.

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

$$f_y(150, 0.6) = \frac{d}{dy} f(150, y)|_{y=0.6} = \lim_{h \to 0} \frac{f(150, 0.6 + h) - f(150, 0.6)}{h}.$$

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 11.3.4 The first-order **partial derivatives of f with respect to x and y** at a point (a, b) are, respectively,

$$f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}, \text{ and}$$
$$f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h},$$

provided the limits exist. ◊

Activity 11.3.2 Consider the function f defined by

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

at the point $(1, 2)$.

a. Write the trace $f(x, 2)$ at the fixed value $y = 2$. On the left side of Figure 11.3.5, draw the graph of the trace with $y = 2$ around the point where $x = 1$, indicating the scale and labels on the axes. Also, sketch the tangent line at the point $x = 1$.

Figure 11.3.5 Traces of $f(x, y) = \frac{xy^2}{x+1}$.

b. Find the partial derivative $f_x(1, 2)$ and relate its value to the sketch you just made.

c. Write the trace $f(1, y)$ at the fixed value $x = 1$. On the right side of Figure 11.3.5, draw the graph of the trace with $x = 1$ indicating the scale and labels on the axes. Also, sketch the tangent line at the point $y = 2$.

d. Find the partial derivative $f_y(1, 2)$ and relate its value to the sketch you just made.

As these examples show, each partial derivative at a point arises as the derivative of a one-variable function defined by fixing one of the coordinates. In addition, we may consider each partial derivative as defining a new function of 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 one-variable derivatives and partial derivatives, we will often use Leibniz-style notation to denote partial derivatives by writing

$$\frac{\partial f}{\partial x}(a, b) = f_x(a, b), \text{ and } \frac{\partial f}{\partial y}(a, b) = f_y(a, b).$$

To calculate the partial derivative f_x, we hold y fixed and thus 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

$$\frac{d}{dx}[3x^2 - 2x + 3] = 3\frac{d}{dx}[x^2] - 2\frac{d}{dx}[x] + \frac{d}{dx}[3] = 3 \cdot 2x - 2,$$

$$\text{and } \frac{\partial}{\partial x}[x^2y - xy + 2y] = y\frac{\partial}{\partial x}[x^2] - y\frac{\partial}{\partial x}[x] + \frac{\partial}{\partial x}[2y] = y \cdot 2x - y$$

Thus, computing partial derivatives is straightforward: we use the standard rules of single variable calculus, but do so while holding one (or more) of the variables constant.

Activity 11.3.3

a. If $f(x, y) = 3x^3 - 2x^2y^5$, find the partial derivatives f_x and f_y.

b. If $f(x, y) = \dfrac{xy^2}{x + 1}$, find the partial derivatives f_x and f_y.

c. If $g(r, s) = rs\cos(r)$, find the partial derivatives g_r and g_s.

d. Assuming $f(w, x, y) = (6w + 1)\cos(3x^2 + 4xy^3 + y)$, find the partial derivatives f_w, f_x, and f_y.

e. Find all possible first-order partial derivatives of $q(x, t, z) = \dfrac{x2^tz^3}{1 + x^2}$.

11.3.2 Interpretations of First-Order Partial Derivatives

Recall that the derivative of a single variable function has a geometric interpretation 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 11.3.6.

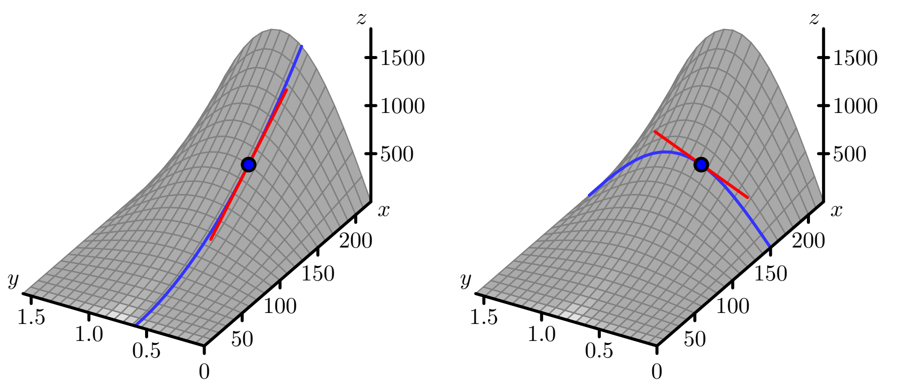


Figure 11.3.6 Tangent lines to two traces of the distance function.

Now we consider the first-order partial derivatives in context. Recall that the difference quotient $\frac{f(a+h)-f(a)}{h}$ for a function f of a single variable x at a point where $x = a$ 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 these same concepts to explain the meanings of the partial derivatives in context.

Activity 11.3.4 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 in which each of the partial derivatives, C_T, C_S and C_D, are expressed and explain the physical meaning of each.

b. 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 $D = 100$. What does the sign of each partial derivatives tell us about the behavior of the function C at the point $(10, 35, 100)$?

11.3.3 Using tables and contours to estimate partial derivatives

Remember that functions of two variables are often represented as either a table of data or a contour plot. In single variable calculus, we saw how we can use the difference quotient to approximate derivatives if, instead of an algebraic formula, we only know the value of the function at a few points. The same idea applies to partial derivatives.

Activity 11.3.5 The wind chill, as frequently reported, is a measure of how cold it feels outside when the wind is blowing. In Table 11.3.7, the wind chill w, 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. We thus view w as being of the form $w = w(v, T)$.

Table 11.3.7 Wind chill as a function of wind speed and temperature.

$v\backslash T$	-30	-25	-20	-15	-10	-5	0	5	10	15	20
5	-46	-40	-34	-28	-22	-16	-11	-5	1	7	13
10	-53	-47	-41	-35	-28	-22	-16	-10	-4	3	9
15	-58	-51	-45	-39	-32	-26	-19	-13	-7	0	6
20	-61	-55	-48	-42	-35	-29	-22	-15	-9	-2	4
25	-64	-58	-51	-44	-37	-31	-24	-17	-11	-4	3
30	-67	-60	-53	-46	-39	-33	-26	-19	-12	-5	1
35	-69	-62	-55	-48	-41	-34	-27	-21	-14	-7	0
40	-71	-64	-57	-50	-43	-36	-29	-22	-15	-8	-1

a. Estimate the partial derivative $w_v(20, -10)$. What are the units on this quantity and what does it mean? (Recall that we can estimate a partial derivative of a single variable function f using the symmetric difference quotient $\frac{f(x+h)-f(x-h)}{2h}$ for small values of h. A partial derivative is a derivative of an appropriate trace.)

b. Estimate the partial derivative $w_T(20, -10)$. What are the units on this quantity and what does it mean?

c. Use your results to estimate the wind chill $w(18, -10)$. (Recall from single variable calculus that for a function f of x, $f(x + h) \approx f(x) + hf'(x)$.)

d. Use your results to estimate the wind chill $w(20, -12)$.

e. Consider how you might combine your previous results to estimate the wind chill $w(18, -12)$. Explain your process.

Activity 11.3.6 Shown below in Figure 11.3.8 is a contour plot of a function f. The values of the function on a few of the contours are indicated to the left of the figure.

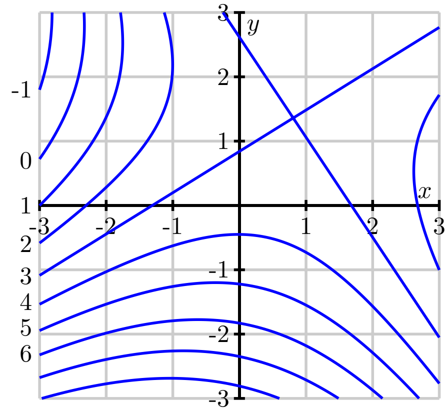


Figure 11.3.8 A contour plot of f.

a. Estimate the partial derivative $f_x(-2,-1)$. (Hint: How can you find values of f that are of the form $f(-2+h)$ and $f(-2-h)$ so that you can use a symmetric difference quotient?)

b. Estimate the partial derivative $f_y(-2,-1)$.

c. Estimate the partial derivatives $f_x(-1,2)$ and $f_y(-1,2)$.

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 $f_y(x,y)>0$.

g. Suppose you have a different function g, and you know that $g(2,2)=4$, $g_x(2,2)>0$, and $g_y(2,2)>0$. Using this information, sketch a possibility for the contour $g(x,y)=4$ passing through $(2,2)$ on the left side of Figure 11.3.9. Then include possible contours $g(x,y)=3$ and $g(x,y)=5$.

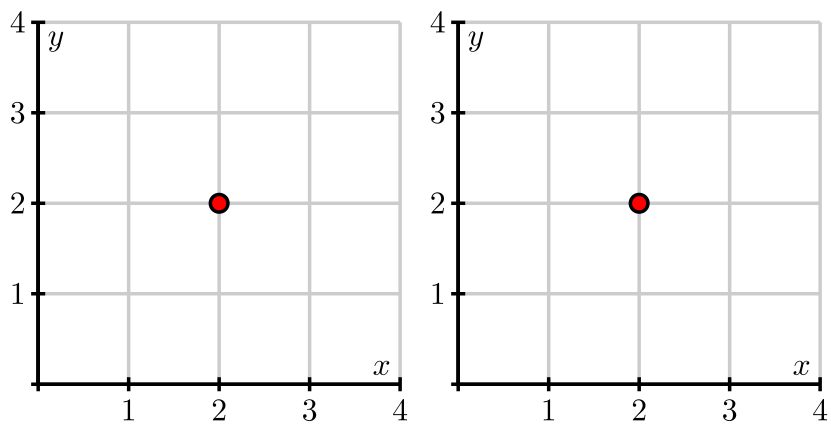


Figure 11.3.9 Plots for contours of g and h.

h. Suppose you have yet another function h, and you know that $h(2,2)=4$, $h_x(2,2)<0$, and $h_y(2,2)>0$. Using this information, sketch a possible contour $h(x,y)=4$ passing through $(2,2)$ on the right side of Figure 11.3.9. Then include possible contours $h(x,y)=3$ and $h(x,y)=5$.

11.3.4 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 $(a, b, f(a, b))$.

11.3.5 Exercises

1. Find the first partial derivatives of
$f(x, y) = \dfrac{x - 4y}{x + 4y}$ at the point $(x, y) = (3, 3)$.
$\dfrac{\partial f}{\partial x}(3, 3) =$ ______________________
$\dfrac{\partial f}{\partial y}(3, 3) =$ ______________________

2. Find the first partial derivatives of $f(x, y) = \sin(x - y)$ at the point (-4, -4).
A. $f_x(-4, -4) =$ ______________________
B. $f_y(-4, -4) =$ ______________________

3. Find the partial derivatives of the function
$$w = \sqrt{4r^2 + 2s^2 + 5t^2}$$
$\frac{\partial w}{\partial r} =$ ______________________
$\frac{\partial w}{\partial s} =$ ______________________
$\frac{\partial w}{\partial t} =$ ______________________

4. Suppose that $f(x, y)$ is a smooth function and that its partial derivatives have the values, $f_x(0, 6) = 4$ and $f_y(0, 6) = 4$. Given that $f(0, 6) = -1$, 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) $f(0, 7)$ ______________
Estimate of (integer value) $f(1, 6)$ ______________
Estimate of (integer value) $f(1, 7)$ ______________

5. 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)

6. Find the first partial derivatives of $f(x, y, z) = z\ \arctan(\frac{y}{x})$ at the point (4, 4, -4).

A. $\frac{\partial f}{\partial x}(4, 4, -4) =$ ______

B. $\frac{\partial f}{\partial y}(4, 4, -4) =$ ______

C. $\frac{\partial f}{\partial z}(4, 4, -4) =$ ______

7. Find the partial derivatives of the function

$$f(x, y) = \int_y^x \cos(7t^2 + 5t - 7)\, dt$$

$f_x(x, y) =$ ______

$f_y(x, y) =$ ______

8. Let $f(x, y) = e^{-2x} \sin(3y)$.

(a) Using difference quotients with $\Delta x = 0.1$ and $\Delta y = 0.1$, we estimate

$f_x(2, -1) \approx$ ______

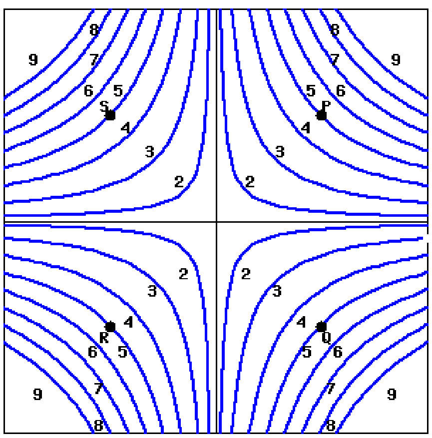
$f_y(2, -1) \approx$ ______

(b) Using difference quotients with $\Delta x = 0.01$ and $\Delta y = 0.01$, we find better estimates:

$f_x(2, -1) \approx$ ______

$f_y(2, -1) \approx$ ______

9. 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 Q,

f_x is (☐ positive ☐ negative) and

f_y is (☐ positive ☐ negative) .

(b) At point R,

f_x is (☐ positive ☐ negative) and

f_y is (☐ positive ☐ negative) .

(c) At point S,
f_x is (☐ positive ☐ negative) and
f_y is (☐ positive ☐ negative) .

10. 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 15.
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.

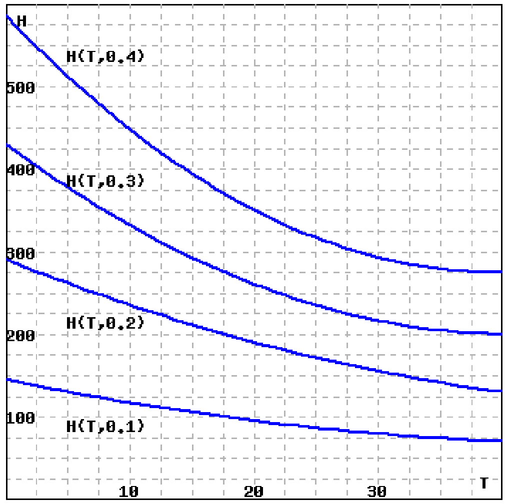
11. An experiment to measure the toxicity of formaldehyde yielded the data in the table below. The values show the percent, $P = f(t, c)$, of rats surviving an exposure to formaldehyde at a concentration of c (in parts per million, ppm) after t months.

	$t = 14$	$t = 16$	$t = 18$	$t = 20$	$t = 22$	$t = 24$
$c = 0$	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 $f_t(18, 2)$:
$f_t(18, 2) \approx$ ______________________________
(b) Estimate $f_c(18, 2)$:
$f_c(18, 2) \approx$ ______________________________

(Be sure that you can give the practical meaning of these two values in terms of formaldehyde toxicity.)

12. 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:
$T = 10, w = 0.2:\ H_T(T, w) \approx$ ______________________
$T = 20, w = 0.2:\ H_T(T, w) \approx$ ______________________
$T = 10, w = 0.3:\ H_T(T, w) \approx$ ______________________
$T = 20, w = 0.3:\ H_T(T, w) \approx$ ______________________

(c) Repeat ***(b)*** to find $H_w(T, w)$ for each of the following:
$T = 10, w = 0.2:\ H_w(T, w) \approx$ ______________________
$T = 20, w = 0.2:\ H_w(T, w) \approx$ ______________________
$T = 10, w = 0.3:\ H_w(T, w) \approx$ ______________________
$T = 20, w = 0.3:\ H_w(T, w) \approx$ ______________________
(Be sure you can interpret this partial derivative in practical terms.)

13. 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 11.3.10.

Table 11.3.10 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.

14. 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 11.3.11) 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.

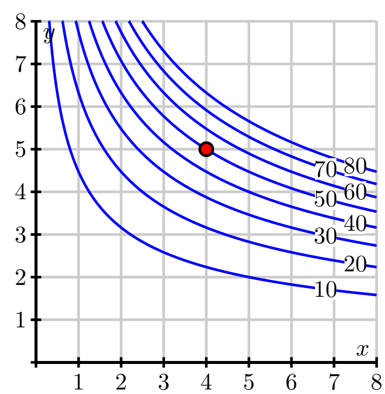


Figure 11.3.11 The graph of $f(x,y) = \frac{1}{2}xy^2$.

15. 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 $\frac{\partial C}{\partial x}|_{(x,y)}$ and $\frac{\partial C}{\partial y}|_{(x,y)}$.

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?

16. 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?

17. 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.

11.4 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?

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, $f''(x) = \frac{d}{dx}[f'(x)]$, 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 what follows, we begin exploring the four different second-order partial derivatives of a function of two variables and seek to understand what these various derivatives tell us about the function's behavior.

Preview Activity 11.4.1 Once again, let's 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 Figure 11.4.1.

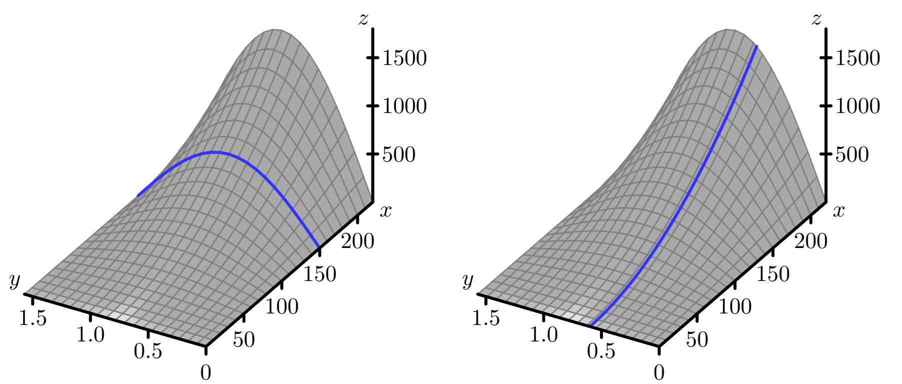


Figure 11.4.1 The distance function with traces $x = 150$ and $y = 0.6$.

a. Compute the partial derivative f_x. 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 $f_{xx} = (f_x)_x$ and show that $f_{xx}(150, 0.6) \approx 0.058$.

b. Figure 11.4.2 shows the trace of f with $y = 0.6$ with three tangent lines included. Explain how your result from part (a) of this preview activity is reflected in this figure.

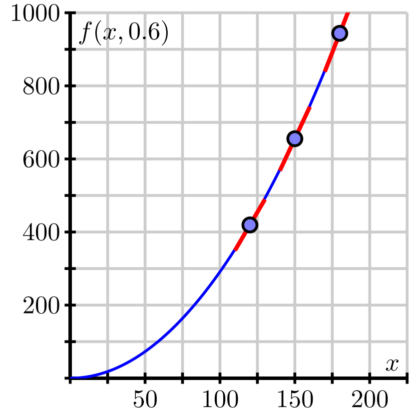


Figure 11.4.2 The trace with $y = 0.6$.

c. Determine the partial derivative f_y, and then find the partial derivative $f_{yy} = (f_y)_y$. Evaluate $f_{yy}(150, 0.6)$.

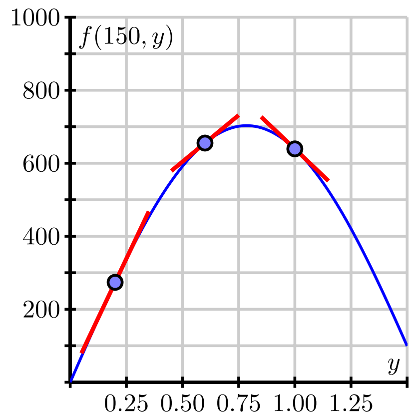


Figure 11.4.3 More traces of the range function.

d. Figure 11.4.3 shows the trace $f(150, y)$ and includes three tangent lines. Explain how the value of $f_{yy}(150, 0.6)$ is reflected in this figure.

e. Because f_x and f_y are each functions of both x and y, they each have two partial derivatives. Not only can we compute $f_{xx} = (f_x)_x$, but also $f_{xy} = (f_x)_y$; likewise, in addition to $f_{yy} = (f_y)_y$, but also $f_{yx} = (f_y)_x$. For the range function $f(x, y) = \frac{x^2 \sin(2y)}{32}$, use your earlier computations of f_x and f_y to now determine f_{xy} and f_{yx}. Write one sentence to explain how you calculated these "mixed" partial derivatives.

11.4.1 Second-Order Partial Derivatives

A function f of two independent variables x and y has two first order partial derivatives, f_x and f_y. As we saw in Preview Activity 11.4.1, each of these first-order partial derivatives has two partial derivatives, giving a total of four *second-order* partial derivatives:

- $f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$,
- $f_{yy} = (f_y)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$,

- $f_{xy} = (f_x)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$,
- $f_{yx} = (f_y)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$.

The first two are called *unmixed* second-order partial derivatives while the last two are called the *mixed* second-order partial derivatives.

One aspect of this notation can be a little confusing. The notation

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

means that we first differentiate with respect to x and then 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 x. This means that

$$\frac{\partial^2 f}{\partial y \partial x} = f_{xy}, \text{ and } \frac{\partial^2 f}{\partial x \partial y} = f_{yx}.$$

Be sure to note carefully the difference between Leibniz notation and subscript notation and the order in which x and y appear in each. In addition, remember that anytime we compute a partial derivative, we hold constant the variable(s) other than the one we are differentiating with respect to.

Activity 11.4.2 Find all second order partial derivatives of the following functions. For each partial derivative you calculate, state explicitly which variable is being held constant.

a. $f(x, y) = x^2 y^3$

b. $f(x, y) = y\cos(x)$

c. $g(s, t) = st^3 + s^4$

d. How many second order partial derivatives does the function h defined by $h(x, y, z) = 9x^9 z - xyz^9 + 9$ have? Find h_{xz} and h_{zx} (you do not need to find the other second order partial derivatives).

In Preview Activity 11.4.1 and Activity 11.4.2, you may have noticed that the mixed second-order partial derivatives are equal. This observation holds generally and is known as Clairaut's Theorem.

Clairaut's Theorem.

Let f be a function of several variables for which the partial derivatives f_{xy} and f_{yx} are continuous near the point (a, b). Then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

11.4.2 Interpreting the Second-Order Partial Derivatives

Recall from single variable calculus that the second derivative measures the instantaneous rate of change of the derivative. This observation is the key to understanding the meaning of the second-order partial derivatives.

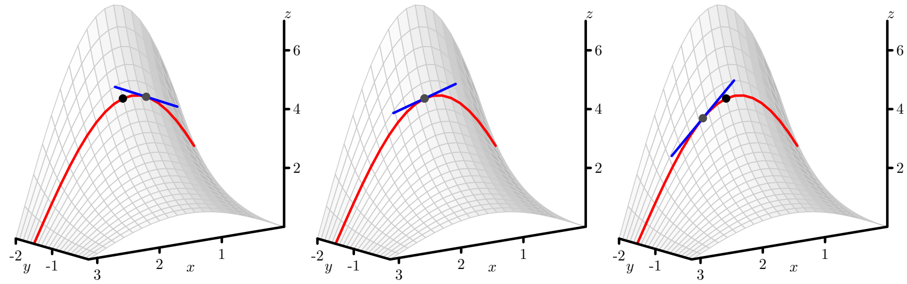


Figure 11.4.4 The tangent lines to a trace with increasing x.

Furthermore, we remember that the second derivative of a function at a point provides us with information about the concavity of the function at that point. Since the unmixed second-order partial derivative f_{xx} requires us to hold y constant and differentiate twice with respect to x, we may simply 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.

Consider, for example, $f(x,y) = \sin(x)e^{-y}$. Figure 11.4.4 shows the graph of this function along with the trace given by $y = -1.5$. Also shown are three tangent lines to this trace, with increasing x-values from left to right among the three plots in Figure 11.4.4.

That the slope of the tangent line is decreasing as x increases is reflected, as it is in one-variable calculus, in the fact that the trace is concave down. Indeed, we see that $f_x(x,y) = \cos(x)e^{-y}$ and so $f_{xx}(x,y) = -\sin(x)e^{-y} < 0$, since $e^{-y} > 0$ for all values of y, including $y = -1.5$.

In the following activity, we further explore what second-order partial derivatives tell us about the geometric behavior of a surface.

Activity 11.4.3 We continue to consider the function f defined by $f(x,y) = \sin(x)e^{-y}$.

a. In Figure 11.4.5, we see the trace of $f(x,y) = \sin(x)e^{-y}$ that has x held constant with $x = 1.75$. We also see three different lines that are tangent to the trace of f in the y direction at values of y that are increasing from left to right in the figure. Write a couple of sentences that describe whether the slope of the tangent lines to this curve increase or decrease as y increases, and, after computing $f_{yy}(x,y)$, explain how this observation is related to the value of $f_{yy}(1.75, y)$. Be sure to address the notion of concavity in your response.(You need to be careful about the directions in which x and y are increasing.)

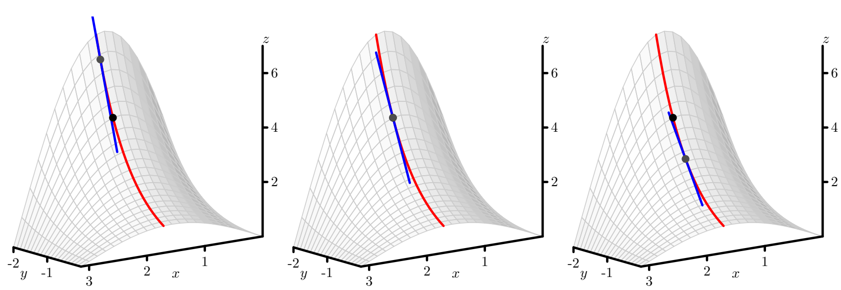


Figure 11.4.5 The tangent lines to a trace with increasing y.

b. In Figure 11.4.6, we start to think about the mixed partial derivative, f_{xy}. Here, we first hold y constant to generate the first-order partial derivative f_x, and then we hold x constant to compute f_{xy}. This leads to first thinking about a trace with x being constant, followed by slopes of tangent lines in the x-direction that slide along the original trace. You might think of sliding your pencil down the trace with x constant in a way that its slope indicates $(f_x)_y$ in order to further animate the three snapshots shown in the figure.

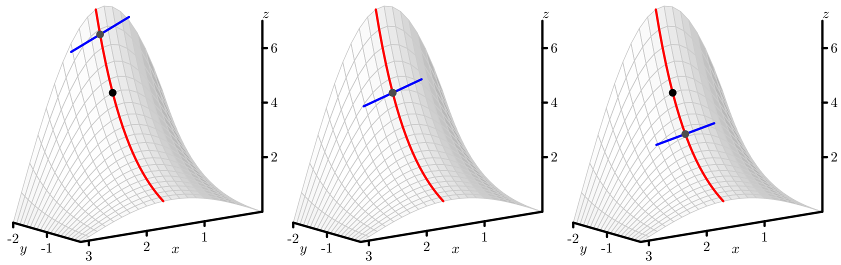


Figure 11.4.6 The trace of $z = f(x, y) = \sin(x)e^{-y}$ with $x = 1.75$, along with tangent lines in the y-direction at three different points.

Based on Figure 11.4.6, is $f_{xy}(1.75, -1.5)$ positive or negative? Why?

c. Determine the formula for $f_{xy}(x, y)$, and hence evaluate $f_{xy}(1.75, -1.5)$. How does this value compare with your observations in (b)?

d. 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?

e. On Figure 11.4.6, sketch the trace with $y = -1.5$, and sketch three tangent lines whose slopes correspond to the value of $f_{yx}(x, -1.5)$ for three different values of x, the middle of which is $x = -1.5$. Is $f_{yx}(1.75, -1.5)$ positive or negative? Why? What does $f_{yx}(1.75, -1.5)$ measure?

Just as with the first-order partial derivatives, we can approximate second-order partial derivatives in the situation where we have only partial information about the function.

Activity 11.4.4 As we saw in Activity 11.3.5, 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 Table 11.4.7.

Table 11.4.7 Wind chill as a function of wind speed and temperature.

$v\backslash T$	-30	-25	-20	-15	-10	-5	0	5	10	15	20
5	-46	-40	-34	-28	-22	-16	-11	-5	1	7	13
10	-53	-47	-41	-35	-28	-22	-16	-10	-4	3	9
15	-58	-51	-45	-39	-32	-26	-19	-13	-7	0	6
20	-61	-55	-48	-42	-35	-29	-22	-15	-9	-2	4
25	-64	-58	-51	-44	-37	-31	-24	-17	-11	-4	3
30	-67	-60	-53	-46	-39	-33	-26	-19	-12	-5	1
35	-69	-62	-55	-48	-41	-34	-27	-21	-14	-7	0
40	-71	-64	-57	-50	-43	-36	-29	-22	-15	-8	-1

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 several sentences that explain what the values $w_{TT}(20,-10)$, $w_{vv}(20,-10)$, and $w_{Tv}(20,-10)$ indicate regarding the behavior of $w(v,T)$.

As we have found in Activities 11.4.3 and Activity 11.4.4, 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 $f_{xy} = f_{yx}$, we see that the amount of twisting is the same in both directions. This twisting is perhaps more easily seen in Figure 11.4.8, which shows the graph of $f(x,y) = -xy$, for which $f_{xy} = -1$.

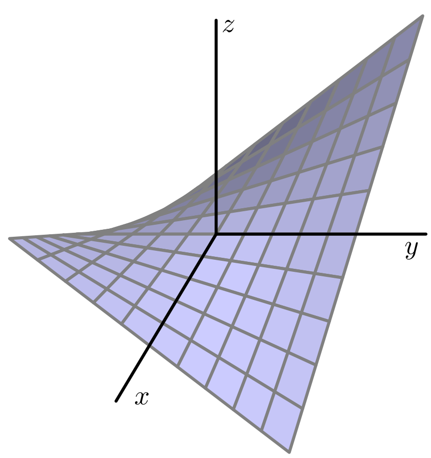


Figure 11.4.8 The graph of $f(x,y) = -xy$.

11.4.3 Summary

- There are four second-order partial derivatives of a function f of two independent variables x and y:

$$f_{xx} = (f_x)_x, f_{xy} = (f_x)_y, f_{yx} = (f_y)_x, \text{ and } f_{yy} = (f_y)_y.$$

- 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.

11.4.4 Exercises

1. Calculate all four second-order partial derivatives of $f(x,y) = 4x^2y + 7xy^3$.

$f_{xx}(x,y) =$ ______

$f_{xy}(x,y) =$ ______

$f_{yx}(x,y) =$ ______

$f_{yy}(x,y) =$ ______

2. Find all the first and second order partial derivatives of $f(x,y) = 3\sin(2x+y) - 3\cos(x-y)$.

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

$$f(x,y) = xye^{4y}$$

$f_x(x,y) =$ ______

$f_y(x,y) =$ ______

$f_{xy}(x,y) =$ ______

$f_{yx}(x,y) =$ ______

4. Calculate all four second-order partial derivatives of $f(x,y) = \sin\left(\frac{5x}{y}\right)$.

$f_{xx}(x,y) =$ ______

$f_{xy}(x,y) =$ ______

$f_{yx}(x,y) =$ ______

$f_{yy}(x,y) =$ ______

5. Given $F(r,s,t) = r\left(9s^3 - 4t^3\right)$, compute:

$F_{rst} =$ ______

6. 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.

$$f(x,y) = e^{2xy}$$

$f_{xx} =$ ______

$f_{yy} =$ ______

$f_{xy} =$ ______

$f_{yx} =$ ______

7. Calculate all four second-order partial derivatives of $f(x, y) = (2x + 4y)\, e^{y}$.

$f_{xx}\,(x, y) =$ ____________________

$f_{xy}\,(x, y) =$ ____________________

$f_{yx}\,(x, y) =$ ____________________

$f_{yy}\,(x, y) =$ ____________________

8. Let $f(x, y) = (-(2x + y))^6$. Then

$$\frac{\partial^2 f}{\partial x \partial y} = \_\_\_\_\_\_\_\_\_\_$$

$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \_\_\_\_\_\_\_\_\_\_$$

$$\frac{\partial^3 f}{\partial x^2 \partial y} = \_\_\_\_\_\_\_\_\_\_$$

9. If $z_{xy} = 5y$ and all of the second order partial derivatives of z are continuous, then

(a) $z_{yx} =$ ____________

(b) $z_{xyx} =$ ____________

(c) $z_{xyy} =$ ____________

10. If $z = f(x) + yg(x)$, what can we say about z_{yy}?

- $z_{yy} = 0$
- $z_{yy} = g(x)$
- $z_{yy} = z_{xx}$
- $z_{yy} = y$
- We cannot say anything

11. Shown in Figure 11.4.9 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.

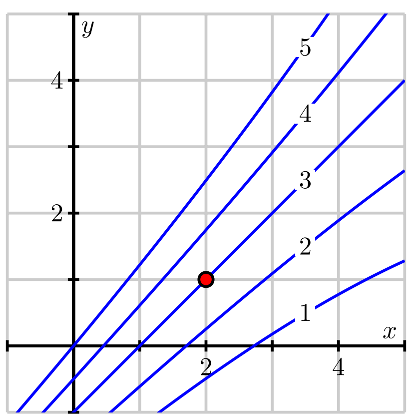


Figure 11.4.9 A contour plot of $f(x, y)$.

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 11.4.10.

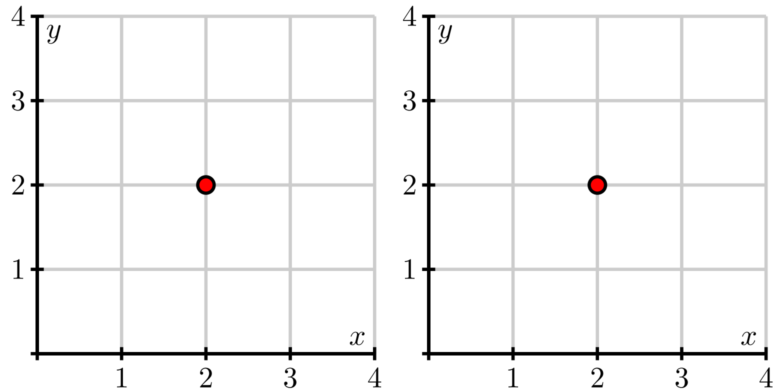


Figure 11.4.10 Plots for contours of g and h.

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 11.4.10.

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 11.4.11.

Table 11.4.11 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

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.

11.5 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?
- What is the differential of a multivariable function of two variables and what are its uses?

One of the central concepts in single variable calculus is that the graph of a differentiable function, when viewed on a very small scale, looks like a line. We call this line the tangent line and measure its slope with the derivative. In this section, we will extend this concept to functions of several variables.

Let's see what happens when we look at the graph of a two-variable function on a small scale. To begin, let's consider the function f defined by

$$f(x, y) = 6 - \frac{x^2}{2} - y^2,$$

whose graph is shown in Figure 11.5.1.

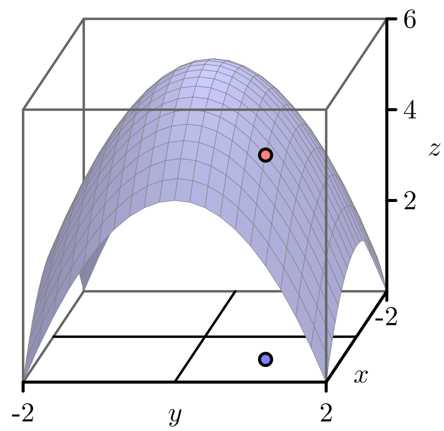


Figure 11.5.1 The graph of $f(x, y) = 6 - x^2/2 - y^2$.

We choose to study the behavior of this function near the point $(x_0, y_0) = (1, 1)$. In particular, we wish to view the graph on an increasingly small scale around this point, as shown in the two plots in Figure 11.5.2

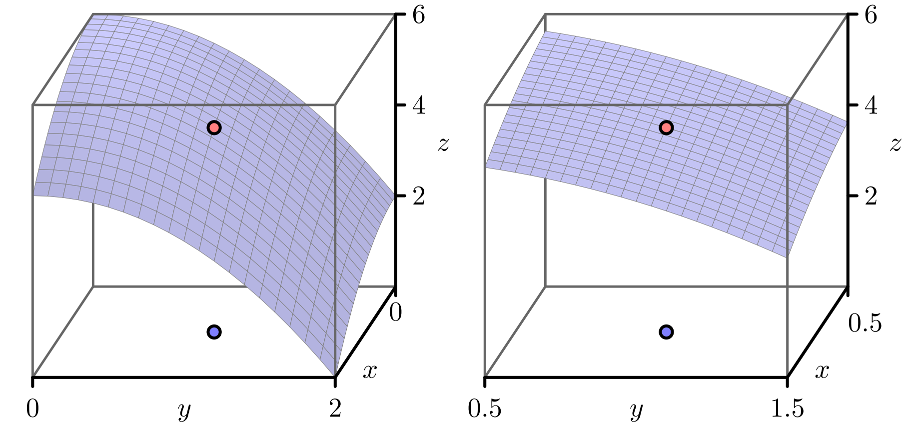


Figure 11.5.2 The graph of $f(x, y) = 6 - x^2/2 - y^2$.

Just as the graph of a differentiable single-variable function looks like a line when viewed on a small scale, we see that the graph of this particular two-variable function looks like a plane, as seen in Figure 11.5.3. In the following preview activity, we explore how to find the equation of this plane.

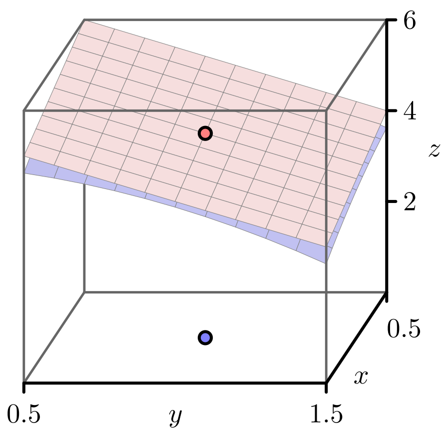


Figure 11.5.3 The graph of $f(x, y) = 6 - x^2/2 - y^2$.

In what follows, we will also use the important fact[1] that the plane passing through (x_0, y_0, z_0) may be expressed in the form $z = z_0 + a(x - x_0) + b(y - y_0)$, where a and b are constants.

Preview Activity 11.5.1 Let $f(x, y) = 6 - \frac{x^2}{2} - y^2$, and let $(x_0, y_0) = (1, 1)$.

a. Evaluate $f(x, y) = 6 - \frac{x^2}{2} - y^2$ and its partial derivatives at (x_0, y_0); that is, find $f(1, 1)$, $f_x(1, 1)$, and $f_y(1, 1)$.

b. We know one point on the tangent plane; namely, the z-value of the tangent plane agrees with the z-value on the graph of $f(x, y) = 6 - \frac{x^2}{2} - y^2$ at the point (x_0, y_0). In other words, both the tangent plane and the graph

[1] As we saw in Section 9.6, 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 and hence 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.

of the function f contain the point (x_0, y_0, z_0). Use this observation to determine z_0 in the expression $z = z_0 + a(x - x_0) + b(y - y_0)$.

c. 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 in Figure 11.5.4.

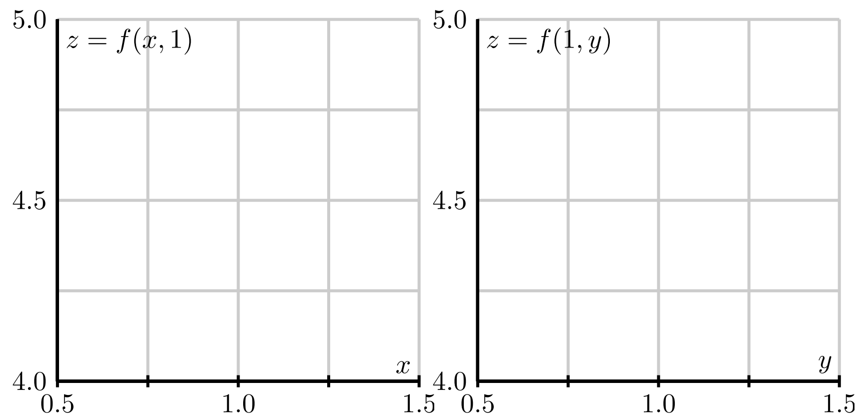


Figure 11.5.4 The traces of $f(x, y)$ with $y = y_0 = 1$ and $x = x_0 = 1$.

d. Determine the equation of the tangent line of the trace that you sketched in the previous part with $y = 1$ (in the x direction) at the point $x_0 = 1$.

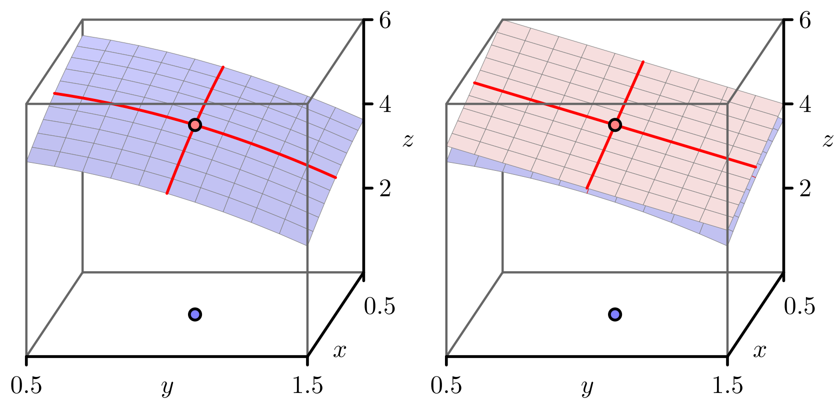


Figure 11.5.5 The traces of $f(x, y)$ and the tangent plane.

e. Figure 11.5.5 shows the traces of the function and the traces of the tangent plane. Explain how the tangent line of the trace of f, whose equation you found in the last part of this activity, is related to the tangent plane. How does this observation help you determine the constant a in the equation for the tangent plane $z = z_0 + a(x - x_0) + b(y - y_0)$? (Hint: How do you think $f_x(x_0, y_0)$ should be related to $z_x(x_0, y_0)$?)

f. In a similar way to what you did in (d), determine the equation of the tangent line of the trace with $x = 1$ at the point $y_0 = 1$. Explain how this tangent line is related to the tangent plane, and use this observation to determine the constant b in the equation for the tangent plane $z = z_0 + a(x - x_0) + b(y - y_0)$. (Hint: How do you think $f_y(x_0, y_0)$ should be related to $z_y(x_0, y_0)$?)

g. Finally, write the equation $z = z_0 + a(x - x_0) + b(y - y_0)$ of the tangent plane to the graph of $f(x, y) = 6 - x^2/2 - y^2$ at the point $(x_0, y_0) = (1, 1)$.

11.5.1 The Tangent Plane

Before stating the formula for the equation of the tangent plane at a point for a general function $f = f(x, y)$, we need to discuss a technical condition. As we have noted, when we look at the graph of a single-variable function on a small scale near a point x_0, we expect to see a line; in this case, we say that f *is locally linear near* x_0 since the graph looks like a linear function locally around x_0. Of course, there are functions, such as the absolute value function given by $f(x) = |x|$, that are not locally linear at every point. In single-variable calculus, we learn that if the derivative of a function exists at a point, then the function is guaranteed to be locally linear there.

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). How can we tell when a function of two variables is locally linear at a point?

It is not unreasonable to expect that if $f_x(a, b)$ and $f_y(a, b)$ exist for some function f at a point (a, b), then f is locally linear at (a, b). This is not sufficient, however. As an example, consider the function f defined by $f(x, y) = x^{1/3}y^{1/3}$. In Exercise 11.5.5.11 you are asked to show that $f_x(0, 0)$ and $f_y(0, 0)$ both exist, but that f is not locally linear at $(0, 0)$ (see Figure **??**). So the existence of the two first order partial derivatives at a point does not guarantee local linearity at that point.

It would take us too far afield to provide a rigorous dicussion of differentiability of functions of more than one variable (see Exercise 11.5.5.15 for a little more detail), so we will be content to define stronger, but more easily verified, conditions that ensure local linearity.

Differentiablity.

If f is a function of the independent variables x and y and 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).

As a consequence, whenever a function $z = f(x, y)$ is continuously differentiable at a point (x_0, y_0), it follows that the function has a tangent plane at (x_0, y_0). Viewed up close, the tangent plane and the function are then virtually indistinguishable. (We won't formally define differentiability of multivariable functions here, and for our purposes continuous differentiability is the only condition we will ever need to use. It is important to note that continuous differentiability is a stronger condition than differentiability. All of the results we encounter will apply to differentiable functions, and so also apply to continuously differentiable functions.) In addition, as in Preview Activity 11.5.1, we find the following general formula for the tangent plane.

The tangent plane.

If $f(x, y)$ has continuous first-order partial derivatives, 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

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{11.5.1}$$

Important Note: As can be seen in Exercise 11.5.5.11, it is possible that $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ can exist for a function f, and so the plane $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$ exists even though f is not locally linear at (x_0, y_0) (because the graph of f does not look linear when we zoom in around the point (x_0, y_0)). In such a case this plane is not tangent to the graph. Differentiability for a function of two variables implies the existence of a tangent plane, but the existence of the two first order partial derivatives of a function at a point does not imply differentiaility. This is quite different than what happens in single variable calculus.

Finally, one important note about the form of the equation for the tangent plane, $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$. Say, for example, that we have the particular tangent plane $z = 7-2(x-3)+4(y+1)$. Observe that we can immediately read from this form that $f_x(3, -1) = -2$ and $f_y(3, -1) = 4$; furthermore, $f_x(3, -1) = -2$ is the slope of the trace to both f and the tangent plane in the x-direction at $(3, -1)$. In the same way, $f_y(3, -1) = 4$ is the slope of the trace of both f and the tangent plane in the y-direction at $(3, -1)$.

Activity 11.5.2

a. Find the equation of the tangent plane to $f(x, y) = 2 + 4x - 3y$ at the point $(1, 2)$. Simplify as much as possible. Does the result surprise you? Explain.

b. Find the equation of the tangent plane to $f(x, y) = x^2y$ at the point $(1, 2)$.

11.5.2 Linearization

In single variable calculus, an important use of the tangent line is to approximate the value of 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 near x_0, as shown in Figure 11.5.6.

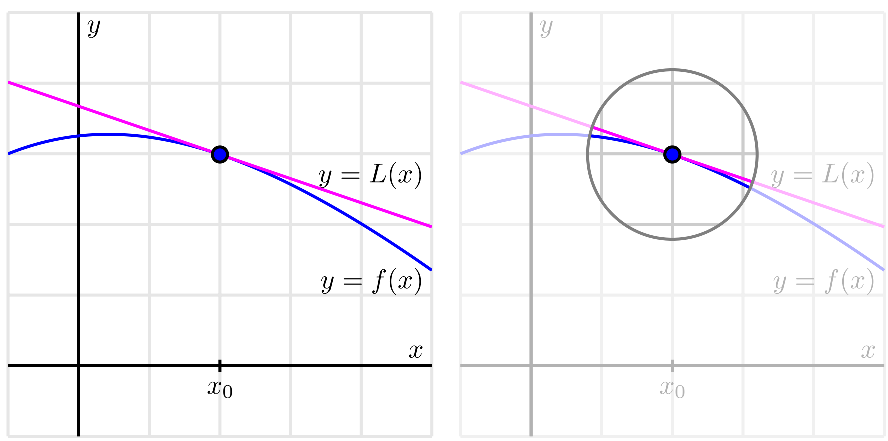


Figure 11.5.6 The linearization of the single-variable function $f(x)$.

In this single-variable setting, 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, observe that $f(x) \approx L(x)$ near x_0. We call L the *linearization* of f.

In the same way, the tangent plane to the graph of a differentiable function $z = f(x, y)$ at a point (x_0, y_0) provides a good approximation of $f(x, y)$ near (x_0, y_0). Here, we define the linearization, L, to be the two-variable function whose graph is the tangent plane, and thus

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Finally, note that $f(x, y) \approx L(x, y)$ for points near (x_0, y_0). This is illustrated in Figure 11.5.7.

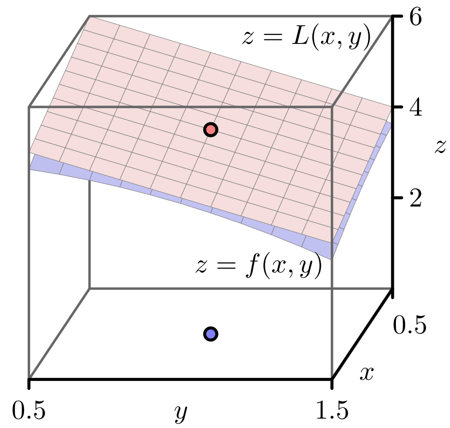


Figure 11.5.7 The linearization of $f(x, y)$.

Activity 11.5.3 In what follows, we 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 $(1, 2)$. Then use the linearization to estimate the value of $g(0.8, 2.3)$.

b. Table 11.5.8 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.

Table 11.5.8 Wind chill as a function of wind speed and temperature.

$v\backslash T$	-30	-25	-20	-15	-10	-5	0	5	10	15	20
5	-46	-40	-34	-28	-22	-16	-11	-5	1	7	13
10	-53	-47	-41	-35	-28	-22	-16	-10	-4	3	9
15	-58	-51	-45	-39	-32	-26	-19	-13	-7	0	6
20	-61	-55	-48	-42	-35	-29	-22	-15	-9	-2	4
25	-64	-58	-51	-44	-37	-31	-24	-17	-11	-4	3
30	-67	-60	-53	-46	-39	-33	-26	-19	-12	-5	1
35	-69	-62	-55	-48	-41	-34	-27	-21	-14	-7	0
40	-71	-64	-57	-50	-43	-36	-29	-22	-15	-8	-1

Use the data to first estimate the appropriate partial derivatives, and then find the linearization $L(v, T)$ at the point $(20, -10)$. Finally, use the linearization to estimate $w(10, -10)$, $w(20, -12)$, and $w(18, -12)$. Compare your results to what you obtained in Activity 11.3.5

c. Figure 11.5.9 gives a contour plot of a continuously differentiable function f.

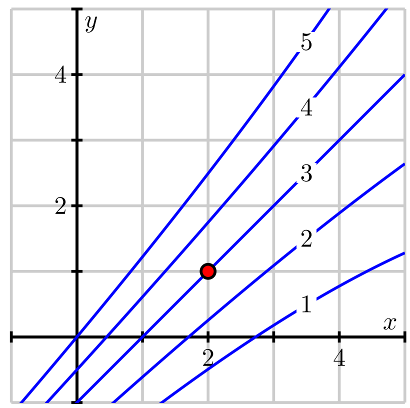


Figure 11.5.9 A contour plot of $f(x, y)$.

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 $f(2.2, 0.8)$.

11.5.3 Differentials

As we have seen, the linearization $L(x, y)$ enables us to estimate the value of $f(x, y)$ for points (x, y) near the base point (x_0, y_0). Sometimes, however, we are more interested in the *change* in f as we move from the base point (x_0, y_0) to another point (x, y).

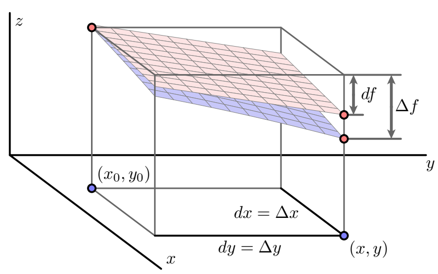


Figure 11.5.10 The differential df approximates the change in $f(x, y)$.

Figure 11.5.10 illustrates this situation. Suppose we are at the point (x_0, y_0), and we know the value $f(x_0, y_0)$ of f at (x_0, y_0). If we consider the displacement $\langle \Delta x, \Delta y\rangle$ to a new point $(x, y) = (x_0 + \Delta x, y_0 + \Delta y)$, we would like to know how much the function has changed. We denote this change by Δf, where

$$\Delta f = f(x, y) - f(x_0, y_0).$$

A simple way to estimate the change Δf is to approximate it by df, which represents the change in the linearization $L(x, y)$ as we move from (x_0, y_0) to (x, y). This gives

$$\begin{aligned}
\Delta f \approx df &= L(x, y) - f(x_0, y_0)\\
&= [f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)] - f(x_0, y_0)\\
&= f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y.
\end{aligned}$$

For consistency, we will denote the change in the independent variables as $dx = \Delta x$ and $dy = \Delta y$, and thus

$$\Delta f \approx df = f_x(x_0, y_0)\ dx + f_y(x_0, y_0)\ dy. \tag{11.5.2}$$

Expressed equivalently in Leibniz notation, we have

$$df = \frac{\partial f}{\partial x}\ dx + \frac{\partial f}{\partial y}\ dy. \tag{11.5.3}$$

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. Equations (11.5.2) and (11.5.3) express the relationship between these changes. Equation (11.5.3) resembles an important idea from single-variable calculus: when y depends on x, it follows in the notation of differentials that

$$dy = y'\ dx = \frac{dy}{dx}\ dx.$$

We will illustrate the use of differentials with an example.

Example 11.5.11 Suppose we have a machine that manufactures rectangles of width $x = 20$ cm and height $y = 10$ cm. However, the machine isn't perfect, and therefore the width could be off by $dx = \Delta x = 0.2$ cm and the height could be off by $dy = \Delta y = 0.4$ cm.

The area of the rectangle is

$$A(x, y) = xy,$$

so 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 manufactured rectangle could differ from the perfect rectangle. We will estimate the uncertainty in the area using (11.5.2), and find that

$$\Delta A \approx dA = A_x(20, 10)\ dx + A_y(20, 10)\ dy.$$

Since $A_x = y$ and $A_y = x$, we have

$$\Delta A \approx dA = 10\ dx + 20\ dy = 10 \cdot 0.2 + 20 \cdot 0.4 = 10.$$

That is, we estimate that the area in our rectangles could be off by as much as 10 square centimeters. □

Activity 11.5.4 The questions in this activity explore the differential in several different contexts.

a. Suppose that the elevation of a landscape 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 easterly and northerly directions, respectively, from some base point $(0, 0)$. 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 units; that is, $dx = \Delta x = 0.2$ and $dy = \Delta y = 0.2$. Estimate the uncertainty in your elevation using differentials.

b. The pressure, volume, and temperature of an ideal gas are related by the equation

$$P = P(T, V) = 8.31T/V,$$

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 differentials to estimate the change in the pressure when the volume increases to 12.3 liters and the temperature decreases to 305 K.

c. Refer to Table 11.5.8, the table of values of 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. Suppose your anemometer says the wind is blowing at 25 miles per hour and your thermometer 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.

11.5.4 Summary

- A function f of two independent variables is locally linear at a point (x_0, y_0) if the graph of f 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 tangent plane $L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$, when considered as a function, 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).

- A function f of two independent variables is 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 f that result from small changes in x and y.

11.5.5 Exercises

1. Find the linearization $L(x, y)$ of the function $f(x, y) = \sqrt{34 - 9x^2 - 1y^2}$ at $(1, -3)$.

$L(x, y) =$ ______________________________

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 $z = e^{2x/17} \ln(4y)$ at the point $(-3, 2, 1.461)$.

z = ______________________________

Note: Your answer should be an expression of x and y; e.g. “5x + 2y - 3”

11.6 The Chain Rule

Motivating Questions

- What is the Chain Rule and how do we use it to find a derivative?
- How can we use a tree diagram to guide us in applying the Chain Rule?

In single-variable calculus, we encountered situations in which some quantity z depends on y and, in turn, y depends on x. A change in x produces a change in y, which consequently produces a change in z. Using the language of differentials that we saw in the previous section, these changes are naturally related by

$$dz = \frac{dz}{dy}\,dy \text{ and } dy = \frac{dy}{dx}\,dx.$$

In terms of instantaneous rates of change, we then have

$$dz = \frac{dz}{dy}\frac{dy}{dx}\,dx = \frac{dz}{dx}\,dx$$

and thus

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

This most recent equation we call the *Chain Rule*.

In the case of a function f of two variables where $z = f(x, y)$, it might be that both x and y depend on another variable t. A change in t then produces changes in both x and y, which then cause z to change. In this section we will see how to find 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 11.6.1 Suppose you are driving around in the xy-plane in such a way that your position $\vec{r}(t)$ at time t is given by function

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

The path taken is shown on the left of Figure 11.6.1.

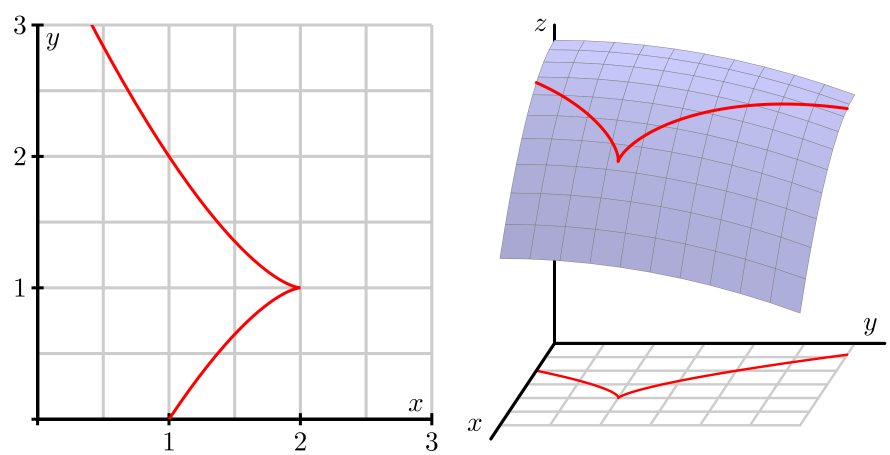


Figure 11.6.1 Left: Your position in the plane. Right: The corresponding temperature.

Suppose, furthermore, that the temperature at a point in the plane is given by

$$T(x, y) = 10 - \frac{1}{2}x^2 - \frac{1}{5}y^2,$$

and note that the surface generated by T is shown on the right of Figure 11.6.1. Therefore, as time passes, your position $(x(t), y(t))$ changes, and, as your position changes, the temperature $T(x, y)$ also changes.

a. The position function $\vec{r}$ provides a parameterization $x = x(t)$ and $y = y(t)$ of the position at time t. By substituting $x(t)$ for x and $y(t)$ for y in the formula for T, we can write $T = T(x(t), y(t))$ as a function of t. Make these substitutions to write T as a function of t and then use the Chain Rule from single variable calculus to find $\frac{dT}{dt}$. (Do not do any algebra to simplify the derivative, either before taking the derivative, nor after.)

b. Now we want to understand how the result from part (a) can be obtained from T as a multivariable function. Recall from the previous section that small changes in x and y produce a change in T that is approximated by

$$\Delta T \approx T_x \Delta x + T_y \Delta y.$$

The Chain Rule tells us about the instantaneous rate of change of T, and this can be found as

$$\lim_{\Delta t \to 0} \frac{\Delta T}{\Delta t} = \lim_{\Delta t \to 0} \frac{T_x \Delta x + T_y \Delta y}{\Delta t}. \tag{11.6.1}$$

Use equation (11.6.1) to explain why the instantaneous rate of change of T that results from a change in t is

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}. \tag{11.6.2}$$

c. Using the original formulas for T, x, and y in the problem statement, calculate all of the derivatives in Equation (11.6.2) (with T_x and T_y in terms of x and y, and x' and y' in terms of t), and hence write the right-hand side of Equation (11.6.2) in terms of x, y, and t.

d. Compare the results of parts (a) and (c). Write a couple of sentences that identify specifically how each term in (c) relates to a corresponding terms in (a). This connection between parts (a) and (c) provides a multivariable version of the Chain Rule.

11.6.1 The Chain Rule

As Preview Activity 11.4.1 suggests, the following version of the Chain Rule holds in general.

The Chain Rule.

Let $z = f(x, y)$, where f is a differentiable function of the independent variables x and y, and let x and y each be differentiable functions of an independent variable t. Then

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{11.6.3}$$

It is important to note the differences among the derivatives in (11.6.3). 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 $y(t)$, we then have z as a function of the single variable t, making the derivative of z with respect to t a standard derivative from single variable calculus as well.

To understand why this Chain Rule works in general, suppose that some quantity z depends on x and y so that

$$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy. \tag{11.6.4}$$

Next, suppose that x and y each depend on another quantity t, so that

$$dx = \frac{dx}{dt}\, dt \text{ and } dy = \frac{dy}{dt}\, dt. \tag{11.6.5}$$

Combining Equations (11.6.4) and (11.6.5), 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 (11.6.3).

Activity 11.6.2 In the following questions, we apply the Chain Rule in several different contexts.

a. Suppose that we have a function z defined by $z(x, y) = x^2 + xy^3$. 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

$$x(t) = 2\cos(t), \text{ and } y(t) = 2\sin(t).$$

   i. Use the Chain Rule to find the resulting 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 $\frac{dz}{dt}$ directly. Compare to the result of part (i.).

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 an ellipse with $x = 6\cos(t)$ and $y = 3\sin(t)$, where t is measured in minutes. Find $\frac{dT}{dt}$ at $t = \pi/6$, $t = \pi/4$, and $t = \pi/3$. What does this seem to 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. 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 $\vec{v} = \langle -1, 2\rangle$. If some contours of f are as shown in Figure 11.6.2, estimate the rate of change df/dt when you pass through $(2, 1)$.

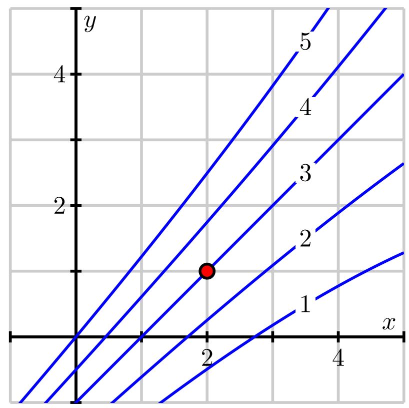


Figure 11.6.2 Some contours of f.

11.6.2 Tree Diagrams

Up to this point, we have applied the Chain Rule to situations where we have a function z of variables x and y, with both x and y depending on another single quantity t. We may apply the Chain Rule, however, 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 11.6.3. 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.

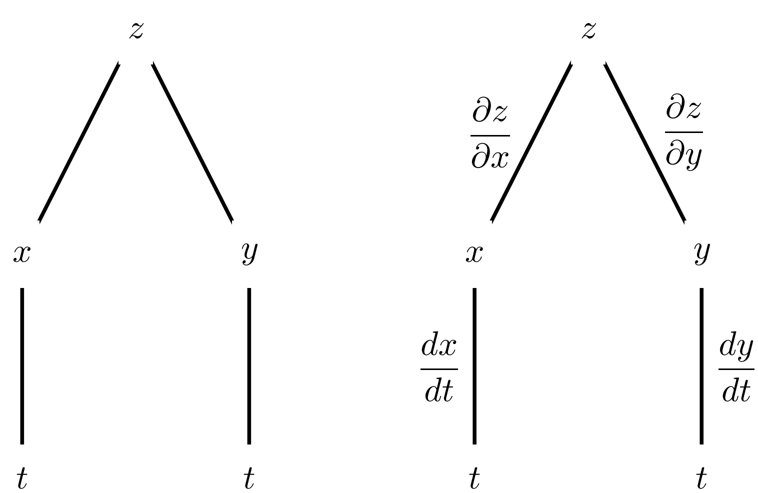


Figure 11.6.3 A tree diagram illustrating dependencies.

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 11.6.3. 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 11.6.3 illustrates the Chain Rule

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Activity 11.6.3

a. Figure 11.6.4 shows the tree diagram we construct when (a) z depends on w, x, and y, (b) w, x, and y each depend on u and v, and (c) u and v depend on t.

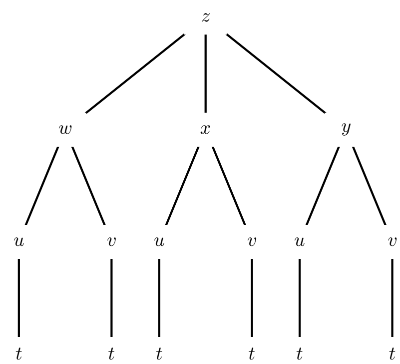


Figure 11.6.4 Three levels of dependencies

i. Label the edges with the appropriate derivatives.

ii. Use the Chain Rule to write $\frac{dz}{dt}$.

b. Suppose that $z = x^2 - 2xy^2$ and that

$$x = r\cos(\theta)$$
$$y = r\sin(\theta).$$

i. Construct a tree diagram representing the dependencies of z on x and y and x and y on r and θ.

ii. Use the tree diagram to find $\frac{\partial z}{\partial r}$.

iii. Now suppose that $r = 3$ and $\theta = \pi/6$. Find the values of x and y that correspond to these given values of r and θ, and then use the Chain Rule to find the value of the partial derivative $\frac{\partial z}{\partial \theta}|_{(3,\frac{\pi}{6})}$.

11.6.3 Summary

- The Chain Rule is a tool for differentiating a composite for 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

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

- A tree diagram can be used to represent the dependence of variables on other variables. By following the links in the tree diagram, we can form chains of partial derivatives or derivatives that can be combined to give a desired partial derivative.

11.6.4 Exercises

1. Use the chain rule to find $\frac{dz}{dt}$, where

$$z = x^2y + xy^2, \quad x = 1 - t^2, \quad y = 1 - t^7$$

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

$$z = e^{xy}\tan y, x = 4s + 4t, y = \frac{7s}{6t}$$

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 $w = \frac{x}{y} + \frac{y}{z}$, where
$x = e^{5t}$, $y = 2 + \sin(2t)$, and $z = 2 + \cos(4t)$.
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.
$\frac{\partial z}{\partial u} =$ ______________________
$\frac{\partial z}{\partial v} =$ ______________________

5. If

$$z = \sin\left(x^2 + y^2\right), \qquad x = v\cos(u), \qquad y = v\sin(u),$$

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 $z = g(u, v)$ and $u(r, s), v(r, s)$. 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

$$u(1, 0) = -4, u_s(1, 0) = -9, u_t(1, 0) = -9$$

$$v(1,0) = -1, v_s(1,0) = 9, v_t(1,0) = 9$$

$$F_u(-4,-1) = -1, F_v(-4,-1) = -6$$

$W_s(1,0) =$ ____________________ $W_t(1,0) =$ ____________________

8. The radius of a right circular cone is increasing at a rate of 3 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 30 inches and the height is 10 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 0.02 volts per second and, as the resistor heats up, the resistance is increasing at 0.02 ohms per second. When the resistance is 400 ohms and the current is 0.04 amperes, at what rate is the current changing?
____________________________________ amperes per second

10. Suppose $z = x^2 \sin y$, $x = 2s^2 + 3t^2$, $y = -2st$.
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 $(s,t) = (1,3)$.
$\frac{\partial z}{\partial s}(1,3) =$ ____
$\frac{\partial z}{\partial t}(1,3) =$ ____

11. 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 $z = u^2 - v$

12. 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 $\frac{\partial z}{\partial x}\Big|_{(0,\frac{2\pi}{3})}$ and $\frac{\partial z}{\partial y}\Big|_{(0,\frac{2\pi}{3})}$.

13. Suppose that $T = x^2 + y^2 - 2z$ where

$$x = \rho\sin(\phi)\cos(\theta)$$
$$y = \rho\sin(\phi)\sin(\theta)$$
$$z = \rho\cos(\phi)$$

a. Construct a tree diagram representing the dependencies among the variables.

b. Apply the chain rule to find the partial derivatives

$$\frac{\partial T}{\partial \rho}, \frac{\partial T}{\partial \phi}, \text{ and } \frac{\partial T}{\partial \theta}.$$

14. 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.

15. 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.

16. 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 $t = 2\pi/3$.

17. 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 11.6.5. The total resistance R in the circuit is then given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$

[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.

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?

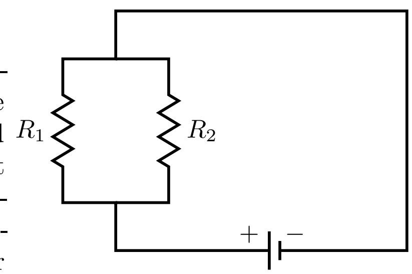


Figure 11.6.5 Resistors in parallel.

11.7 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 function and what does it tell us?

The partial derivatives of a function tell us the instantaneous rate at which the function changes as we hold all but one independent variable constant and allow the remaining independent variable to change. It is natural to wonder how we can measure the rate at which a function changes in directions other than parallel to a coordinate axes. In what follows, we investigate this question, and see how the rate of change in any given direction is connected to the rates of change given by the standard partial derivatives.

Preview Activity 11.7.1 Let's consider the function f defined by

$$f(x, y) = 30 - x^2 - \frac{1}{2}y^2,$$

and suppose that f measures the temperature, in degrees Celsius, at a given point in the plane, where x and y are measured in feet. Assume that the positive x-axis points due east, while the positive y-axis points due north. A contour plot of f is shown in Figure 11.7.1

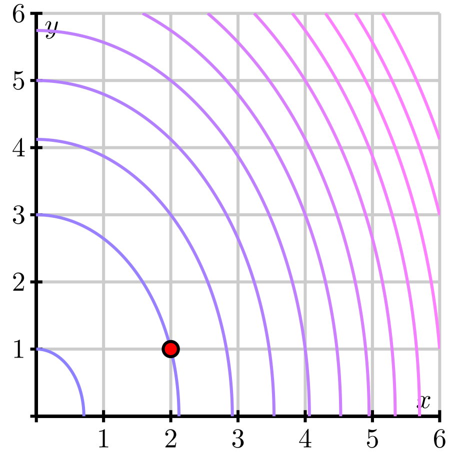


Figure 11.7.1 A contour plot of $f(x, y) = 30 - x^2 - \frac{1}{2}y^2$.

a. Suppose that a person is walking due east, and thus parallel to the x-axis. At what instantaneous rate is the temperature changing with respect to x at the moment the walker passes the point $(2, 1)$? What are the units on this rate of change?

b. Next, determine the instantaneous rate of change of temperature with respect to distance at the point $(2, 1)$ if the person is instead walking due north. Again, include units on your result.

c. Now, rather than walking due east or due north, let's suppose that the person is walking with velocity given by the vector $\vec{v} = \langle 3, 4\rangle$, where time is measured in seconds. Note that the person's speed is thus $|\vec{v}| = 5$ feet per second. Find parametric equations for the person's path; that is, parameterize the line through $(2, 1)$ using the direction vector $\vec{v} = \langle 3, 4\rangle$. Let $x(t)$ denote the x-coordinate of the line, and $y(t)$ its y-coordinate. Make sure your parameterization places the walker at the point $(2, 1)$ when $t = 0$.

d. With the parameterization in (c), we can now view the temperature f as not only a function of x and y, but also of time, t. Hence, use the chain rule to determine the value of $\frac{df}{dt}\Big|_{t=0}$. What are the units on your answer? What is the practical meaning of this result?

11.7.1 Directional Derivatives

Given a function $z = f(x, y)$, the partial derivative $f_x(x_0, y_0)$ measures the instantaneous rate of change of f as only the x variable changes; likewise, $f_y(x_0, y_0)$ measures the rate of change of f at (x_0, y_0) as only y changes. Note particularly that $f_x(x_0, y_0)$ is measured in "units of f per unit of change in x," and that the units on $f_y(x_0, y_0)$ are similar.

In Preview Activity 11.7.1, we saw how we could measure the rate of change of f in a situation where both x and y were changing; in that activity, however, we found that this rate of change was measured in "units of f per unit of *time*." In a given unit of time, we may move more than one unit of distance. In fact, in Preview Activity 11.7.1, in each unit increase in time we move a distance of $|\vec{v}| = 5$ feet. To generalize the notion of partial derivatives to any direction of our choice, we instead want to have a rate of change whose units are "units of f per unit of distance in the given direction."

In this light, in order to formally define the derivative in a particular direction of motion, we want to represent the change in f for a given *unit* change in the direction of motion. We can represent this unit change in direction with a unit vector, say $\vec{u} = \langle u_1, u_2\rangle$. If we move a distance h in the direction of $\vec{u}$ from a fixed point (x_0, y_0), we then arrive at the new point $(x_0 + u_1h, y_0 + u_2h)$. It now follows that the slope of the secant line to the curve on the surface through (x_0, y_0) in the direction of $\vec{u}$ through the points (x_0, y_0) and (x_0+u_1h, y_0+u_2h) is

$$m_{\text{sec}} = \frac{f(x_0 + u_1h, y_0 + u_2h) - f(x_0, y_0)}{h}. \tag{11.7.1}$$

To get the instantaneous rate of change of f in the direction $\vec{u} = \langle u_1, u_2\rangle$, we must take the limit of the quantity in Equation (11.7.1) as $h \to 0$. Doing so results in the formal definition of the directional derivative.

Definition 11.7.2 Let $f = f(x, y)$ be given. 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

$$D_{\vec{u}}f(x, y) = \lim_{h\to 0} \frac{f(x + u_1h, y + u_2h) - f(x, y)}{h} \tag{11.7.2}$$

for those values of x and y for which the limit exists. ♢

The quantity $D_{\vec{u}}f(x, y)$ is called a *directional derivative*. 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 $\vec{u}$. In addition, the quantity $D_{\vec{u}}f(x_0, y_0)$ tells us the slope of the line tangent to the surface in the direction of $\vec{u}$ at the point $(x_0, y_0, f(x_0, y_0))$.

11.7.2 Computing the Directional Derivative

In a similar way to how we developed shortcut rules for standard derivatives in single variable calculus, and for partial derivatives in multivariable calculus, we can also find a way to evaluate directional derivatives without resorting to the limit definition found in Equation (11.7.2). We do so using a very similar approach to our work in Preview Activity 11.7.1.

Suppose we consider the situation where we are interested in the instantaneous rate of change of f at a point (x_0, y_0) in the direction $\vec{u} = \langle u_1, u_2\rangle$, where $\vec{u}$ is a unit vector. The variables x and y are therefore changing according to the parameterization

$$x = x_0 + u_1t \qquad \text{and} \qquad y = y_0 + u_2t.$$

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, it follows that a point moving along this line moves one unit of distance per one unit of time; that is, each single unit of time corresponds to movement of a single unit of distance in that direction. This observation allows us to use the Chain Rule to calculate the directional derivative, which measures the instantaneous rate of change of f with respect to change in the direction $\vec{u}$.

In particular, by the Chain Rule, it follows that

$$\begin{aligned} D_{\vec{u}}f(x_0, y_0) &= f_x(x_0, y_0)\frac{dx}{dt}\Big|_{(x_0,y_0)} + f_y(x_0, y_0)\frac{dy}{dt}\Big|_{(x_0,y_0)} \\ &= f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2. \end{aligned}$$

This now allows us to compute the directional derivative at an arbitrary point according to the following formula.

Calculating a directional derivative.

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

$$D_{\vec{u}} f(x, y) = f_x(x, y)u_1 + f_y(x, y)u_2. \tag{11.7.3}$$

Note well: To use Equation (11.7.3), 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 scale the vector to have length 1.

Activity 11.7.2 Let $f(x, y) = 3xy - x^2y^3$.

a. Determine $f_x(x, y)$ and $f_y(x, y)$.

b. Use Equation (11.7.3) to determine $D_{\hat{\imath}} f(x, y)$ and $D_{\hat{\jmath}} f(x, y)$. What familiar function is $D_{\hat{\imath}} f$? What familiar function is $D_{\hat{\jmath}} f$? (Recall 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 (11.7.3) to find the derivative of f in the direction of the vector $\vec{v} = \langle 2, 3 \rangle$ at the point $(1, -1)$. Remember that a unit direction vector is needed.

11.7.3 The Gradient

Via the Chain Rule, we have seen that for a given function $f = f(x, y)$, its instantaneous rate of change in the direction of a unit vector $\vec{u} = \langle u_1, u_2 \rangle$ is given by

$$D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2. \tag{11.7.4}$$

Recalling that the dot product of two vectors $\vec{v} = \langle v_1, v_2 \rangle$ and $\vec{u} = \langle u_1, u_2 \rangle$ is computed by

$$\vec{v} \cdot \vec{u} = v_1u_1 + v_2u_2,$$

we see that we may recast Equation (11.7.4) in a way that has geometric meaning. In particular, we see that $D_{\vec{u}} f(x_0, y_0)$ is the dot product of the vector $\langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle$ and the vector $\vec{u}$.

We call this vector formed by the partial derivatives of f the *gradient* of f and denote it

$$\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle .$$

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.

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

$$D_{\vec{u}} f(x, y) = \nabla f(x, y) \cdot \vec{u}. \tag{11.7.5}$$

[1] The symbol ∇ is called *nabla*, which comes from a Greek word for a certain type of harp that has a similar shape.

In the following activity, we investigate some of what the gradient tells us about the behavior of a function f.

Activity 11.7.3 Let's consider the function f defined by $f(x,y) = x^2 - y^2$. Some contours for this function are shown in Figure 11.7.3.

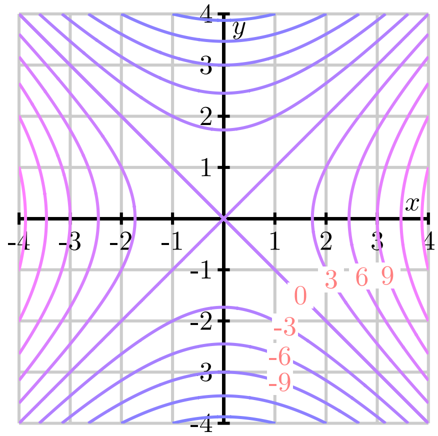


Figure 11.7.3 Contours of $f(x,y) = x^2 - y^2$.

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, but we'd like to draw them to scale; to do so, scale each vector by a factor of $1/4$.

- $(x_0, y_0) = (2, 0)$
- $(x_0, y_0) = (0, 2)$
- $(x_0, y_0) = (2, 2)$
- $(x_0, y_0) = (2, 1)$
- $(x_0, y_0) = (-3, 2)$
- $(x_0, y_0) = (-2, -4)$
- $(x_0, y_0) = (0, 0)$

c. What do you notice about the relationship between the gradient at (x_0, y_0) and the contour passing through that point?

d. Does f increase or decrease in the direction of $\nabla f(x_0, y_0)$? Provide a justification for your response.

As a vector, $\nabla f(x_0, y_0)$ defines a direction and a length. As we will soon see, both of these convey important information about the behavior of f near (x_0, y_0).

11.7.4 The Direction of the Gradient

Remember that the dot product also conveys information about the angle between the two vectors. If θ is the angle between $\nabla f(x_0, y_0)$ and $\vec{u}$ (where $\vec{u}$ is a unit vector), then we also have that

$$D_{\vec{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \vec{u} = |\nabla f(x_0, y_0)||\vec{u}| \cos(\theta).$$

In particular, when θ is a right angle, as shown on the left of Figure 11.7.4, then $D_{\vec{u}}f(x_0, y_0) = 0$, because $\cos(\theta) = 0$. Since the value of the directional derivative is 0, this means that f is unchanging in this direction, and hence $\vec{u}$ must be tangent to the contour of f that passes through (x_0, y_0). In other words, $\nabla f(x_0, y_0)$ is orthogonal to the contour through (x_0, y_0). This shows that the gradient vector at a given point is always perpendicular to the contour passing through the point, confirming that what we saw in part (c) of Activity 11.7.3 holds in general.

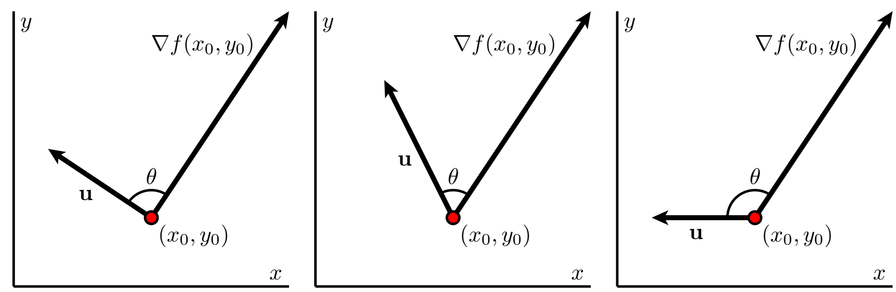


Figure 11.7.4 The sign of $D_{\vec{u}}f(x_0, y_0)$ is determined by θ.

Moreover, when θ is an acute angle, it follows that $\cos(\theta) > 0$ so since

$$D_{\vec{u}}f(x_0, y_0) = |\nabla f(x_0, y_0)||\vec{u}|\cos(\theta),$$

and therefore $D_{\vec{u}}f(x_0, y_0) > 0$, as shown in the middle image in Figure 11.7.4. This means that f is increasing in any direction where θ is acute. In a similar way, when θ is an obtuse angle, then $\cos(\theta) < 0$ so $D_{\vec{u}}f(x_0, y_0) < 0$, as seen on the right in Figure 11.7.4. This means that f is decreasing in any direction for which θ is obtuse.

Finally, as we can see in the following activity, we may also use the gradient to determine the directions in which the function is increasing and decreasing most rapidly.

Activity 11.7.4 In this activity we investigate how the gradient is related to the directions of greatest increase and decrease of a function. Let f be a differentiable function and $\vec{u}$ a unit vector.

a. Let θ be the angle between $\nabla f(x_0, y_0)$ and $\vec{u}$. Use the relationship between the dot product and the angle between two vectors to explain why

$$D_{\vec{u}}f(x_0, y_0) = |\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle| \cos(\theta). \tag{11.7.6}$$

b. At the point (x_0, y_0), the only quantity in Equation (11.7.6) that can change is θ (which determines the direction $\vec{u}$ of travel). Explain why $\theta = 0$ makes the quantity

$$|\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle| \cos(\theta)$$

as large as possible.

c. When $\theta = 0$, in what direction does the unit vector $\vec{u}$ point relative to $\nabla f(x_0, y_0)$? Why? What does this tell us about the direction of greatest increase of f at the point (x_0, y_0)?

d. In what direction, relative to $\nabla f(x_0, y_0)$, does f decrease most rapidly at the point (x_0, y_0)?

e. State the unit vectors $\vec{u}$ and $\vec{v}$ (in terms of $\nabla f(x_0, y_0)$) that provide the directions of greatest increase and decrease for the function f at the point (x_0, y_0). What important assumption must we make regarding $\nabla f(x_0, y_0)$ in order for these vectors to exist?

11.7.5 The Length of the Gradient

Having established in Activity 11.7.4 that the direction in which a function increases most rapidly at a point (x_0, y_0) is the unit vector $\vec{u}$ in the direction of the gradient, (that is, $\vec{u} = \frac{1}{|\nabla f(x_0,y_0)|}\nabla f(x_0, y_0)$, provided that $\nabla f(x_0, y_0) \neq \vec{0}$), it is also natural to ask, "in the direction of greatest increase for f at (x_0, y_0), what is the *value* of the rate of increase?" In this situation, we are asking for the value of $D_{\vec{u}}f(x_0, y_0)$ where $\vec{u} = \frac{1}{|\nabla f(x_0,y_0)|}\nabla f(x_0, y_0)$.

Using the now familiar way to compute the directional derivative, we see that

$$D_{\vec{u}}f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \left(\frac{1}{|\nabla f(x_0, y_0)|}\nabla f(x_0, y_0)\right).$$

Next, we recall two important facts about the dot product: (i) $\vec{w} \cdot (c\vec{v}) = c(\vec{w} \cdot \vec{v})$ for any scalar c, and (ii) $\vec{w} \cdot \vec{w} = |\vec{w}|^2$. Applying these properties to the most recent equation involving the directional derivative, we find that

$$D_{\vec{u}}f(x_0, y_0) = \frac{1}{|\nabla f(x_0, y_0)|}(\nabla f(x_0, y_0)\cdot\nabla f(x_0, y_0)) = \frac{1}{|\nabla f(x_0, y_0)|}|\nabla f(x_0, y_0)|^2.$$

Finally, since $\nabla f(x_0, y_0)$ is a nonzero vector, its length $|\nabla f(x_0, y_0)|$ is a nonzero scalar, and thus we can simplify the preceding equation to establish that

$$D_{\vec{u}}f(x_0, y_0) = |\nabla f(x_0, y_0)|.$$

We summarize our most recent work by stating two important facts about the gradient.

Important facts about the gradient.

Let f be a differentiable function and (x_0, y_0) a point for which $\nabla f(x_0, y_0) \neq \vec{0}$. Then $\nabla f(x_0, y_0)$ points in the direction of greatest increase of f at (x_0, y_0), and the instantaneous rate of change of f in that direction is the length of the gradient vector. That is, 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 $D_{\vec{u}}f(x_0, y_0) = |\nabla f(x_0, y_0)|$.

Activity 11.7.5 Consider the function f defined by $f(x, y) = -x + 2xy - y$.

a. Find the gradient $\nabla f(1, 2)$ and sketch it on Figure 11.7.5.

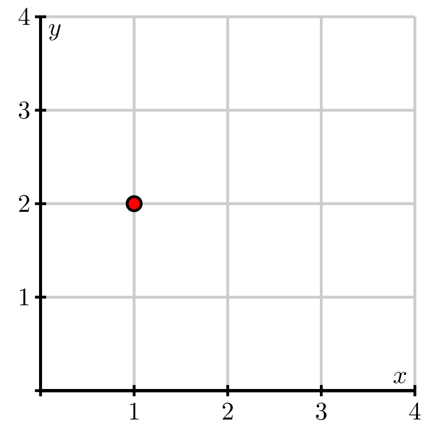


Figure 11.7.5 A plot for the gradient $\nabla f(1, 2)$.

b. Sketch the unit vector $\vec{z} = \left\langle -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\rangle$ on Figure 11.7.5 with its tail at $(1, 2)$. Now find the directional derivative $D_{\vec{z}} f(1, 2)$.

c. What is the slope of the graph of f in the direction $\vec{z}$? What does the sign of the directional derivative tell you?

d. Consider the vector $\vec{v} = \langle 2, -1 \rangle$ and sketch $\vec{v}$ on Figure 11.7.5 with its tail at $(1, 2)$. Find a unit vector $\vec{w}$ pointing in the same direction of $\vec{v}$. Without computing $D_{\vec{w}} f(1, 2)$, what do you know about the sign of this directional derivative? Now verify your observation by computing $D_{\vec{w}} f(1, 2)$.

e. In which direction (that is, for what unit vector $\vec{u}$) is $D_{\vec{u}} f(1, 2)$ the greatest? What is the slope of the graph in this direction?

f. Corresponding, in which direction is $D_{\vec{u}} f(1, 2)$ least? What is the slope of the graph in this direction?

g. Sketch two unit vectors $\vec{u}$ for which $D_{\vec{u}} f(1, 2) = 0$ and then find component representations of these vectors.

h. Suppose you are standing at the point $(3, 3)$. In which direction should you move to cause f to increase as rapidly as possible? At what rate does f increase in this direction?

11.7.6 Applications

The gradient finds many natural applications. For example, situations often arise — for instance, constructing a road through the mountains or planning the flow of water across a landscape — where we are interested in knowing the direction in which a function is increasing or decreasing most rapidly.

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 11.7.6.

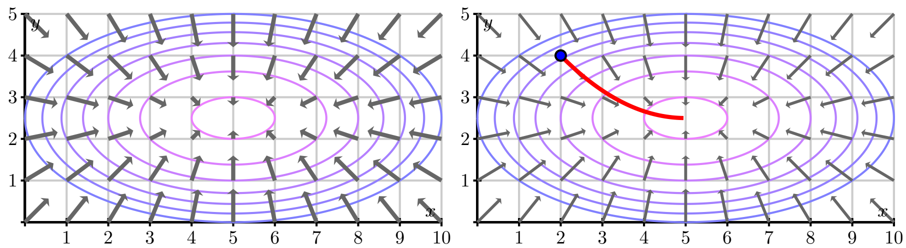


Figure 11.7.6 Contours and gradient for $T(x, y)$ and the missile's path.

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 11.7.6.

In the final activity of this section, we consider several questions related to this context of a heat-seeking missile, and foreshadow some upcoming work in Section 11.8.

Activity 11.7.6

a. The temperature $T(x, y)$ has its maximum value at the fighter jet's location. State the fighter jet's location and explain how Figure 11.7.6 tells you this.

b. Determine ∇T at the fighter jet's location and give a justification for your response.

c. Suppose that a different function f has a local maximum value at (x_0, y_0). Sketch the behavior of some possible contours near this point. What is $\nabla f(x_0, y_0)$? (Hint: What is the direction of greatest increase in f at the local maximum?)

d. Suppose that a function g has a local minimum value at (x_0, y_0). Sketch the behavior of some possible contours near this point. What is $\nabla g(x_0, y_0)$?

e. If a function g has a local minimum at (x_0, y_0), what is the direction of greatest increase of g at (x_0, y_0)?

11.7.7 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, 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 opposite of the gradient points in the direction of greatest decrease, and that rate of decrease is the opposite of the length of the gradient.

11.7.8 Exercises

1. Consider the function $f(x, y, z) = xy + yz^2 + xz^3$.

Find the gradient of f:

$\langle \underline{\hspace{3cm}}, \underline{\hspace{3cm}}, \underline{\hspace{3cm}} \rangle$

Find the gradient of f at the point (2, 1, 2).

$\langle \underline{\hspace{3cm}}, \underline{\hspace{3cm}}, \underline{\hspace{3cm}} \rangle$

Find the rate of change of the function f at the point (2, 1,2) in the direction $\mathbf{u} = \langle -2/\sqrt{54}, 5/\sqrt{54}, 5/\sqrt{54} \rangle$.

2. If $f(x, y) = 3x^2 - 2y^2$, find the value of the directional derivative at the point $(-3, -4)$ in the direction given by the angle $\theta = \frac{2\pi}{6}$.

3. Find the directional derivative of $f(x, y, z) = 4xy + z^2$ at the point $(5, 1, 3)$ in the direction of the maximum rate of change of f.

f<B><I>u</B></I> $(5, 1, 3) =$ D<B><I>u</B></I> $f(5, 1, 3) =$ $\underline{\hspace{5cm}}$

4. The temperature at any point in the plane is given by $T(x, y) = \dfrac{100}{x^2 + y^2 + 3}$.

(a) What shape are the level curves of T?

- ⊙ lines
- ⊙ circles
- ⊙ hyperbolas
- ⊙ ellipses
- ⊙ parabolas
- ⊙ none of the above

(b) At what point on the plane is it hottest? $\underline{\hspace{6cm}}$

What is the maximum temperature? $\underline{\hspace{6cm}}$

(c) Find the direction of the greatest increase in temperature at the point $(3, -3)$. $\underline{\hspace{6cm}}$

What is the value of this maximum rate of change, that is, the maximum value of the directional derivative at $(3, -3)$? $\underline{\hspace{6cm}}$

(d) Find the direction of the greatest decrease in temperature at the point $(3, -3)$. $\underline{\hspace{6cm}}$

What is the value of this most negative rate of change, that is, the minimum value of the directional derivative at $(3, -3)$? $\underline{\hspace{6cm}}$

5. 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, 5).

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)?

6. If $f(x,y,z) = 2zy^2$, then the gradient at the point $(6,5,6)$ is

$\nabla f(6,5,6) =$ ____________________________

7. The concentration of salt in a fluid at (x,y,z) is given by $F(x,y,z) = 2x^2 + 3y^4 + x^2z^2$ 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 4 cm/sec. How fast is the concentration changing?

rate of change = ____

8. 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

9. Suppose that you are climbing a hill whose shape is given by $z = 409 - 0.07x^2 - 0.09y^2$, and that you are at the point (20, 30, 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)?

10. Are the following statements true or false?

(a) 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.

(b) 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.

(c) If $\vec{u}$ is a unit vector, then $f_{\vec{u}}(a,b)$ is a vector.

(d) $f_{\vec{u}}\,(a,b)$ is parallel to $\vec{u}$.

(e) $\nabla f(a,b)$ is a vector in 3-dimensional space.

(f) The gradient vector $\nabla f(a,b)$ is tangent to the contour of f at (a,b).

(g) $f_{\vec{u}}\,(a,b) = ||\nabla f(a,b)||$.

(h) If $\vec{u}$ is perpendicular to $\nabla f(a,b)$, then $f_{\vec{u}}\,(a,b) = \langle 0,0\rangle$.

11. 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.

12. Find all directions in which the directional derivative of $f(x, y) = ye^{-xy}$ is 1 at the point $(0, 2)$.

13. 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.

14. 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 11.7.7. Illustrate this direction on the surface.

c. A contour diagram of f is shown at right in Figure 11.7.7. Illustrate your calculation from (b) on this contour diagram.

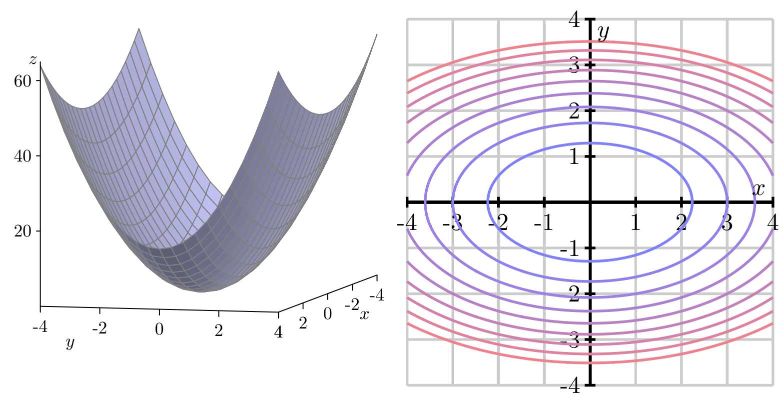


Figure 11.7.7 Left: Graph of $f(x, y) = x^2 + 3y^2$. Right: Contours.

d. Find a direction $\vec{w}$ for which the derivative of f in the direction of $\vec{w}$ is zero.

15. The properties of the gradient that we have observed for functions of two variables also hold for functions of more variables. In this problem, we consider a situation where there are three independent variables. 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 $(1, 2, -1)$.

16. Figure 11.7.8 shows a plot of the gradient ∇f at several points for some function $f = f(x, y)$.

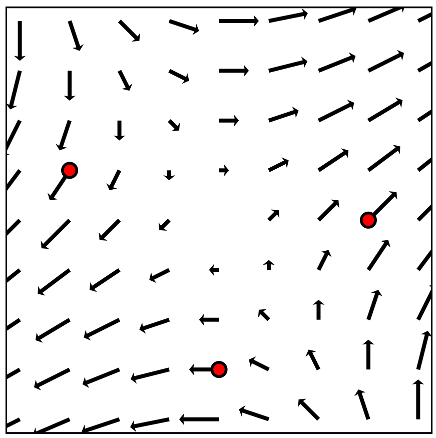


Figure 11.7.8 The gradient ∇f.

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.

11.8 Optimization

Motivating Questions

- How can we find the points at which $f(x, y)$ has a local maximum or minimum?
- 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?

We learn in single-variable calculus that the derivative is a useful tool for finding the local maxima and minima of functions, and that these ideas may often be employed in applied settings. In particular, if a function f, such as the one shown in Figure 11.8.1 is everywhere differentiable, we know that the tangent line is horizontal at any point where f has a local maximum or minimum. This, of course, means that the derivative f' is zero at any such point. Hence, one way that we seek extreme values of a given function is to first find where the derivative of the function is zero.

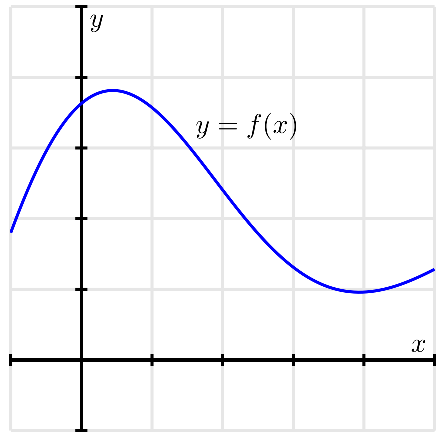


Figure 11.8.1 The graph of $y = f(x)$.

In multivariable calculus, we are often similarly interested in finding the greatest and/or least value(s) that a function may achieve. Moreover, there are many applied settings in which a quantity of interest depends on several different variables. In the following preview activity, we begin to see how some key ideas in multivariable calculus can help us answer such questions by thinking about the geometry of the surface generated by a function of two variables.

Preview Activity 11.8.1 Let $z = f(x, y)$ be a differentiable function, and suppose that at the point (x_0, y_0), f achieves a local maximum. That is, the value of $f(x_0, y_0)$ is greater than the value of $f(x, y)$ for all (x, y) nearby (x_0, y_0). You might find it helpful to sketch a rough picture of a possible function f that has this property.

a. If we consider the trace given by holding $y = y_0$ constant, then the single-variable function defined by $f(x, y_0)$ must have a local maximum at x_0. What does this say about the value of the partial derivative $f_x(x_0, y_0)$?

b. In the same way, the trace given by holding $x = x_0$ constant has a local maximum at $y = y_0$. What does this say about the value of the partial derivative $f_y(x_0, y_0)$?

c. What may we now conclude about the gradient $\nabla f(x_0, y_0)$ at the local maximum? How is this consistent with the statement "f increases most rapidly in the direction $\nabla f(x_0, y_0)$?"

d. How will the tangent plane to the surface $z = f(x, y)$ appear at the point $(x_0, y_0, f(x_0, y_0))$?

e. By first computing the partial derivatives, find any points at which $f(x, y) = 2x - x^2 - (y + 2)^2$ may have a local maximum.

11.8.1 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 naturally extend the concept of local maxima and minima to several-variable functions.

Definition 11.8.2 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 11.8.3 illustrates the graphs of two functions that have an absolute maximum and minimum, respectively, at the origin $(x_0, y_0) = (0, 0)$.

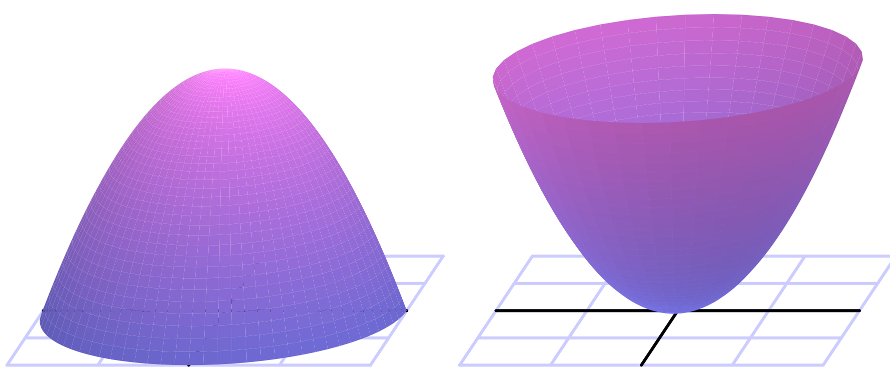


Figure 11.8.3 An absolute maximum and an absolute minimum

In single-variable calculus, we saw that the extrema of a continuous function f always occur at *critical points*, values of x where f fails to be differentiable or where $f'(x) = 0$. Said differently, critical points provide the locations where extrema of a function may appear. Our work in Preview Activity 11.8.1 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 value 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 11.8.4 A **critical point** (x_0, y_0) of a function $f = f(x, y)$ is a point in the domain of f at which $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$, or such that one of $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ fails to exist. ◊

We can therefore 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

$$\begin{aligned} f_x(x, y) &= 0 \\ f_y(x, y) &= 0. \end{aligned}$$

Activity 11.8.2 Find the critical points of each of the following functions. Then, using appropriate technology, plot the graphs of the surfaces near each critical point and compare the graph to your work.

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. $f(x, y) = 2xy - 4x + 2y - 3$.

11.8.2 Classifying Critical Points: The Second Derivative Test

While the extrema of a continuous function f always occur at critical points, 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. We know that $x_0 = 0$ is a critical point since $f'(x_0) = 0$, but $x_0 = 0$ is neither a local maximum nor a local minimum of f.

A similar situation may arise in a multivariable setting. Consider the function f defined by $f(x, y) = x^2 - y^2$ whose graph and contour plot are shown in Figure 11.8.5. 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 or minimum; 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.

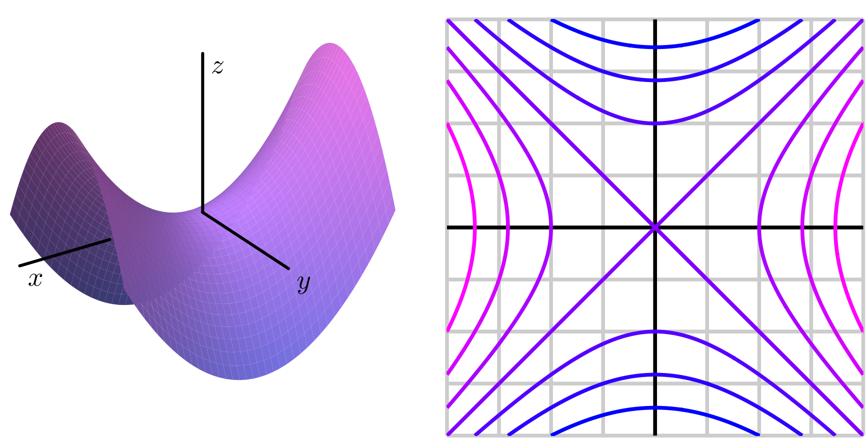


Figure 11.8.5 A saddle point.

As in single-variable calculus, we would like to have some sort of test to help us identify whether a critical point is a local maximum, local maximum, or neither.

Activity 11.8.3 Recall that the *Second Derivative Test* for single-variable functions states that if x_0 is a critical point of a function f so that $f'(x_0) = 0$ and if $f''(x_0)$ exists, then

- if $f''(x_0) < 0$, x_0 is a local maximum,
- if $f''(x_0) > 0$, x_0 is a local minimum, and
- if $f''(x_0) = 0$, this test yields no information.

Our goal in this activity is to understand a similar test for classifying extreme values of functions of two variables. Consider the following three functions:

$$f_1(x, y) = 4 - x^2 - y^2, \quad f_2(x, y) = x^2 + y^2, \quad f_3(x, y) = x^2 - y^2.$$

You can verify that each function has a critical point at the origin $(0, 0)$. You should check this.

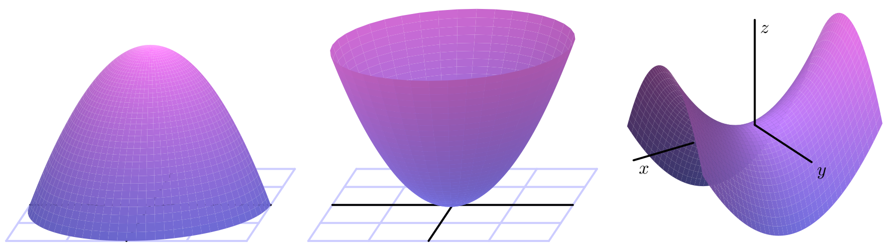


Figure 11.8.6 Three surfaces.

a. The graphs of these three functions are shown in Figure 11.8.6, with $z = 4 - x^2 - y^2$ at left, $z = x^2 + y^2$ in the middle, and $z = x^2 - y^2$ at right. Use the graphs to decide if a function has a local maximum, local minimum, saddle point, or none of the above at the origin.

b. There is no single second derivative of a function of two variables, so we consider a quantity that combines the second order partial derivatives. Let $D = f_{xx}f_{yy} - f_{xy}^2$. Calculate D at the origin for each of the functions f_1, f_2, and f_3. What difference do you notice between the values of D when a function has a maximum or minimum value at the origin versus when a function has a saddle point at the origin?

c. Now consider the cases where $D > 0$. It is in these cases that a function has a local maximum or minimum at a point. What is necessary in these cases is to find a condition that will distinguish between a maximum and a minimum. In the cases where $D > 0$ at the origin, evaluate $f_{xx}(0, 0)$. What value does $f_{xx}(0, 0)$ have when f has a local maximum value at the origin? When f has a local minimum value at the origin? Explain why. (Hint: This should look very similar to the Second Derivative Test for functions of a single variable.) What would happen if we considered the values of $f_{yy}(0, 0)$ instead?

Activity 11.8.3 provides the basic ideas for the Second Derivative Test for functions of two variables.

The Second Derivative Test.

Suppose (x_0, y_0) is a critical point of the function f for which $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$. Let D be the quantity defined by

$$D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.$$

- 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 yields no information about what happens at (x_0, y_0).

The quantity D is called the *discriminant* of the function f at (x_0, y_0).

To properly understand the origin of the Second Derivative Test, we could 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 Derivative Test requires key ideas from linear algebra that are beyond the scope of this course, so instead of presenting a detailed explanation, we will accept this test as stated. In Activity 11.8.4, we apply the test to more complicated examples.

Activity 11.8.4 Find the critical points of the following functions and use the Second Derivative Test to classify the critical points.

a. $f(x, y) = 3x^3 + y^2 - 9x + 4y$

b. $f(x, y) = xy + \dfrac{2}{x} + \dfrac{4}{y}$

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 11.8.5 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 Old Navy 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 as

$$q_1 = 150 - 2p_1 - p_2 \tag{11.8.1}$$

$$q_2 = 200 - p_1 - 3p_2. \tag{11.8.2}$$

The seller would like 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 (11.8.1) and (11.8.2), 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 Figure 11.8.7.

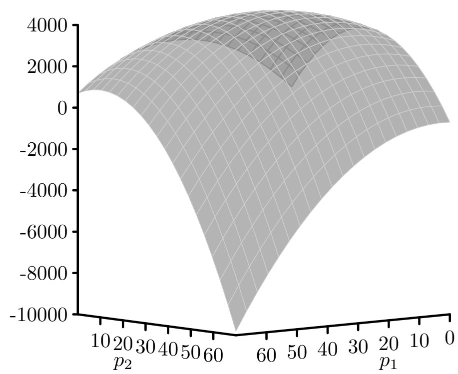


Figure 11.8.7 A revenue function.

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 Derivative 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?

11.8.3 Optimization on a Restricted Domain

The Second Derivative Test helps us classify critical points of a function, but it does not tell us if the function actually has an absolute maximum or minimum at each such point. For single-variable functions, the Extreme Value Theorem told us that a continuous function on a closed interval $[a, b]$ always has both an absolute maximum and minimum on that interval, and that these absolute extremes must occur at either an endpoint or at a critical point. Thus, to find the absolute maximum and minimum, we determine the critical points in the interval and then evaluate the function at these critical points and at the endpoints of the interval. A similar approach works for functions of two variables.

For functions of two variables, closed and bounded regions 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, while its interior $x^2 + y^2 < 1$ is not, for example), while 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 Extreme Value Theorem.

Let $f = f(x, y)$ be a continuous function on a closed and bounded region R. Then f has an absolute maximum and an absolute minimum in R.

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 11.8.8 Suppose the temperature T at each point on the circular plate $x^2 + y^2 \leq 1$ is given by

$$T(x, y) = 2x^2 + y^2 - y.$$

The domain $R = \{(x, y) : x^2 + y^2 \leq 1\}$ is a closed and bounded region, as shown on the left of Figure 11.8.9, so the Extreme Value Theorem assures us that T has an absolute maximum and minimum on the plate. The graph of T over its domain R is shown in Figure 11.8.9. We will find the hottest and coldest points on the plate.

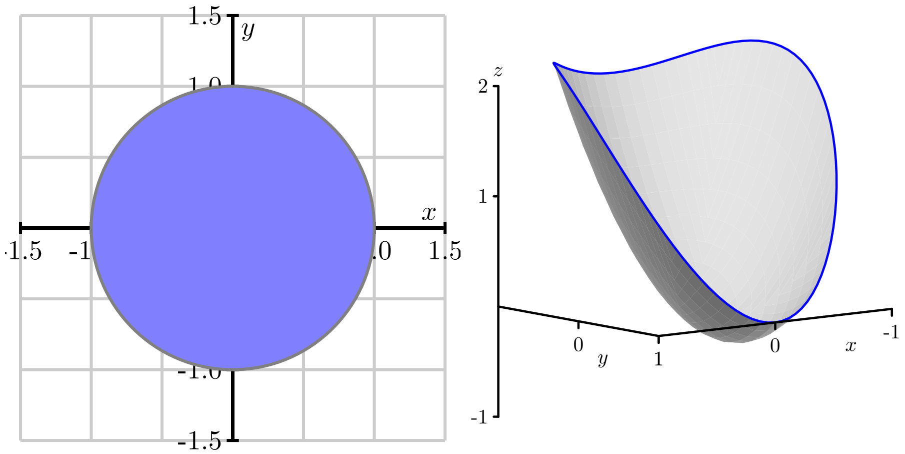


Figure 11.8.9 Domain of the temperature $T(x, y) = 2x^2 + y^2 - y$ and its graph.

If the absolute maximum or minimum occurs inside the disk, it will be at a critical point so 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 lies inside the disk.

We now find the hottest and coldest points on the boundary of the disk, which is the circle of radius 1. As we have seen, the points on the unit circle can be parametrized 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

$$\frac{dT}{dt} = -4\cos(t)\sin(t) + 2\cos(t)\sin(t) - \cos(t)$$

$$
\begin{aligned}
&= -2\cos(t)\sin(t) - \cos(t) = \cos(t)(-2\sin(t) - 1) \\
&= 0.
\end{aligned}
$$

This shows that we have critical points when $\cos(t) = 0$ or $\sin(t) = -1/2$. This occurs when $t = \pi/2$, $3\pi/2$, $7\pi/6$, and $11\pi/6$. Since we have $x(t) = \cos(t)$ 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 and so 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

- $T\left(0, \frac{1}{2}\right) = -\frac{1}{4}$,
- $T\,(0, 1) = 0$,
- $T\,(0, -1) = 2$,
- $T\left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$,
- $T\left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$.

So the maximum value of T 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. $\square$

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. Working on the boundary of the domain reduces this part of the problem to one or more single variable optimization problems. Note that 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 11.8.6 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 and a graph of f on the domain appear in Figure 11.8.10.

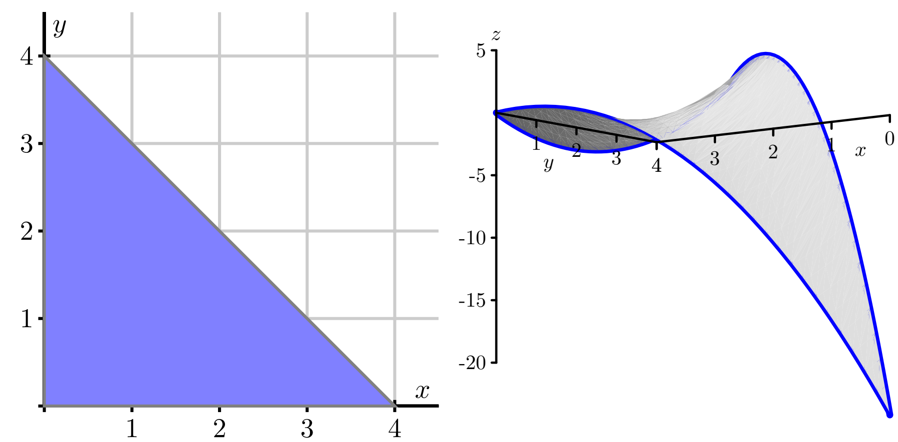


Figure 11.8.10 The domain of $f(x, y) = x^2 - 3y^2 - 4x + 6y$ and its graph.

a. Find all of the critical points of f in R.

b. Parameterize the horizontal leg of the triangular domain, and find the critical points of f on that leg. (Hint: You may need to consider endpoints.)

c. Parameterize the vertical leg of the triangular domain, and find the critical points of f on that leg. (Hint: You may need to consider endpoints.)

d. Parameterize the hypotenuse of the triangular domain, and find the critical points of f on the hypotenuse. (Hint: You may need to consider endpoints.)

e. Find the absolute maximum and absolute minimum values of f on R.

11.8.4 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 $f_x = 0$ and $f_y = 0$.
- The Second Derivative 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 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.

11.8.5 Exercises

1. The function

$$k(x, y) = e^{-y^2} \cos(5x)$$

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)

2. Consider the function $f(x, y) = (18x - x^2)(16y - y^2)$.

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)

3. 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 $\mathbf{R}^2$?

4. 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 = ____

5. 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 $f(x, y) = e^{-2x^2-3y^2}$, type of critical point: (☐ Local Maximum ☐ Local Minimum ☐ Saddle Point ☐ No Critical Point)
For $f(x, y) = e^{2x^2-3y^2}$, type of critical point: (☐ Local Maximum ☐ Local Minimum ☐ Saddle Point ☐ No Critical Point)
For $f(x, y) = 2x^2 + 3y^2 + 3$, type of critical point: (☐ Local Maximum ☐ Local Minimum ☐ Saddle Point ☐ No Critical Point)
For $f(x, y) = 2x^2 + 3y + 3$, type of critical point: (☐ Local Maximum ☐ Local Minimum ☐ Saddle Point ☐ No Critical Point)

6. Find the absolute minimum and absolute maximum of

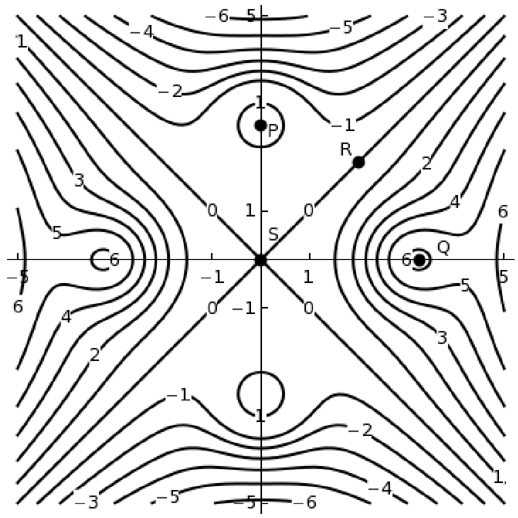
$$f(x, y) = 9 - 5x + 10y$$

on the closed triangular region with vertices $(0, 0), (10, 0)$ and $(10, 11)$.
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 ____________________

7. Find the maximum and minimum values of $f(x, y) = xy$ on the ellipse $5x^2 + y^2 = 2$.
maximum value = ____________________
minimum value = ____________________

8. Find A and B so that $f(x, y) = x^2 + Ax + y^2 + B$ has a local minimum at the point $(6, 0)$, with z-coordinate 45.
A = ___________________________________
B = ___________________________________

9. 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)

10. Consider the three points $(5, 1)$, $(6, -1)$, and $(8, -2)$.

(a) Supposed that at $(5, 1)$, 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 $(5, 1)$? (☐ (5,1) is a local maximum ☐ (5,1) is a local minimum ☐ (5,1) is a saddle point ☐ (5,1) is a none of these)

(b) Supposed that at $(6, -1)$, 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 $(6, -1)$? (☐ (6,-1) is a local maximum ☐ (6,-1) is a local minimum ☐ (6,-1) is a saddle point ☐ (6,-1) is a none of these)

(c) Supposed that at $(8, -2)$, 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 $(8, -2)$? (☐ (8,-2) is a local maximum ☐ (8,-2) is a local minimum ☐ (8,-2) is a saddle point ☐ (8,-2) is a none of these)

Using this information, on a separate sheet of paper sketch a possible contour diagram for f.

11. Find three positive real numbers whose sum is 94 and whose product is a maximum.

Enter the three numbers separated by commas: ____________________

12. A closed rectangular box has volume $20\,\text{cm}^3$. What are the lengths of the edges giving the minimum surface area?

lengths = ____________________________________

(Give the three lengths as a comma separated list.)

13. An open rectangular box has volume 26 cm^3. What are the lengths of the edges giving the minimum surface area?

lengths = ____

(Give the three lengths as a comma separated list.)

14. What is the shortest distance from the surface $xy + 6x + z^2 = 41$ to the origin?
distance = ______________________

15. Find the volume of the largest rectangular box with edges parallel to the axes that can be inscribed in the ellipsoid

$$\frac{x^2}{16} + \frac{y^2}{9} + \frac{z^2}{49} = 1$$

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: ______________________

16. Design a rectangular milk carton box of width w, length l, and height h which holds 498 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.) (Help with units.[1])

17. Respond to each of the following prompts to solve the given optimization problem.

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.

18. 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

[1] `/webwork2_files/helpFiles/Units.html`

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.

19. 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.

20. 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.

21. 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.

22. 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.)

23.

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 11.8.11., 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 11.8.11 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), $\ldots$, (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.

11.9 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?

We previously considered how to find the extreme values of functions on both unrestricted domains and on 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 often called *constrained optimization* problems and can be solved with the method of Lagrange Multipliers, which we study in this section.

Preview Activity 11.9.1 According to U.S. postal regulations, the girth plus the length of a parcel sent by mail may not exceed 108 inches, where by "girth" we mean the perimeter of the smallest end. Our goal is to find the largest possible volume of a rectangular parcel with a square end that can be sent by mail. (We solved this applied optimization problem in single variable *Active Calculus*, so it may look familiar. We take a different approach in this section, and this approach allows us to view most applied optimization problems from single variable calculus as constrained optimization problems, as well as provide us tools to solve a greater variety of optimization problems.) If we let x be the length of the side of one square end of the package and y the length of the package, then we want to maximize the volume $f(x, y) = x^2 y$ of the box subject to the constraint that the girth $(4x)$ plus the length (y) is as large as possible, or $4x + y = 108$. The equation $4x + y = 108$ is thus an external constraint on the variables.

a. The constraint equation involves the function g that is given by

$$g(x, y) = 4x + y.$$

Explain why the constraint is a contour of g, and is therefore a two-dimensional curve.

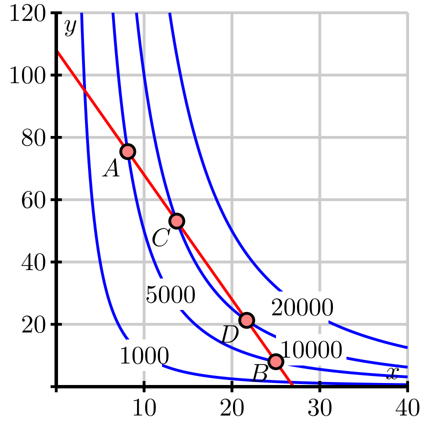


Figure 11.9.1 Contours of f and the constraint equation $g(x, y) = 108$.

b. Figure 11.9.1 shows the graph of the constraint equation $g(x, y) = 108$ along with a few contours of the volume function f. Since our goal is to find the maximum value of f subject to the constraint $g(x, y) = 108$, we want to find the point on our constraint curve that intersects the contours of f at which f has its largest value.

   i. Points A and B in Figure 11.9.1 lie on a contour of f and on the constraint equation $g(x, y) = 108$. Explain why neither A nor B provides a maximum value of f that satisfies the constraint.

   ii. Points C and D in Figure 11.9.1 lie on a contour of f and on the constraint equation $g(x, y) = 108$. Explain why neither C nor D provides a maximum value of f that satisfies the constraint.

   iii. Based on your responses to parts i. and ii., draw the contour of f on which you believe f will achieve a maximum value subject to the constraint $g(x, y) = 108$. Explain why you drew the contour you did.

c. Recall that $g(x, y) = 108$ is a contour of the function g, and that the gradient of a function is always orthogonal to its contours. With this in mind, how should ∇f and ∇g be related at the optimal point? Explain.

11.9.1 Constrained Optimization and Lagrange Multipliers

In Preview Activity 11.9.1, 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 108 inches. We saw that we can create a function g from the constraint, specifically $g(x, y) = 4x + y$. The constraint equation is then just a contour of g, $g(x, y) = c$, where c is a constant (in our case 108). Figure 11.9.2 illustrates that the volume function f is maximized, subject to the constraint $g(x, y) = c$, when the graph of $g(x, y) = c$ is tangent to a contour of f. Moreover, the value of f on this contour is the sought maximum value.

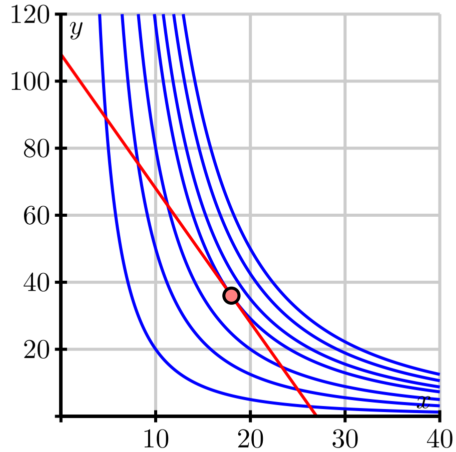


Figure 11.9.2 Contours of f and the constraint contour.

To find this point where the graph of the constraint is tangent to a contour of f, recall that ∇f is perpendicular to the contours of f and ∇g is perpendicular to the contour of g. At such a point, the vectors ∇g and ∇f are parallel, and thus we need to determine the points where this occurs. Recall that two vectors are parallel if one is a nonzero scalar multiple of the other, so we therefore look for values of a parameter λ that make

$$\nabla f = \lambda \nabla g. \tag{11.9.1}$$

The constant λ is called a *Lagrange multiplier*.

To find the values of λ that satisfy (11.9.1) for the volume function in Preview Activity 11.9.1, we calculate both ∇f and ∇g. Observe that

$$\nabla f = 2xy\hat{\imath} + x^2\hat{\jmath} \qquad \text{and} \qquad \nabla g = 4\hat{\imath} + \hat{\jmath},$$

and thus we need a value of λ so that

$$2xy\hat{\imath} + x^2\hat{\jmath} = \lambda(4\hat{\imath} + \hat{\jmath}).$$

Equating components in the most recent equation and incorporating the original constraint, we have three equations

$$2xy = \lambda(4) \tag{11.9.2}$$
$$x^2 = \lambda(1) \tag{11.9.3}$$
$$4x + y = 108 \tag{11.9.4}$$

in the three unknowns x, y, and λ. First, note that if $\lambda = 0$, then equation (11.9.3) shows that $x = 0$. From this, Equation (11.9.4) tells us that $y = 108$. So the point $(0, 108)$ is a point we need to consider. Next, provided that $\lambda \neq 0$ (from which it follows that $x \neq 0$ by Equation (11.9.3)), we may divide both sides of Equation (11.9.2) by the corresponding sides of (11.9.3) to eliminate λ, and thus find that

$$\frac{2y}{x} = 4, \text{ so}$$
$$y = 2x.$$

Substituting into Equation (11.9.4) gives us

$$4x + 2x = 108$$

or

$$x = 18.$$

Thus we have $y = 2x = 36$ and $\lambda = x^2 = 324$ as another point to consider. So the points at which the gradients of f and g are parallel, and thus at which f may have a maximum or minimum subject to the constraint, are $(0, 108)$ and $(18, 36)$. By evaluating the function f at these points, we see that we maximize the volume when the length of the square end of the box is 18 inches and the length is 36 inches, for a maximum volume of $f(18, 36) = 11664$ cubic inches. Since $f(0, 108) = 0$, we obtain a minimum value at this point.

We summarize the process of Lagrange multipliers as follows.

The 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.

Activity 11.9.2 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. 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?

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 f and which plays the role of g in our discussion of Lagrange multipliers.)

c. Find λ and the values of your variables that satisfy Equation (11.9.1) in the context of this problem.

d. Determine the dimensions of the pop can that give the desired solution to this constrained optimization problem.

The method of Lagrange multipliers also works for functions of more than two variables.

Activity 11.9.3 Use the method of Lagrange multipliers to find the dimensions of the least expensive packing crate with a volume of 240 cubic feet when the material for the top costs \$2 per square foot, the bottom is \$3 per square foot and the sides are \$1.50 per square foot.

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 $g(x, y, z) = k$, the optimal point (x, y, z) lies on the level surface S defined by the constraint $g(x, y, z) = k$. As we did in Preview Activity 11.9.1, we can argue that the optimal value occurs at the level surface $f(x, y, z) = c$ that is tangent to S. Thus, the gradients of f and g are parallel at this optimal point. So, just as in the two variable case, we can optimize $f = f(x, y, z)$ subject to the constraint $g(x, y, z) = k$ by finding all points (x, y, z) that satisfy $\nabla f = \lambda \nabla g$ and $g(x, y, z) = k$.

11.9.2 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.

11.9.3 Exercises

1. Use Lagrange multipliers to find the maximum and minimum values of $f(x, y) = 4x - 3y$ subject to the constraint $x^2 + 2y^2 = 82$, 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 $(1, 1, 11)$ 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 $h(x, y) = x^2 + y^2 - \lambda(2x + 8y - 18)$ 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 $2x + 8y = 18$ 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 ______________, occuring at ______________ points (positive integer or "infinitely many").
Minimum value is ______________, occuring 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 ______________, occuring at ______________ points (positive integer or "infinitely many").
Minimum value is ______________, occuring 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 ____,
occuring at (______________________,
______________________,______________________).
Minimum value is ____,
occuring 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

$$\nabla f(x_0, y_0, z_0) \cdot \vec{r}\,'(t_0) = 0,$$
$$\nabla g(x_0, y_0, z_0) \cdot \vec{r}\,'(t_0) = 0, \text{ and}$$

$$\nabla h(x_0, y_0, z_0) \cdot \vec{r}'(t_0) = 0.$$

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 (or by approximately how much the optimal value of the function f will change if we increase the value of c by 1 unit).

c. Suppose that $\lambda = 324$ at the point where the package described in Preview Activity 11.9.1 has its maximum volume. Explain in context what the value 324 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 Double Riemann Sums and Double Integrals over Rectangles

Motivating Questions

- What is a double Riemann sum?
- How is the double integral of a continuous function $f = f(x, y)$ defined?
- What are two things the double integral of a function can tell us?

In single-variable calculus, recall that 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. The general process involved subdividing 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 will extend this process in this section to its three-dimensional analogs, double Riemann sums and double integrals over rectangles.

Preview Activity 12.1.1 In this activity we introduce the concept of a double Riemann sum.

a. Review the concept of the Riemann sum from single-variable calculus. Then, explain how we define the definite integral $\int_a^b f(x)\, dx$ of a continuous function of a single variable x on an interval $[a, b]$. Include a sketch of a continuous function on an interval $[a, b]$ with appropriate labeling in order to illustrate your definition.

b. In our upcoming study of integral calculus for multivariable functions, we will first extend the idea of the single-variable definite integral to functions of two variables over rectangular domains. To do so, we will need to understand how to partition a rectangle into subrectangles. Let R be rectangular domain $R = \{(x, y) : 0 \leq x \leq 6, 2 \leq y \leq 4\}$ (we can also represent this domain with the notation $[0, 6] \times [2, 4]$), as pictured in Figure 12.1.1.

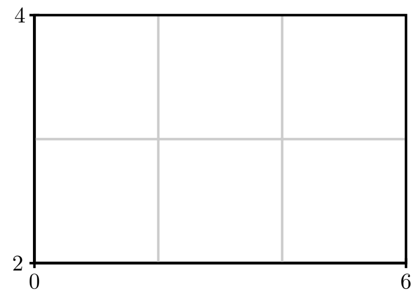


Figure 12.1.1 Rectangular domain R with subrectangles.

To form a partition of the full rectangular region, R, we will partition both intervals $[0, 6]$ and $[2, 4]$; in particular, we choose to partition the interval $[0, 6]$ into three uniformly sized subintervals and the interval $[2, 4]$ into two evenly sized subintervals as shown in Figure 12.1.1. In the following questions, we discuss how to identify the endpoints of each subinterval and the resulting subrectangles.

i. Let $0 = x_0 < x_1 < x_2 < x_3 = 6$ be the endpoints of the subintervals of $[0, 6]$ after partitioning. What is the length Δx of each subinterval $[x_{i-1}, x_i]$ for i from 1 to 3?

ii. Explicitly identify x_0, x_1, x_2, and x_3. On Figure 12.1.1 or your own version of the diagram, label these endpoints.

iii. Let $2 = y_0 < y_1 < y_2 = 4$ be the endpoints of the subintervals of $[2, 4]$ after partitioning. What is the length Δy of each subinterval $[y_{j-1}, y_j]$ for j from 1 to 2? Identify y_0, y_1, and y_2 and label these endpoints on Figure 12.1.1.

iv. Let R_{ij} denote the subrectangle $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$. Appropriately label each subrectangle in your drawing of Figure 12.1.1. How does the total number of subrectangles depend on the partitions of the intervals $[0, 6]$ and $[2, 4]$?

v. What is area ΔA of each subrectangle?

12.1.1 Double Riemann Sums over Rectangles

For the definite integral in single-variable calculus, we considered a continuous function over a closed, bounded interval $[a, b]$. In multivariable calculus, we will eventually develop the idea of a definite integral over a closed, bounded region (such as the interior of a circle). We begin with a simpler situation by thinking only about rectangular domains, and will address more complicated domains in Section 12.3.

Let $f = f(x, y)$ be a continuous function defined on a rectangular domain $R = \{(x, y) : a \le x \le b, c \le y \le d\}$. As we saw in Preview Activity 12.1.1, the domain is a rectangle R and we want to partition R into subrectangles. We do this by partitioning each of the intervals $[a, b]$ and $[c, d]$ into subintervals and using those subintervals to create a partition of R into subrectangles. In the first activity, we address the quantities and notations we will use in order to define double Riemann sums and double integrals.

Activity 12.1.2 Let $f(x, y) = 100 - x^2 - y^2$ be defined on the rectangular domain $R = [a, b] \times [c, d]$. Partition the interval $[a, b]$ into four uniformly sized subintervals and the interval $[c, d]$ into three evenly sized subintervals as shown in Figure 12.1.2. As we did in Preview Activity 12.1.1, we will need a method for identifying the endpoints of each subinterval and the resulting subrectangles.

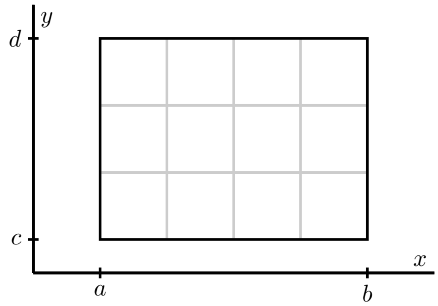


Figure 12.1.2 Rectangular domain with subrectangles.

a. Let $a = x_0 < x_1 < x_2 < x_3 < x_4 = b$ be the endpoints of the subintervals of $[a, b]$ after partitioning. Label these endpoints in Figure 12.1.2.

b. What is the length Δx of each subinterval $[x_{i-1}, x_i]$? Your answer should be in terms of a and b.

c. Let $c = y_0 < y_1 < y_2 < y_3 = d$ be the endpoints of the subintervals of $[c, d]$ after partitioning. Label these endpoints in Figure 12.1.2.

d. What is the length Δy of each subinterval $[y_{j-1}, y_j]$? Your answer should be in terms of c and d.

e. The partitions of the intervals $[a, b]$ and $[c, d]$ partition the rectangle R into subrectangles. How many subrectangles are there?

f. Let R_{ij} denote the subrectangle $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$. Label each subrectangle in Figure 12.1.2.

g. What is area ΔA of each subrectangle?

h. Now let $[a, b] = [0, 8]$ and $[c, d] = [2, 6]$. Let (x^*_{11}, y^*_{11}) be the point in the upper right corner of the subrectangle R_{11}. Identify and correctly label this point in Figure 12.1.2. Calculate the product

$$f(x^*_{11}, y^*_{11})\Delta A.$$

Explain, geometrically, what this product represents.

i. For each i and j, choose a point (x^*_{ij}, y^*_{ij}) in the subrectangle $R_{i,j}$. Identify and correctly label these points in Figure 12.1.2. Explain what the product

$$f(x^*_{ij}, y^*_{ij})\Delta A$$

represents.

j. If we were to add all the values $f(x^*_{ij}, y^*_{ij})\Delta A$ for each i and j, what does the resulting number approximate about the surface defined by f on the domain R? (You don't actually need to add these values.)

k. Write a double sum using summation notation that expresses the arbitrary sum from part (j).

12.1.2 Double Riemann Sums and Double Integrals

Now we use the process from the most recent activity to formally define double Riemann sums and double integrals.

Definition 12.1.3 Let f be a continuous function on a rectangle $R = \{(x, y) : a \le x \le b, c \le y \le d\}$. A **double Riemann sum for f over R** is created as follows.

- -.

  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 $a = x_0 < x_1 < x_2 < \cdots < x_m = b$.

- -.

  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 $c = y_0 < y_1 < y_2 < \cdots < y_n = d$.

- -.

  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 $\Delta A = \Delta x \cdot \Delta y$.

- Choose a point (x^*_{ij}, y^*_{ij}) in each rectangle R_{ij}. Then, a double Riemann sum for f over R is given by

$$\sum_{j=1}^{n}\sum_{i=1}^{m} f(x^*_{ij}, y^*_{ij}) \cdot \Delta A.$$

◊

If $f(x, y) \ge 0$ on the rectangle R, we may ask to find the volume of the solid bounded above by f over R, as illustrated on the left of Figure 12.1.4. This volume is approximated by a Riemann sum, which sums the volumes of the rectangular boxes shown on the right of Figure 12.1.4.

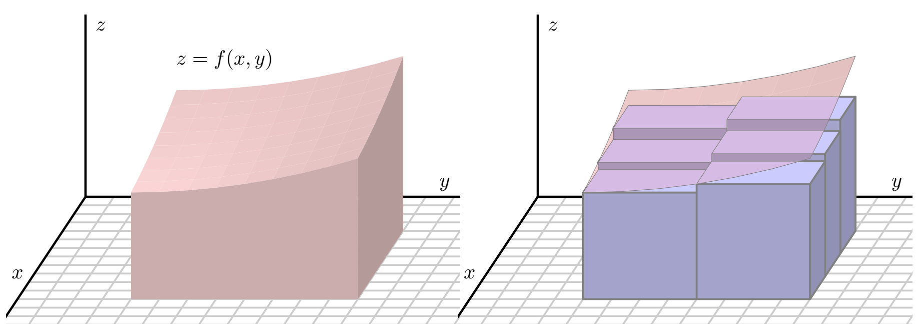


Figure 12.1.4 The volume under a graph approximated by a Riemann Sum.

As we let the number of subrectangles increase without bound (in other words, as both m and n in a double Riemann sum go to infinity), as illustrated in Figure 12.1.5, the sum of the volumes of the rectangular boxes approaches the volume of the solid bounded above by f over R. The value of this limit, provided it exists, is the double integral.

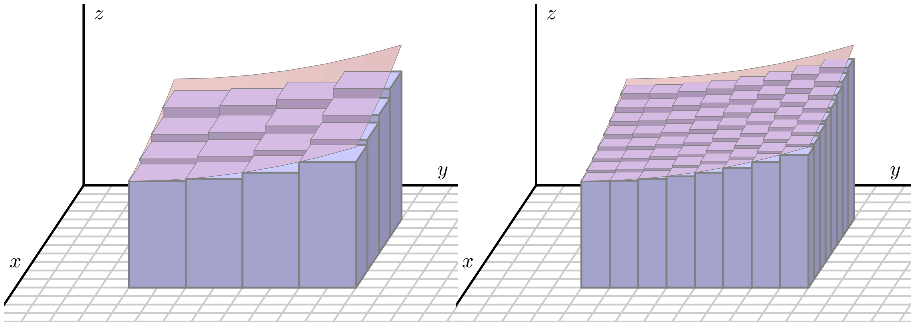


Figure 12.1.5 Finding better approximations by using smaller subrectangles.

Definition 12.1.6 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.

12.1.3 Interpretation of Double Riemann Sums and Double integrals.

At the moment, there are two ways we can interpret the value of the double integral.

- Suppose that $f(x,y)$ assumes both positive and negatives values on the rectangle R, as shown on the left of Figure 12.1.7. When constructing a Riemann sum, for each i and j, the product $f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$ 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

  $$\sum_{j=1}^{n} \sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$$

  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.

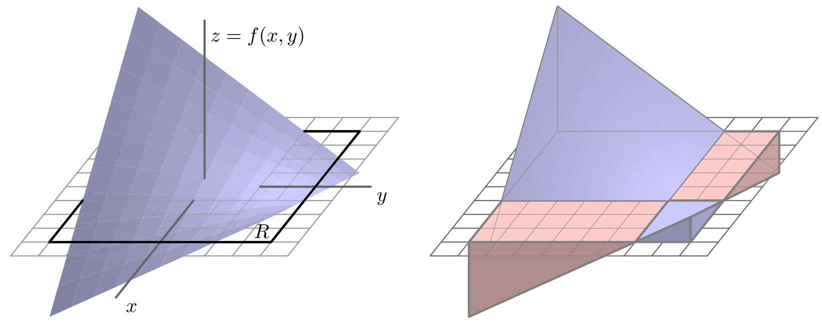


Figure 12.1.7 The integral measures signed volume.

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 on the right of Figure 12.1.7.

- The average of the finitely many mn values $f(x_{ij}^*, y_{ij}^*)$ that we take in a double Riemann sum is given by

$$\text{Avg}_{mn} = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*).$$

If we take the limit as m and n go to infinity, we obtain what we define as the average value of f over the region R, which is connected to the value of the double integral. First, to view Avg $_{mn}$ as a double Riemann sum, note that

$$\Delta x = \frac{b-a}{m} \qquad \text{and} \qquad \Delta y = \frac{d-c}{n}.$$

Thus,

$$\frac{1}{mn} = \frac{\Delta x \cdot \Delta y}{(b-a)(d-c)} = \frac{\Delta A}{A(R)},$$

where $A(R)$ denotes the area of the rectangle R. Then, the average value of the function f over R, $f_{\text{AVG}(R)}$, is given by

$$\begin{aligned} f_{\text{AVG}(R)} &= \lim_{m,n\to\infty} \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \\ &= \lim_{m,n\to\infty} \frac{1}{A(R)}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*)\cdot \Delta A \\ &= \frac{1}{A(R)}\iint_R f(x,y)\,dA. \end{aligned}$$

Therefore, the double integral of f over R divided by the area of R gives us the average value of the function f on R. Finally, 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.

Activity 12.1.3 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. Give two interpretations for the meaning of the sum you just calculated.

Activity 12.1.4 Let $f(x,y) = \sqrt{4-y^2}$ on the rectangular domain $R = [1,7] \times [-2,2]$. Partition $[1,7]$ into 3 equal length subintervals and $[-2,2]$ into 2 equal length subintervals. A table of values of f at some points in R is given in Table 12.1.8, and a graph of f with the indicated partitions is shown in Figure 12.1.9.

	-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

Table 12.1.8 Table of values of $f(x,y) = \sqrt{4-y^2}$.

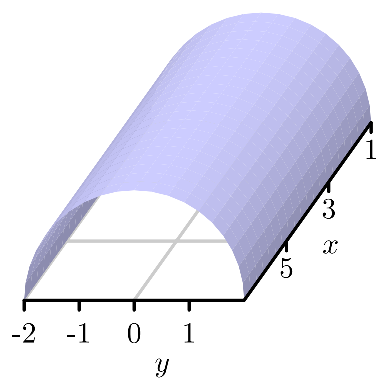


Figure 12.1.9 Graph of $f(x,y) = \sqrt{4-y^2}$ on R.

a. Sketch the region R in the plane using the values in Table 12.1.8 as the partitions.

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. Describe one way we could obtain a better approximation using the given data.

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. $\iint_R (f(x, y) + g(x, y))\, dA = \iint_R f(x, y)\, dA + \iint_R g(x, y)\, dA$.
2. $\iint_R kf(x, y)\, dA = k \iint_R f(x, y)\, dA$.
3. If $f(x, y) \geq g(x, y)$ on R, then $\iint_R f(x, y)\, dA \geq \iint_R g(x, y)\, dA$.

12.1.4 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.
  - -.
    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 $a = x_0 < x_1 < x_2 < \cdots < x_m = b$.
  - -.
    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 $c = y_0 < y_1 < y_2 < \cdots < y_n = d$.
  - -.
    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 $\Delta A = \Delta x \cdot \Delta y$.
  - -.
    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

$$\iint_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f(x_{ij}^*, y_{ij}^*) \cdot \Delta A.$$

- Two 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;
  - -.
    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.

12.1.5 Exercises

1. Suppose $f(x,y) = 25 - x^2 - y^2$ and R is the rectangle with vertices (0,0), (6,0), (6,4), (0,4). In each part, estimate $\displaystyle\iint\limits_R f(x,y)\,dA$ 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 = ________________________

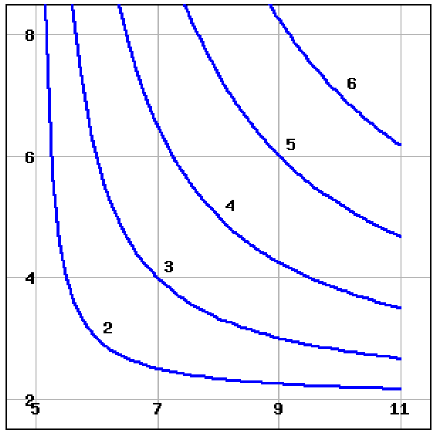
2. Consider the solid that lies above the square (in the xy-plane) $R = [0,1] \times [0,1]$, and below the elliptic paraboloid $z = 100 - x^2 + 6xy - 2y^2$.
Estimate the volume by dividing R into 9 equal squares and choosing the sample points to lie in the midpoints of each square.

3. Let R be the rectangle with vertices $(0,0)$, $(2,0)$, $(2,2)$, and $(0,2)$ and let $f(x,y) = \sqrt{0.25xy}$.
(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: $\int_R f\,dA \approx$ ________________________

4. Using Riemann sums with four subdivisions in each direction, find upper and lower bounds for the volume under the graph of $f(x,y) = 4 + xy$ above the rectangle R with $0 \le x \le 1, \quad 0 \le y \le 5$.
upper bound = ________________________
lower bound = ________________________

5. Consider the solid that lies above the square (in the xy-plane) $R = [0,2] \times [0,2]$,
and below the elliptic paraboloid $z = 36 - x^2 - 2y^2$.
(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)?

6. The figure below shows contours of $g(x,y)$ on the region R, with $5 \le x \le 11$ and $2 \le y \le 8$.

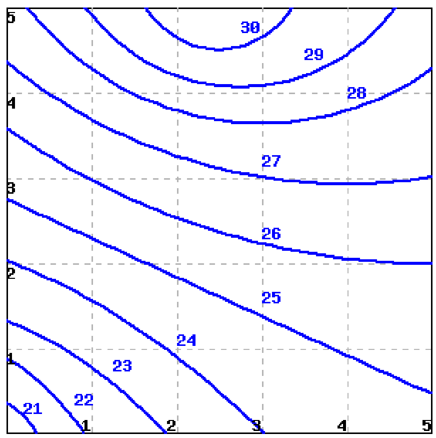


Using $\Delta x = \Delta y = 2$, find an overestimate and an underestimate for $\int_R g(x,y)\,dA$.

Overestimate = ______________________________

Underestimate = ______________________________

7. 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 = ______________________________

8. 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	4	0	-5
2.6	-3	0	8
3.0	6	6	-4

$\int_R f(x,y)\,da \approx$ ____________________

9. Values of $f(x,y)$ are shown in the table below.

	$x = 3$	$x = 3.2$	$x = 3.4$
$y = 5$	7	8	11
$y = 5.4$	6	7	8
$y = 5.8$	5	6	17

Let R be the rectangle $3 \leq x \leq 3.4$, $5 \leq y \leq 5.8$. 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 $\Delta y = 0.4$.
over-estimate: ________________________
under-estimate: ________________________

10. 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, $T_{\text{AVG}(R)}$, over the region R.

d. Do you think your estimate in (c) is an over- or under-estimate of the true temperature? Why?

11. 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.

12. The wind chill, as frequently reported, is a measure of how cold it feels outside when the wind is blowing. In Table 12.1.10, 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 12.1.10 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

13. 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?

12.2 Iterated Integrals

Motivating Questions

- How do we evaluate a double integral over a rectangle as an iterated integral, and why does this process work?

Recall that we defined the double integral of a continuous function $f = f(x, y)$ over a rectangle $R = [a, b] \times [c, d]$ as

$$\iint_R f(x, y)\, dA = \lim_{m,n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{m} f\left(x_{ij}^*, y_{ij}^*\right) \cdot \Delta A,$$

where the different variables and notation are as described in Section 12.1. Thus $\iint_R f(x, y)\, dA$ is a limit of double Riemann sums, but while this definition tells us exactly what a double integral is, it is not very helpful for determining 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 12.2.1, we saw that in some circumstances we could slice the solid perpendicular to an axis and have each slice be approximately a circular disk. From there, we were able to find the volume of each disk, and then use an integral to add the volumes of the slices. In what follows, we are able to use single integrals to generalize this approach to handle even more general geometric shapes.

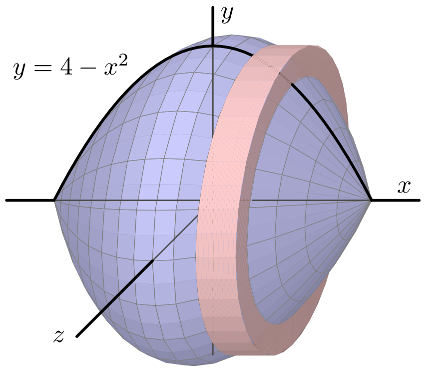


Figure 12.2.1 A solid of revolution.

Preview Activity 12.2.1 Let $f(x,y) = 25 - x^2 - y^2$ on the rectangular domain $R = [-3,3] \times [-4,4]$.

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. Now we investigate what happens if we integrate instead of differentiate.

a. Choose a fixed value of x in the interior of $[-3,3]$. Let

$$A(x) = \int_{-4}^{4} f(x,y)\, dy.$$

What is the geometric meaning of the value of $A(x)$ relative to the surface defined by f. (Hint: Think about the trace determined by the fixed value of x, and consider how $A(x)$ is related to the image at left in Figure 12.2.2.)

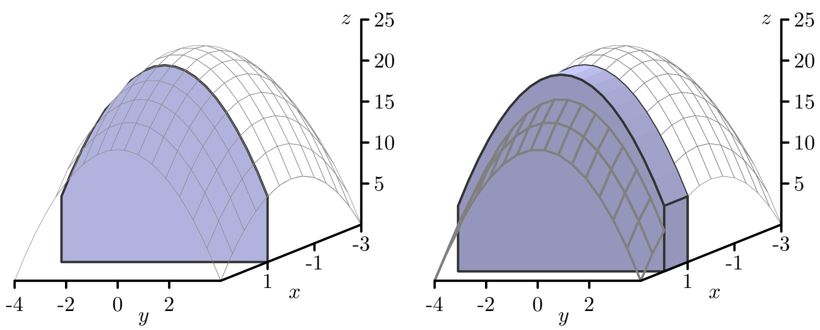


Figure 12.2.2 Left: A cross section with fixed x. Right: A cross section with fixed x and Δx.

b. For a fixed value of x, say x_i^*, what is the geometric meaning of $A(x_i^*)\,\Delta x$? (Hint: Consider how $A(x_i^*)\Delta x$ is related to the image at right in Figure 12.2.2.)

c. Since f is continuous on R, we can define the function $A = A(x)$ at every value of x in $[-3,3]$. Now think about subdividing the x-interval $[-3,3]$

into m subintervals, and choosing a value x_i^* in each of those subintervals. What will be the meaning of the sum $\sum_{i=1}^{m} A(x_i^*)\ \Delta x$?

d. Explain why $\int_{-3}^{3} A(x)\, dx$ will determine the exact value of the volume under the surface $z = f(x, y)$ over the rectangle R.

12.2.1 Iterated Integrals

The ideas that we explored in Preview Activity 12.2.1 work more generally and lead to the idea of an iterated integral. Let f be a continuous function on a rectangular domain $R = [a, b] \times [c, d]$, and let

$$A(x) = \int_c^d f(x, y)\, dy.$$

The function $A = A(x)$ determines the value of the cross sectional area (by area we mean "signed" area) in the y direction for the fixed value of x of the solid bounded between the surface defined by f and the xy-plane.

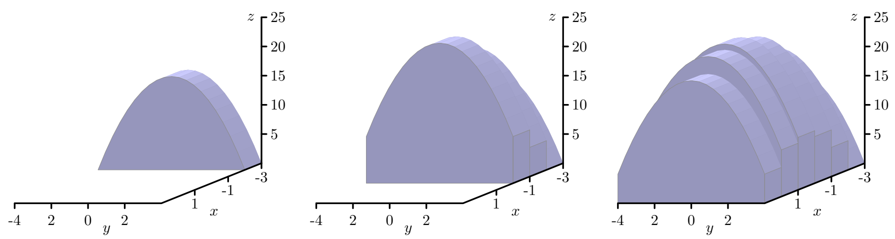


Figure 12.2.3 Summing volumes of cross section slices.

The value of this cross sectional area is determined by the input x in A. Since A is a function of x, it follows that we can integrate A with respect to x. In doing so, we use a partition of $[a, b]$ and make an approximation to the integral given by

$$\int_a^b A(x)\, dx \approx \sum_{i=1}^{m} A(x_i^*)\Delta x,$$

where x_i^* is any number in the subinterval $[x_{i-1}, x_i]$. Each term $A(x_i^*)\Delta x$ in the sum represents an approximation of a fixed cross sectional slice of the surface in the y direction with a fixed width of Δx as illustrated in Figure 12.2.3. We add the signed volumes of these slices as shown in the frames in Figure 12.2.3 to obtain an approximation of the total signed volume.

As we let the number of subintervals in the x direction approach infinity, we can see that the Riemann sum $\sum_{i=1}^{m} A(x_i^*)\Delta x$ approaches a limit and that limit is the sum of signed volumes bounded by the function f on R. Therefore, since $A(x)$ is itself determined by an integral, we have

$$\iint_R f(x, y)\, dA = \lim_{m\to\infty} \sum_{i=1}^{m} A(x_i^*)\Delta x = \int_a^b A(x)\, dx = \int_a^b \left(\int_c^d f(x, y)\, dy \right)\, dx.$$

Hence, we can compute the double integral of f over R by first integrating f with respect to y on $[c,d]$, then integrating the resulting function of x with respect to x on $[a,b]$. The nested integral

$$\int_a^b \left(\int_c^d f(x,y)\, dy \right) dx = \int_a^b \int_c^d f(x,y)\, dy\, dx$$

is called an *iterated integral*, and we see that each double integral may be represented by two single integrals.

We made a choice to integrate first with respect to y. The same argument shows that we can also find the double integral as an iterated integral integrating with respect to x first, or

$$\iint_R f(x,y)\, dA = \int_c^d \left(\int_a^b f(x,y)\, dx \right) dy = \int_c^d \int_a^b f(x,y)\, dx\, dy.$$

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 iterated integrals without resorting to the limit definition. Instead, 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.

Activity 12.2.2 Let $f(x,y) = 25 - x^2 - y^2$ on the rectangular domain $R = [-3,3] \times [-4,4]$.

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 $\iint_R f(x,y)\, dA$? What are two different ways we may interpret the meaning of this value?

Activity 12.2.3 Let $f(x,y) = x + y^2$ on the rectangle $R = [0,2] \times [0,3]$.

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).

12.2.2 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:

  - -.
    $\int_a^b \left(\int_c^d f(x,y)\,dy \right)\,dx$, or
  - -.
    $\int_c^d \left(\int_a^b f(x,y)\,dx \right)\,dy$.

  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.

12.2.3 Exercises

1. Evaluate the iterated integral $\int_0^4 \int_0^3 3x^2y^3\,dxdy$

2. Evaluate the iterated integral $\int_4^5 \int_2^3 (2x+y)^{-2}\,dydx$

3. Find $\int_0^6 \int_7^8 (x + \ln y)\,dydx$

4. Find $\int_0^4 \int_0^5 xye^{x+y}\,dydx$

5. Calculate the double integral $\int\int_{\mathbf{R}} (4x+4y+16)\,dA$ where **R** is the region: $0 \le x \le 2, 0 \le y \le 2$.

6. Calculate the double integral $\int\int_{\mathbf{R}} x\cos(x+y)\,dA$ where **R** is the region: $0 \le x \le \frac{\pi}{3}, 0 \le y \le \frac{\pi}{2}$

7. 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.

8. If $\displaystyle\int_1^2 f(x)\ dx = -3$ and $\displaystyle\int_3^6 g(x)\ dx = -2$, what is the value of $\displaystyle\int\int_D f(x)g(y)\,dA$ where D is the rectangle: $1 \le x \le 2, \quad 3 \le y \le 6$?

9. Find the average value of $f(x,y) = 4x^6y^2$ over the rectangle R with vertices $(-2,0), (-2,3), (2,0), (2,3)$.

Average value = ____________________

10. Find the average value of $f(x,y) = 7e^y\sqrt{x+e^y}$ over the rectangle $R = [0,6] \times [0,4]$.

Average value = ____________________

11. Evaluate each of the following double or iterated integrals exactly.

a. $\displaystyle\int_1^3 \left(\int_2^5 xy\,dy \right)\,dx$

b. $\displaystyle\int_0^{\pi/4}\left(\int_0^{\pi/3}\sin(x)\cos(y)\,dx\right)dy$

c. $\displaystyle\int_0^1\left(\int_0^1 e^{-2x-3y}\,dy\right)dx$

d. $\iint_R \sqrt{2x+5y}\,dA$, where $R=[0,2]\times[0,3]$.

12. 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.

13. 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?

12.3 Double Integrals over General Regions

Motivating Questions

- How do we define a double integral over a non-rectangular region?
- What general form does an iterated integral over a non-rectangular region have?

Recall that we defined the double integral of a continuous function $f=f(x,y)$ over a rectangle $R=[a,b]\times[c,d]$ as

$$\iint_R f(x,y)\,dA=\lim_{m,n\to\infty}\sum_{j=1}^{n}\sum_{i=1}^{m} f(x_{ij}^*,y_{ij}^*)\cdot\Delta A,$$

where the notation is as described in Section 12.1. Furthermore, we have seen that we can evaluate a double integral $\iint_R f(x,y)\,dA$ over R as an iterated integral of either of the forms

$$\int_a^b\int_c^d f(x,y)\,dy\,dx \qquad \text{or} \qquad \int_c^d\int_a^b f(x,y)\,dx\,dy.$$

It is natural to wonder how we might define and evaluate a double integral over a non-rectangular region; we explore one such example in the following preview activity.

Preview Activity 12.3.1 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 at left in Figure 12.3.1. If we place one vertex at the origin and let vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ be determined by the edges of the tetrahedron that have one end at the origin, then a formula that tells us the volume V of the tetrahedron is

$$V = \frac{1}{6}|\vec{a} \cdot (\vec{b} \times \vec{c})|. \tag{12.3.1}$$

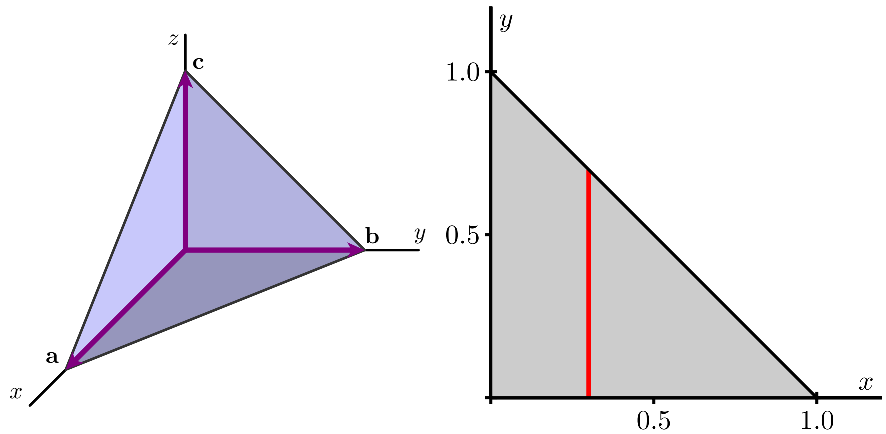


Figure 12.3.1 Left: The tetrahedron T. Right: Projecting T onto the xy-plane.

a. Use the formula (12.3.1) to find the volume of the tetrahedron T.

b. Instead of memorizing or looking up the formula for the volume of a tetrahedron, we can use a double integral to calculate the volume of the tetrahedron T. To see how, notice that the top face of the tetrahedron T is the plane whose equation is

$$z = 1 - (x + y).$$

Provided that we can use an iterated integral on a non-rectangular region, the volume of the tetrahedron will be given by an iterated integral of the form

$$\int_{x=?}^{x=?} \int_{y=?}^{y=?} 1 - (x + y)\, dy\, dx.$$

The issue that is new here is how we find the limits on the integrals; note that the outer integral's limits are in x, while the inner ones are in y, since we have chosen $dA = dy\, dx$. To see the domain over which we need to integrate, think of standing way above the tetrahedron looking straight down on it, which means we are projecting the entire tetrahedron onto the xy-plane. The resulting domain is the triangular region shown at right in Figure 12.3.1. Explain why we can represent the triangular region with the inequalities

$$0 \leq y \leq 1 - x \quad \text{and} \quad 0 \leq x \leq 1.$$

(Hint: Consider the cross sectional slice shown at right in Figure 12.3.1.)

c. Explain why it makes sense to now write the volume integral in the form

$$\int_{x=?}^{x=?} \int_{y=?}^{y=?} 1-(x+y)\, dy\, dx = \int_{x=0}^{x=1} \int_{y=0}^{y=1-x} 1-(x+y)\, dy\, dx.$$

d. Use the Fundamental Theorem of Calculus to evaluate the iterated integral

$$\int_{x=0}^{x=1} \int_{y=0}^{y=1-x} 1-(x+y)\, dy\, dx$$

and compare to your result from part (a). (As with iterated integrals over rectangular regions, start with the inner integral.)

12.3.1 Double Integrals over General Regions

So far, we have learned that a double integral over a rectangular region may be interpreted in one of two ways:

- $\iint_R f(x,y)\, dA$ 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;
- $\frac{1}{A(R)} \iint_R f(x,y)\, dA$, where $A(R)$ is the area of R tells 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.

As we saw in Preview Activity 12.1.1, 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 and evaluate double integrals over non-rectangular regions. Note that if we can, then the two interpretations of the double integral noted above will naturally extend to solid regions with non-rectangular bases.

So, suppose f is a continuous function on a closed, bounded domain D. For example, consider D as the circular domain shown at left in Figure 12.3.2.

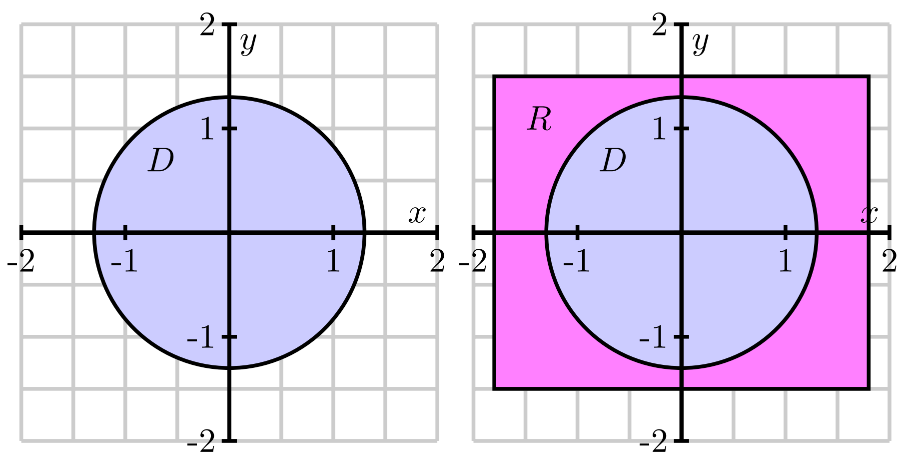


Figure 12.3.2 Left: A non-rectangular domain. Right: Enclosing this domain in a rectangle.

We can enclose D in a rectangular domain R as shown at right in Figure 12.3.2 and extend the function f to be defined over R in order to be able to use the definition of the double integral over a rectangle. We extend f in such a way that its values at the points in R that are not in D contribute 0 to the value of the integral. In other words, define a function $F = F(x, y)$ on R as

$$F(x, y) = \begin{cases} f(x, y), & \text{if } (x, y) \in D, \\ 0, & \text{if } (x, y) \notin D \end{cases}.$$

We then say that the double integral of f over D is the same as the double integral of F over R, and thus

$$\iint_D f(x, y)\, dA = \iint_R F(x, y)\, dA.$$

In practice, we just ignore everything that is in R but not in D, since these regions contribute 0 to the value of the integral.

Just as with double integrals over rectangles, a double integral over a domain D can be evaluated as an iterated integral. If the region D can be described by the inequalities $g_1(x) \le y \le g_2(x)$ and $a \le x \le b$, where $g_1 = g_1(x)$ and $g_2 = g_2(x)$ are functions of only x, then

$$\iint_D f(x, y)\, dA = \int_{x=a}^{x=b} \int_{y=g_1(x)}^{y=g_2(x)} f(x, y)\, dy\, dx.$$

Alternatively, if the region D is described by the inequalities $h_1(y) \le x \le h_2(y)$ and $c \le y \le d$, where $h_1 = h_1(y)$ and $h_2 = h_2(y)$ are functions of only y, we have

$$\iint_D f(x, y)\, dA = \int_{y=c}^{y=d} \int_{x=h_1(y)}^{x=h_2(y)} f(x, y)\, dx\, dy.$$

The structure of an iterated integral is of particular note:

In an iterated double integral:

- the limits on the outer integral must be constants;
- the limits on the inner integral must be constants or in terms of only the remaining variable — that is, if the inner integral is with respect to y, then its limits may only involve x and constants, and vice versa.

We next consider a detailed example.

Example 12.3.3 Let $f(x, y) = x^2 y$ be defined on the triangle D with vertices $(0, 0)$, $(2, 0)$, and $(2, 3)$ as shown at left in Figure 12.3.4.

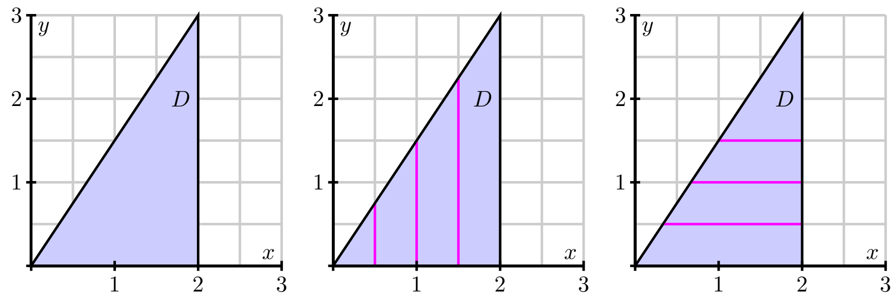


Figure 12.3.4 A triangular domain and slices in the y and x directions.

To evaluate $\iint_D f(x,y)\,dA$, we must first describe the region D in terms of the variables x and y. We take two approaches.

Approach 1: Integrate first with respect to y.

In this case we choose to evaluate the double integral as an iterated integral in the form

$$\iint_D x^2y\,dA = \int_{x=a}^{x=b}\int_{y=g_1(x)}^{y=g_2(x)} x^2y\,dy\,dx,$$

and therefore we need to describe D in terms of inequalities

$$g_1(x) \le y \le g_2(x) \qquad \text{and} \qquad a \le x \le b.$$

Since we are integrating with respect to y first, the iterated integral has the form

$$\iint_D x^2y\,dA = \int_{x=a}^{x=b} A(x)\,dx,$$

where $A(x)$ is a cross sectional area in the y direction. So we are slicing the domain perpendicular to the x-axis and want to understand what a cross sectional area of the overall solid will look like. Several slices of the domain are shown in the middle image in Figure 12.3.4. On a slice with fixed x value, the y values are bounded below by 0 and above by the y coordinate on the hypotenuse of the right triangle. Thus, $g_1(x) = 0$; to find $y = g_2(x)$, we need to write the hypotenuse as a function of x. The hypotenuse connects the points (0,0) and (2,3) and hence has equation $y = \frac{3}{2}x$. This gives the upper bound on y as $g_2(x) = \frac{3}{2}x$. The leftmost vertical cross section is at $x = 0$ and the rightmost one is at $x = 2$, so we have $a = 0$ and $b = 2$. Therefore,

$$\iint_D x^2y\,dA = \int_{x=0}^{x=2}\int_{y=0}^{y=\frac{3}{2}x} x^2y\,dy\,dx.$$

We evaluate the iterated integral by applying the Fundamental Theorem of Calculus first to the inner integral, and then to the outer one, and find that

$$\begin{aligned}
\int_{x=0}^{x=2}\int_{y=0}^{y=\frac{3}{2}x} x^2y\,dy\,dx &= \int_{x=0}^{x=2} \left[x^2\cdot\frac{y^2}{2}\right]\Bigg|_{y=0}^{y=\frac{3}{2}x} dx \\
&= \int_{x=0}^{x=2} \frac{9}{8}x^4\,dx \\
&= \frac{9}{8}\frac{x^5}{5}\Bigg|_{x=0}^{x=2} \\
&= \left(\frac{9}{8}\right)\left(\frac{32}{5}\right) \\
&= \frac{36}{5}.
\end{aligned}$$

Approach 2: Integrate first with respect to x.

In this case, we choose to evaluate the double integral as an iterated integral in the form

$$\iint_D x^2 y \, dA = \int_{y=c}^{y=d} \int_{x=h_1(y)}^{x=h_2(y)} x^2 y \, dx \, dy$$

and thus need to describe D in terms of inequalities

$$h_1(y) \leq x \leq h_2(y) \qquad \text{and} \qquad c \leq y \leq d.$$

Since we are integrating with respect to x first, the iterated integral has the form

$$\iint_D x^2 y \, dA = \int_c^d A(y) \, dy,$$

where $A(y)$ is a cross sectional area of the solid in the x direction. Several slices of the domain — perpendicular to the y-axis — are shown at right in Figure 12.3.4. On a slice with fixed y value, the x values are bounded below by the x coordinate on the hypotenuse of the right triangle and above by 2. So $h_2(y) = 2$; to find $h_1(y)$, we need to write the hypotenuse as a function of y. Solving the earlier equation we have for the hypotenuse ($y = \frac{3}{2}x$) for x gives us $x = \frac{2}{3}y$. This makes $h_1(y) = \frac{2}{3}y$. The lowest horizontal cross section is at $y = 0$ and the uppermost one is at $y = 3$, so we have $c = 0$ and $d = 3$. Therefore,

$$\iint_D x^2 y \, dA = \int_{y=0}^{y=3} \int_{x=(2/3)y}^{x=2} x^2 y \, dx \, dy.$$

We evaluate the resulting iterated integral as before by twice applying the Fundamental Theorem of Calculus, and find that

$$\begin{aligned}
\int_{y=0}^{y=3} \int_{x=\frac{2}{3}y}^{2} x^2 y \, dx \, dy &= \int_{y=0}^{y=3} \left[\frac{x^3}{3}\right]\Bigg|_{x=\frac{2}{3}y}^{x=2} y \, dx \\
&= \int_{y=0}^{y=3} \left[\frac{8}{3}y - \frac{8}{81}y^4\right] dy \\
&= \left[\frac{8}{3}\frac{y^2}{2} - \frac{8}{81}\frac{y^5}{5}\right]\Bigg|_{y=0}^{y=3} \\
&= \left(\frac{8}{3}\right)\left(\frac{9}{2}\right) - \left(\frac{8}{81}\right)\left(\frac{243}{5}\right) \\
&= 12 - \frac{24}{5} \\
&= \frac{36}{5}.
\end{aligned}$$

We see, of course, that in the situation where D can be described in two different ways, the order in which we choose to set up and evaluate the double integral doesn't matter, and the same value results in either case. □

The meaning of a double integral over a non-rectangular region, D, parallels the meaning over a rectangular region. In particular,

- $\iint_D f(x,y)\,dA$ tells us the volume of the solids the graph of f bounds above the xy-plane over the closed, bounded region D minus the volume of the solids the graph of f bounds below the xy-plane under the region D;
- $\frac{1}{A(D)}\iint_R f(x,y)\,dA$, where $A(D)$ is the area of D tells us the average value of the function f on D. If $f(x,y) \geq 0$ on D, we can interpret this average value of f on D as the height of the solid with base D and constant cross-sectional area D that has the same volume as the volume of the surface defined by f over D.

Activity 12.3.2 Consider the double integral $\iint_D (4-x-2y)\,dA$, where D is the triangular region with vertices (0,0), (4,0), and (0,2).

a. Write the given integral as an iterated integral of the form $\iint_D (4-x-2y)\,dy\,dx$. Draw a labeled picture of D with relevant cross sections.

b. Write the given integral as an iterated integral of the form $\iint_D (4-x-2y)\,dx\,dy$. Draw a labeled picture of D with relevant cross sections.

c. Evaluate the two iterated integrals from (a) and (b), and verify that they produce the same value. Give at least one interpretation of the meaning of your result.

Activity 12.3.3 Consider the iterated integral $\int_{x=0}^{x=1}\int_{y=x}^{y=\sqrt{x}} (4x+10y)\,dy\,dx$.

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 the limits of integration for which

$$\iint_D (4x+10y)\,dA = \int_{y=?}^{y=?}\int_{x=?}^{x=?} (4x+10y)\,dx\,dy.$$

c. Evaluate one of the two iterated integrals above. Explain what the value you obtained tells you.

d. Set up and evaluate a single definite integral to determine the exact area of D, $A(D)$.

e. Determine the exact average value of $f(x,y) = 4x+10y$ over D.

Activity 12.3.4 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 developed in (c). Write one sentence to explain the meaning of the value you found.

e. What is the important lesson this activity offers regarding the order in which we set up an iterated integral?

12.3.2 Summary

- For a double integral $\iint_D f(x,y)\,dA$ over a non-rectangular region D, we enclose D in a rectangle R and then extend integrand f to a function F so that $F(x,y)=0$ at all points in R outside of D and $F(x,y)=f(x,y)$ for all points in D. We then define $\iint_D f(x,y)\,dA$ to be equal to $\iint_R F(x,y)\,dA$.
- 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 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 $g_1(x)\le y\le g_2(x)$ and $a\le x\le b$ 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 $h_1(y)\le x\le h_2(y)$ and $c\le y\le d$.

12.3.3 Exercises

1. Evaluate the double integral $I=\displaystyle\iint_{\mathbf{D}} xy\ dA$ where **D** is the triangular region with vertices $(0,0),(6,0),(0,1)$.

2. Evaluate the double integral $I=\displaystyle\iint_{\mathbf{D}} xy\ dA$ where **D** is the triangular region with vertices $(0,0),(1,0),(0,5)$.

3. Evaluate the integral by reversing the order of integration.
$\int_0^1\int_{3y}^3 e^{x^2}\ dxdy =$ ______________________________

4. 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) $\displaystyle\iint\limits_R (y^3+y^5)\,dA$

(b) $\displaystyle\iint\limits_B (y^3+y^5)\,dA$

(c) $\displaystyle\iint\limits_D (y^3+y^5)\,dA$

(d) $\displaystyle\iint_L (y^3 + y^5)\, dA$

(e) $\displaystyle\iint_T (y^3 + y^5)\, dA$

5. The region W lies below the surface $f(x, y) = 7e^{-(x-3)^2-y^2}$ and above the disk $x^2 + y^2 \leq 16$ 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 $x = 3$.

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 $d =$__________

6. Set up a double integral in rectangular coordinates for calculating the volume of the solid under the graph of the function $f(x, y) = 22 - x^2 - y^2$ and above the plane $z = 6$.

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.

$\displaystyle\int_A^B \int_C^D$ ______________________ _____ _____

A = ____________________

B = ____________________

C = ____________________

D = ____________________

7. Find the volume of the solid bounded by the planes x = 0, y = 0, z = 0, and x + y + z = 6.

8. Consider the integral $\displaystyle\int_0^6 \int_0^{\sqrt{36-y}} f(x, y)dxdy$. If we change the order of integration we obtain the sum of two integrals:

$\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)dydx + \int_c^d \int_{g_3(x)}^{g_4(x)} f(x, y)dydx$

$a =$ ____________________ $b =$ ____________________

$g_1(x) =$ ____________________ $g_2(x) =$ ____________________

$c =$ ____________________ $d =$ ____________________

$g_3(x) =$ ____________________ $g_4(x) =$ ____________________

9. A pile of earth standing on flat ground has height 36 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 $x^2 + y^2 = 36 - z$ for $0 \leq z \leq 36$, with x, y, and z in meters.

(a) What equation gives the edge of the base of the pile?

⊙ $x + y = 36$

⊙ $x + y = 6$

⊙ $x^2 + y^2 = 36$

⊙ $x^2 + y^2 = 6$

⊙ 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 $z = 4$?

⊙ $x^2 + y^2 = 4$

⊙ $x^2 + y^2 = 32$

⊙ $x^2 + y^2 = 16$

⊙ $x^2 + y^2 = \sqrt{32}$

⊙ None of the above

(d) What is the area of the cross-section $z = 4$ 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

10. 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) $\displaystyle\int_0^{\frac{1}{\sqrt{3}}} \int_0^{\frac{1}{2}\sqrt{1-3y^2}} \sqrt{1 - 4x^2 - 3y^2}\, dxdy$

(b) $\displaystyle\int_0^2 \int_{-2}^2 \sqrt{4 - y^2}\, dydx$

(c) $\displaystyle\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 - x^2 - y^2\, dydx$

(d) $\displaystyle\int_{-2}^2 \int_4^{4+\sqrt{4-x^2}} 4x + 3y\, dydx$

(e) $\displaystyle\int_0^1 \int_{y^2}^{\sqrt{y}} 4x^2 + 3y^2\, dxdy$

A. Solid under an elliptic paraboloid and over a planar region bounded by two parabolas.

B. Solid under a plane and over one half of a circular disk.

C. One half of a cylindrical rod.

D. One eighth of an ellipsoid.

E. Solid bounded by a circular paraboloid and a plane.

11. 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. $\displaystyle\int_{x=0}^{x=1} \int_{y=x^2}^{y=x} xy\, dy\, dx$

b. $\displaystyle\int_{y=0}^{y=2} \int_{x=-\sqrt{4-y^2}}^{x=0} xy\, dx\, dy$

c. $\displaystyle\int_{x=0}^{x=1} \int_{y=x^4}^{y=x^{1/4}} x + y\, dy\, dx$

d. $\displaystyle\int_{y=0}^{y=2} \int_{x=y/2}^{x=2y} x + y\, dx\, dy$

12. 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.

13. 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 $x = \sqrt{y}$; $p(x, y) = \sqrt{1 + x^2 + y^2}$.

14. Consider the iterated integral $I = \displaystyle\int_{x=0}^{x=4} \int_{y=\sqrt{x}}^{y=2} \cos(y^3)\, dy\, dx$.

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.

12.4 Applications of Double Integrals

Motivating Questions

- If we have a mass density function for a lamina (thin plate), how does a double integral determine the mass of the lamina?
- How may a double integral be used to find the area between two curves?
- Given a mass density function on a lamina, how can we find the lamina's center of mass?
- What is a joint probability density function? How do we determine the probability of an event if we know a probability density function?

So far, we have interpreted the double integral of a function f over a domain D in two different ways. First, $\iint_D f(x, y)\, dA$ tells us a difference of volumes — the volume the surface defined by f bounds above the xy-plane on D minus the volume the surface bounds below the xy-plane on D. In addition, $\frac{1}{A(D)} \iint_D f(x, y)\, dA$ determines the average value of f on D. In this section, we investigate several other applications of double integrals, using the integration process as seen in Preview Activity 12.4.1: 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 how a double integral can be used to determine the density of a thin plate with a mass density distribution. 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 12.4.1 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. Suppose that we partition the plate into subrectangles R_{ij}, where $1 \leq i \leq m$ and $1 \leq j \leq n$, of equal area ΔA, and select a point (x_{ij}^*, y_{ij}^*) in R_{ij} for each i and j. What is the meaning of the quantity $\delta(x_{ij}^*, y_{ij}^*)\Delta A$?

b. State a double Riemann sum that provides 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. Determine 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 12.5.)

12.4.1 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 12.4.1, 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 12.4.2 Let D be a half-disk lamina of radius 3 in quadrants IV and I, centered at the origin as shown in Figure 12.4.1. Assume the density at point (x,y) is given by $\delta(x,y) = x$. Find the exact mass of the lamina.

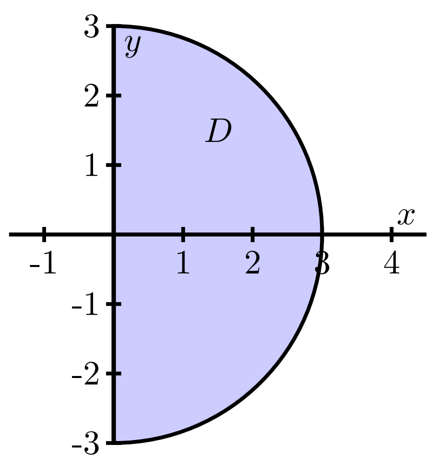


Figure 12.4.1 A half disk lamina.

12.4.2 Area

If we consider the situation where the mass-density distribution is constant, we can also see how a double integral may be used to determine the area of a region. Assuming that $\delta(x,y) = 1$ over a closed bounded region D, where the units of δ are "mass per unit of area," it follows that $\iint_D 1\,dA$ is the mass of the lamina. But since the density is constant, the numerical value of the integral is simply the area.

As the following activity demonstrates, we can also see this fact by considering a three-dimensional solid whose height is always 1.

Activity 12.4.3 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 12.4.2.

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?)

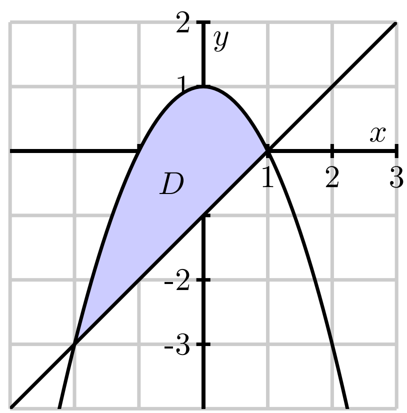


Figure 12.4.2 The graphs of $y = 1 - x^2$ and $y = x - 1$.

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 12.4.3.

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

$$A(D) = \iint_D 1\,dA.$$

12.4.3 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 $m_1, m_2, \ldots, m_N$ 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 12.4.3. 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.

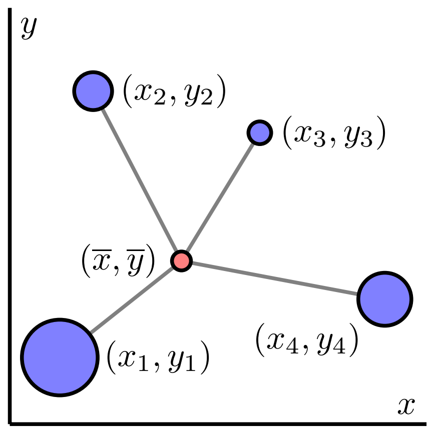


Figure 12.4.3 A center of mass $(\overline{x}, \overline{y})$ of four masses.

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

$$\overline{x} = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij}^* \delta(x_{ij}^*, y_{ij}^*)\Delta A}{\sum_{j=1}^{n} \sum_{i=1}^{m} \delta(x_{ij}^*, y_{ij}^*)\Delta A} \qquad \text{and} \qquad \overline{y} = \frac{\sum_{j=1}^{n} \sum_{i=1}^{m} y_{ij}^* \delta(x_{ij}^*, y_{ij}^*)\Delta A}{\sum_{j=1}^{n} \sum_{i=1}^{m} \delta(x_{ij}^*, y_{ij}^*)\Delta A}.$$

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

$$\overline{x} = \frac{\iint_D x\delta(x, y)\, dA}{\iint_D \delta(x, y)\, dA} \qquad \text{and} \qquad \overline{y} = \frac{\iint_D y\delta(x, y)\, dA}{\iint_D \delta(x, y)\, dA}.$$

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, or the center of mass if 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 12.4.4 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 12.4.4.

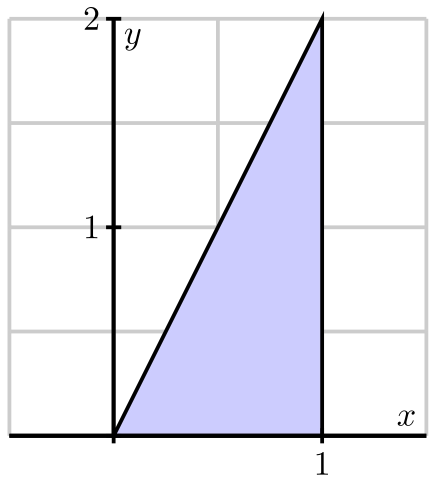


Figure 12.4.4 The lamina bounded by the x-axis and the lines $x = 1$ and $y = 2x$ in the first quadrant.

a. Set up an iterated integral that represents the mass of the lamina.

b. Assume the mass of the lamina is 14. Set up two iterated integrals that represent the coordinates of the center of mass of the lamina.

12.4.4 Probability

Calculating probabilities is a very 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. Let us 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, and is not 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 (assuming that the dart thrower always hits the board itself at some point). Similarly, the probability that the dart strikes a point in the disk D_3 of radius 3 inches 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, but we can easily envision situations where the distribution function varies. For example, if the player is fairly good and is aiming for the bulls eye (the center of D), then the distribution function f could be skewed toward the center, say

$$f(x, y) = Ke^{-(x^2+y^2)}$$

for some constant positive K. If we assume that the player is consistent enough so that the dart always strikes the board, then the probability that the dart strikes the board somewhere is 1, and the distribution function f will have to 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, 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

$$\int_a^b \int_c^d f(x,y)\, dy\, dx,$$

- The probability that the point (x,y) is in D is 1, that is

$$\iint_D f(x,y)\, dA = 1. \tag{12.4.1}$$

Note that it is possible that D could be an infinite region and the limits on the integral in Equation (12.4.1) 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 (the notation we use here is "$P((x,y) \in R)$") is determined by

$$P((x,y) \in R) = \iint_R f(x,y)\, dA.$$

Activity 12.4.5 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 might be given by

$$f(x,y) = e^{-x}e^{-y}.$$

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

[1] This makes $K = \frac{1}{\pi\left(1-e^{-100}\right)}$, which you can check.

that f is a probability density function on D we need to demonstrate that

$$\int\int_D f(x,y)\,dA = 1,$$

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?

12.4.5 Summary

- The mass of a lamina D with a mass density function $\delta = \delta(x,y)$ is $\iint_D \delta(x,y)\,dA$.
- 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 center of mass, $(\overline{x}, \overline{y})$, of a continuous lamina with a variable density $\delta(x,y)$ is given by

$$\overline{x} = \frac{\iint_D x\delta(x,y)\,dA}{\iint_D \delta(x,y)\,dA} \qquad \text{and} \qquad \overline{y} = \frac{\iint_D y\delta(x,y)\,dA}{\iint_D \delta(x,y)\,dA}.$$

- 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

$$\iint_R f(x,y)\,dA.$$

12.4.6 Exercises

1. The masses m_i are located at the points P_i. Find the center of mass of the system.

$m_1 = 1$, $m_2 = 8$, $m_3 = 3$.
$P_1 = (-9,5)$, $P_2 = (-9,4)$, $P_3 = (3,5)$.
$\bar{x}=$ ______________________
$\bar{y}=$ ______________________

2. Find the centroid $(\bar{x}, \bar{y})$ of the triangle with vertices at $(0,0)$, $(3,0)$, and $(0,2)$.

$\bar{x}$= ______________________________

$\bar{y}$= ______________________________

3. Find the mass of the rectangular region $0 \le x \le 1$, $0 \le y \le 4$ with density function $\rho\,(x,y) = 4 - y$.

4. Find the mass of the triangular region with vertices (0, 0), (3, 0), and (0, 4), with density function $\rho\,(x,y) = x^2 + y^2$.

5. A lamina occupies the region inside the circle $x^2 + y^2 = 8y$ but outside the circle $x^2 + y^2 = 16$. The density at each point is inversely proportional to its distance from the orgin.

Where is the center of mass?

(______________, ______________________________)

6. 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 6?

____ ft^3 per hour

B. What is the total amount of water that goes through the sprinkler per hour?

____ ft^3 per hour

7. Let p be the joint density function such that $p(x,y) = \frac{1}{81}xy$ in R, the rectangle $0 \le x \le 6, 0 \le y \le 3$, and $p(x,y) = 0$ outside R. Find the fraction of the population satisfying the constraint $x + y \le 9$

fraction = ______________________________

8. A lamp has two bulbs, each of a type with an average lifetime of 10 hours. The probability density function for the lifetime of a bulb is $f(t) = \frac{1}{10}e^{-t/10}, t \ge 0$.

What is the probability that both of the bulbs will fail within 3 hours?

9. 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) $p(x,y) = 3$, where R is $2 \le x \le 3, 2 \le y \le 4$.

$p(x,y)$ (☐ is a joint density function ☐ is not a joint density function)

(b) $p(x,y) = 1$, where R is $0 \le x \le 2, -1 \le y \le -0.5$.

$p(x,y)$ (☐ is a joint density function ☐ is not a joint density function)

Then, for the region R given by $-1 \le x \le 1, -1 \le y \le 2$, what constant function $p(x,y)$ is a joint density function?

$p(x,y) =$ ______________

10. 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 = ________________________

11. 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.

12. Let D be a half-disk lamina of radius 3 in quadrants IV and I, centered at the origin as in Activity 12.4.2. 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.

13. 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.

12.5 Double Integrals in Polar Coordinates

Motivating Questions

- What are the polar coordinates of a point in two-space?
- How do we convert between polar coordinates and rectangular coordinates?
- What is the area element in polar coordinates?
- How do we convert a double integral in rectangular coordinates to a double integral in polar coordinates?

While we have naturally defined double integrals in the rectangular coordinate system, starting with domains that are rectangular regions, there are many of these integrals that are difficult, if not impossible, to evaluate. For example, consider the domain D that is the unit circle and $f(x, y) = e^{-x^2-y^2}$. To integrate f over D, we would use the iterated integral

$$\iint_D f(x, y)\, dA = \int_{x=-1}^{x=1} \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} e^{-x^2-y^2}\, dy\, dx.$$

For this particular integral, regardless of the order of integration, we are unable to find an antiderivative of the integrand; in addition, even if we were able to find an antiderivative, the inner limits of integration involve relatively complicated functions.

It is useful, therefore, to be able to translate to other coordinate systems where the limits of integration and evaluation of the involved integrals is simpler. In this section we provide a quick discussion of one such system — polar coordinates — and then introduce and investigate their ramifications for double integrals. The rectangular coordinate system allows us to consider domains and graphs relative to a rectangular grid. The polar coordinate system is an alternate coordinate system that allows us to consider domains less suited to rectangular coordinates, such as circles.

Preview Activity 12.5.1 The coordinates of a point determine its location. In particular, the rectangular coordinates of a point P are given by an ordered pair (x, y), where x is the (signed) distance the point lies from the y-axis to P and y is the (signed) distance the point lies from the x-axis to P. In polar coordinates, we locate the point by considering the distance the point lies from the origin, $O = (0, 0)$, and the angle the line segment from the origin to P forms with the positive x-axis.

a. Determine the rectangular coordinates of the following points:

   (a) The point P that lies 1 unit from the origin on the positive x-axis.

   (b) The point Q that lies 2 units from the origin and such that $\overline{OQ}$ makes an angle of $\frac{\pi}{2}$ with the positive x-axis.

   (c) The point R that lies 3 units from the origin such that $\overline{OR}$ makes an angle of $\frac{2\pi}{3}$ with the positive x-axis.

b. Part (a) indicates that the two pieces of information completely determine the location of a point: either the traditional (x, y) coordinates, or alternately, the distance r from the point to the origin along with the angle θ that the line through the origin and the point makes with the positive x-axis. We write "(r, θ)" to denote the point's location in its polar coordinate representation. Find polar coordinates for the points with the given rectangular coordinates.

   i. $(0, -1)$ ii. $(-2, 0)$ iii. $(-1, 1)$

c. For each of the following points whose coordinates are given in polar form, determine the rectangular coordinates of the point.

   i. $(5, \frac{\pi}{4})$ ii. $(2, \frac{5\pi}{6})$ iii. $(\sqrt{3}, \frac{5\pi}{3})$

12.5.1 Polar Coordinates

The rectangular coordinate system is best suited for graphs and regions that are naturally considered over a rectangular grid. The polar coordinate system is an alternative that offers good options for functions and domains that have more circular characteristics. 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, as seen in Preview Activity 12.5.1, 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 at left in Figure 12.5.1.

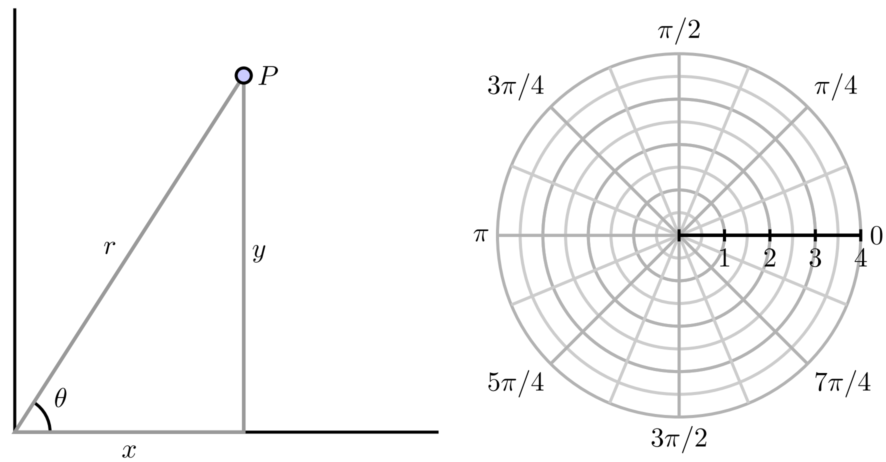


Figure 12.5.1 The polar coordinates of a point and the polar coordinate grid.

Trigonometry and the Pythagorean Theorem allow for straightforward conversion from rectangular to polar, and vice versa.

Converting between rectangular and polar coordinates.

- *Converting from rectangular to polar..*

  If we are given the rectangular coordinates (x, y) of a point P, then the polar coordinates (r, θ) of P satisfy

$$r = \sqrt{x^2 + y^2} \qquad \text{and} \qquad \tan(\theta) = \frac{y}{x}, \text{ assuming } x \neq 0.$$

- *Converting from polar to rectangular..*

  If we are given the polar coordinates (r, θ) of a point P, then the rectangular coordinates (x, y) of P satisfy

$$x = r\cos(\theta) \qquad \text{and} \qquad y = r\sin(\theta).$$

Note: The angle θ in the polar coordinates of a point is not unique. We could replace θ with $\theta + 2\pi$ and still be at the same terminal point. In addition, the sign of $\tan(\theta)$ does not uniquely determine the quadrant in which θ lies, so we have to determine the value of θ from the location of the point. In other words, more care has to be paid when using polar coordinates than rectangular 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 at right in Figure 12.5.1.

Activity 12.5.2 Most polar graphing devices can plot curves in polar coordinates of the form $r = f(\theta)$. Use such a device to complete this activity.

a. Before plotting the polar curve $r = 1$ (where θ can have any value), think about what shape it should have, in light of how r is connected to x and y. Then use appropriate technology to draw the graph and test your intuition.

b. The equation $\theta = 1$ does not define r as a function of θ, so we can't graph this equation on many polar plotters. What do you think the graph of the polar curve $\theta = 1$ looks like? Why?

c. Before plotting the polar curve $r = \theta$, what do you think the graph looks like? Why? Use technology to plot the curve and compare your intuition.

d. What does the region defined by $1 \leq r \leq 3$ (where θ can have any value) look like? (Hint: Compare to your response from part (a).)

e. What does the region defined by $1 \leq r \leq 3$ and $\pi/4 \leq \theta \leq \pi/2$ look like?

f. Consider the curve $r = \sin(\theta)$. For some values of θ we will have $r < 0$. In these situations, we plot the point (r, θ) as $(|r|, \theta + \pi)$ (in other words, when $r < 0$, we reflect the point through the origin). With that in mind, what do you think the graph of $r = \sin(\theta)$ looks like? Plot this curve using technology and compare to your intuition.

12.5.2 Integrating in Polar Coordinates

Consider the double integral

$$\iint_D e^{x^2+y^2}\, dA,$$

where D is the unit disk. While we cannot directly evaluate this integral in rectangular coordinates, a change to polar coordinates will convert it to one we can easily evaluate.

We have seen how to evaluate a double integral $\iint_D f(x,y)\, dA$ as an iterated integral of the form

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\, dy\, dx$$

in rectangular coordinates, because we know that $dA = dy\, dx$ in rectangular coordinates. To make the change to polar coordinates, we not only need to represent the variables x and y in polar coordinates, but we also must understand how to write the area element, dA, in polar coordinates. That is, we must determine how the area element dA can be written in terms of dr and $d\theta$ in the context of polar coordinates. We address this question in the following activity.

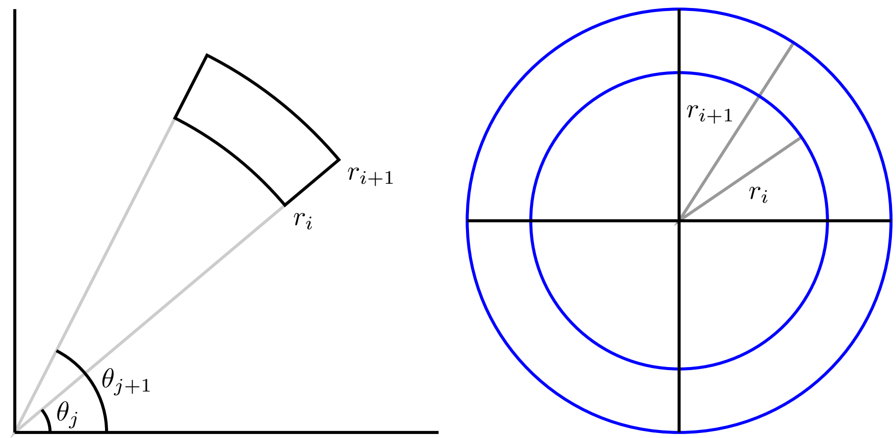


Figure 12.5.2 Left: A polar rectangle. Right: An annulus.

Activity 12.5.3 Consider a polar rectangle R, with r between r_i and r_{i+1} and θ between θ_j and θ_{j+1} as shown at left in Figure 12.5.2. Let $\Delta r = r_{i+1} - r_i$ and $\Delta\theta = \theta_{j+1} - \theta_j$. Let ΔA be the area of this region.

a. Explain why the area ΔA in polar coordinates is not $\Delta r\, \Delta\theta$.

b. Now find ΔA by the following steps:

   i. Find the area of the annulus (the washer-like region) between r_i and r_{i+1}, as shown at right in Figure 12.5.2. This area will be in terms of r_i and r_{i+1}.

   ii. Observe that the region R is only a portion of the annulus, so the area ΔA of R is only a fraction of the area of the annulus. For instance, if $\theta_{i+1} - \theta_i$ were $\frac{\pi}{4}$, then the resulting wedge would be

   $$\frac{\frac{\pi}{4}}{2\pi} = \frac{1}{8}$$

   of the entire annulus. In this more general context, using the wedge between the two noted angles, what fraction of the area of the annulus is the area ΔA?

   iii. Write an expression for ΔA in terms of r_i, r_{i+1}, θ_j, and θ_{j+1}.

   iv. Finally, write the area ΔA in terms of r_i, r_{i+1}, Δr, and $\Delta\theta$, where each quantity appears only once in the expression. (Hint: Think about how to factor a difference of squares.)

c. As we take the limit as Δr and $\Delta\theta$ go to 0, Δr becomes dr, $\Delta\theta$ becomes $d\theta$, and ΔA becomes dA, the area element. Using your work in (iv), write dA in terms of r, dr, and $d\theta$.

From the result of Activity 12.5.3, we see when we convert an integral from rectangular coordinates to polar coordinates, we must not only convert x and y to being in terms of r and θ, but we also have to change the area element to $dA = r\,dr\,d\theta$ in polar coordinates. As we saw in Activity 12.5.3, the reason the additional factor of r in the polar area element is due to the fact that in polar coordinates, the cross sectional area element increases as r increases, while the cross sectional area element in rectangular coordinates is constant. So, given a double integral $\iint_D f(x,y)\,dA$ in rectangular coordinates, to write a corresponding iterated integral in polar coordinates, we replace x with $r\cos(\theta)$, y with $r\sin(\theta)$ and dA with $r\,dr\,d\theta$. Of course, we need to describe the region

D in polar coordinates as well. To summarize:

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 $\iint_D f(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta$.

Example 12.5.3 Let $f(x,y) = e^{x^2+y^2}$ on the disk $D = \{(x,y) : x^2+y^2 \leq 1\}$. We will evaluate $\iint_D f(x,y)\,dA$.

In rectangular coordinates the double integral $\iint_D f(x,y)\,dA$ can be written as the iterated integral

$$\iint_D f(x,y)\,dA = \int_{x=-1}^{x=1} \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} e^{x^2+y^2}\,dy\,dx.$$

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 constraints $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$. 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:

$$\begin{aligned}
\int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1} e^{r^2}\,r\,dr\,d\theta &= \int_{\theta=0}^{2\pi} \left(\frac{1}{2} e^{r^2} \Big|_{r=0}^{r=1} \right) d\theta \\
&= \frac{1}{2} \int_{\theta=0}^{\theta=2\pi} (e-1)\,d\theta \\
&= \frac{1}{2}(e-1) \int_{\theta=0}^{\theta=2\pi} d\theta \\
&= \frac{1}{2}(e-1)\,[\theta] \Big|_{\theta=0}^{\theta=2\pi} \\
&= \pi(e-1).
\end{aligned}$$

□

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 12.5.4 Let $f(x,y) = x+y$ and $D = \{(x,y) : x^2+y^2 \leq 4\}$.

a. Sketch the region D and then write the double integral of f over D as an iterated integral in rectangular coordinates.

b. Write the double integral of f over D as an iterated integral in polar coordinates.

c. Evaluate one of the iterated integrals. Why is the final value you found not surprising?

Activity 12.5.5 Consider the circle given by $x^2 + (y-1)^2 = 1$ as shown in Figure 12.5.4.

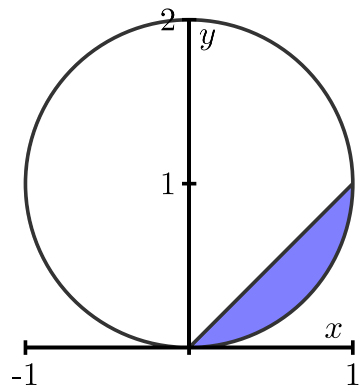


Figure 12.5.4 The graphs of $y = x$ and $x^2 + (y-1)^2 = 1$, for use in Activity 12.5.5.

a. Determine a polar curve in the form $r = f(\theta)$ that traces out the circle $x^2 + (y-1)^2 = 1$. (Hint: Recall that a circle centered at the origin of radius r can be described by the equations $x = r\cos(\theta)$ and $y = r\sin(\theta)$.)

b. Find the exact average value of $g(x, y) = \sqrt{x^2 + y^2}$ over the interior of the circle $x^2 + (y-1)^2 = 1$.

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 (this is the shaded region in Figure 12.5.4).

d. Explain why in both (b) and (c) it is advantageous to use polar coordinates.

12.5.3 Summary

- The polar representation of a point P is the ordered pair (r, θ) where r is the distance from the origin to P and θ is the angle the ray through the origin and P makes with the positive x-axis.

- The polar coordinates r and θ of a point (x, y) in rectangular coordinates satisfy
$$r = \sqrt{x^2 + y^2} \qquad \text{and} \qquad \tan(\theta) = \frac{y}{x};$$
the rectangular coordinates x and y of a point (r, θ) in polar coordinates satisfy
$$x = r\cos(\theta) \qquad \text{and} \qquad y = r\sin(\theta).$$

- 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

$$\iint_D f(r\cos(\theta), r\sin(\theta))\, r\, dr\, d\theta.$$

12.5.4 Exercises

1. For each set of Polar coordinates, match the equivalent Cartesian coordinates.

2. (a) The Cartesian coordinates of a point are $(-1, -\sqrt{3})$.
(i) Find polar coordinates (r, θ) of the point, where $r > 0$ and $0 \le \theta < 2\pi$.
$r =$ ______________________________
$\theta =$ ______________________________
(ii) Find polar coordinates (r, θ) of the point, where $r < 0$ and $0 \le \theta < 2\pi$.
$r =$ ______________________________
$\theta =$ ______________________________
(b) The Cartesian coordinates of a point are $(-2, 3)$.
(i) Find polar coordinates (r, θ) of the point, where $r > 0$ and $0 \le \theta < 2\pi$.
$r =$ ______________________________
$\theta =$ ______________________________
(ii) Find polar coordinates (r, θ) of the point, where $r < 0$ and $0 \le \theta < 2\pi$.
$r =$ ______________________________
$\theta =$ ______________________________

3. (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 $2\pi \le \theta < 4\pi$.
$r =$ _______________
$\theta =$ _______________
(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $0 \le \theta < 2\pi$.
$r =$ _______________
$\theta =$ _______________
(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 $2\pi \le \theta < 4\pi$.
$r =$ _______________
$\theta =$ _______________
(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $-2\pi \le \theta < 0$.
$r =$ _______________
$\theta =$ _______________
(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 $2\pi \le \theta < 4\pi$.
$r =$ _______________
$\theta =$ _______________
(ii) Find another pair of polar coordinates for this point such that $r < 0$ and $0 \le \theta < 2\pi$.

$r =$ ________________
$\theta =$ ________________

4. 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.) $(5, \frac{\pi}{3}), (-5, \frac{-\pi}{3})$
_ b.) $(2, \frac{27\pi}{4}), (2, -\frac{27\pi}{4})$
_ c.) $(0, 5\pi), (0, \frac{3\pi}{4})$
_ d.) $(1, \frac{109\pi}{4}), (-1, \frac{\pi}{4})$
_ e.) $(16, \frac{50\pi}{3}), (-16, \frac{-\pi}{3})$
_ f.)$(5, 8\pi), (-5, 8\pi)$

5. A curve with polar equation

$$r = \frac{5}{7\sin\theta + 50\cos\theta}$$

represents a line. Write this line in the given Cartesian form.
$y =$ ________________
Note: Your answer should be a function of x .

6. 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.
$r =$ ________________________________

7. By changing to polar coordinates, evaluate the integral

$$\iint_D (x^2 + y^2)^{3/2}\, dxdy$$

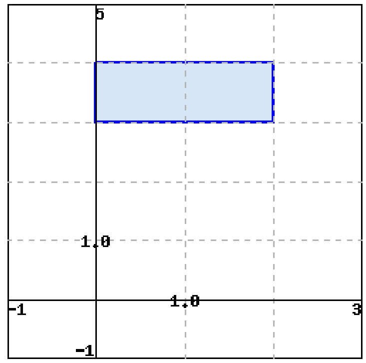
where D is the disk $x^2 + y^2 \leq 36$.
Answer = ________________________________

8. Convert the integral

$$\int_0^{\sqrt{10}} \int_{-x}^{x} dy\, dx$$

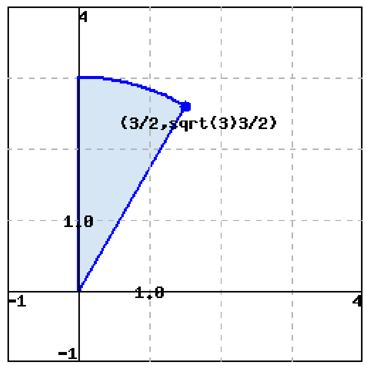
to polar coordinates and evaluate it (use t for θ):
With $a =$ _________, $b =$ _________, $c =$ _________ and $d =$ _________,
$\int_0^{\sqrt{10}} \int_{-x}^{x} dy\, dx = \int_a^b \int_c^d$ _____$dr\, dt$
$= \int_a^b$ __________________________ dt
$=$ __________________________ $\Big|_a^b$
$=$ __________________________.

9. 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_____

10. A Cartesian equation for the polar equation $r = 3$ can be written as:
$x^2 + y^2 =$ __________

11. 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 $x^2-6x+y^2 = 0$.

12. ***(a)*** Graph $r = 1/(4\cos\theta)$ for $-\pi/2 \le \theta \le \pi/2$ 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 = __________________________________

13. Using polar coordinates, evaluate the integral $\displaystyle\iint_R \sin(x^2+y^2)\,dA$ where R is the region $4 \le x^2 + y^2 \le 25$.

14. Sketch the region of integration for the following integral.

$$\int_0^{\pi/4} \int_0^{5/\cos(\theta)} f(r,\theta)\, r\, dr\, d\theta$$

The region of integration is bounded by

- ⊙ $y = 0, y = \sqrt{25 - x^2}$, and $x = 5$
- ⊙ $y = 0, y = x$, and $y = 5$
- ⊙ $y = 0, y = x$, and $x = 5$
- ⊙ $y = 0, x = \sqrt{25 - y^2}$, and $y = 5$
- ⊙ None of the above

15. Use the polar coordinates to find the volume of a sphere of radius 7.

16. Consider the solid under the graph of $z = e^{-x^2-y^2}$ above the disk $x^2+y^2 \leq a^2$, 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 $a \to \infty$?

$\lim\limits_{a\to\infty} V =$ ______________________

17. 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).

18. 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.

19. 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. $\displaystyle\int_{0}^{\sqrt{2}/2}\int_{y}^{\sqrt{1-y^2}} \cos(x^2+y^2)\,dx\,dy.$

12.6 Surfaces Defined Parametrically and Surface Area

Motivating Questions

- What is a parameterization of a surface?
- How do we find the surface area of a parametrically defined surface?

We have now studied at length how curves in space can be defined parametrically by functions of the form $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$, and surfaces can be represented by functions $z = f(x, y)$. In what follows, we will see how we can also define surfaces parametrically. A one-dimensional curve in space results from a vector function that relies upon one parameter, so a two-dimensional surface naturally involves the use of two parameters. If $x = x(s, t)$, $y = y(s, t)$, and $z = z(s, t)$ are functions of independent parameters s and t, then the terminal points of all vectors of the form

$$\vec{r}(s, t) = x(s, t)\hat{\imath} + y(s, t)\hat{\jmath} + z(s, t)\hat{k}$$

form a surface in space. The equations $x = x(s, t)$, $y = y(s, t)$, and $z = z(s, t)$ are the *parametric equations* for the surface, or a *parametrization* of the surface. In Preview Activity 12.6.1 we investigate how to parameterize a cylinder and a cone.

Preview Activity 12.6.1 Recall the standard parameterization of the unit circle that is given by

$$x(t) = \cos(t) \qquad \text{and} \qquad y(t) = \sin(t),$$

where $0 \le t \le 2\pi$.

a. Determine a parameterization of the circle of radius 1 in $\mathbb{R}^3$ that has its center at $(0, 0, 1)$ and lies in the plane $z = 1$.

b. Determine a parameterization of the circle of radius 1 in 3-space that has its center at $(0, 0, -1)$ and lies in the plane $z = -1$.

c. Determine a parameterization of the circle of radius 1 in 3-space that has its center at $(0, 0, 5)$ and lies in the plane $z = 5$.

d. Taking into account your responses in (a), (b), and (c), describe the graph that results from the set of parametric equations

$$x(s,t) = \cos(t), \quad y(s,t) = \sin(t), \quad \text{and} \quad z(s,t) = s,$$

where $0 \le t \le 2\pi$ and $-5 \le s \le 5$. Explain your thinking.

e. Just as a cylinder can be viewed as a "stack" of circles of constant radius, a cone can be viewed as a stack of circles with varying radius. Modify the parametrizations of the circles above in order to construct the parameterization of a cone whose vertex lies at the origin, whose base radius is 4, and whose height is 3, where the base of the cone lies in the plane $z = 3$. Use appropriate technology to plot the parametric equations you develop. (Hint: The cross sections parallel to the xy-plane are circles, with the radii varying linearly as z increases.)

12.6.1 Parametric Surfaces

In a single-variable setting, any function may have its graph expressed parametrically. For instance, given $y = g(x)$, by considering the parameterization $\langle t, g(t)\rangle$ (where t belongs to the domain of g), we generate the same curve. What is more important is that certain curves that are not functions may be represented parametrically; for instance, the circle (which cannot be represented by a single function) can be parameterized by $\langle \cos(t), \sin(t)\rangle$, where $0 \le t \le 2\pi$.

In the same way, in a two-variable setting, the surface $z = f(x,y)$ may be expressed parametrically by considering

$$\langle x(s,t), y(s,t), z(s,t)\rangle = \langle s, t, f(s,t)\rangle,$$

where (s,t) varies over the entire domain of f. Therefore, any familiar surface that we have studied so far can be generated as a parametric surface. But what is more powerful is that there are surfaces that cannot be generated by a single function $z = f(x,y)$ (such as the unit sphere), but that can be represented parametrically. We now consider an important example.

Example 12.6.1 Consider the torus (or doughnut) shown in Figure 12.6.2.

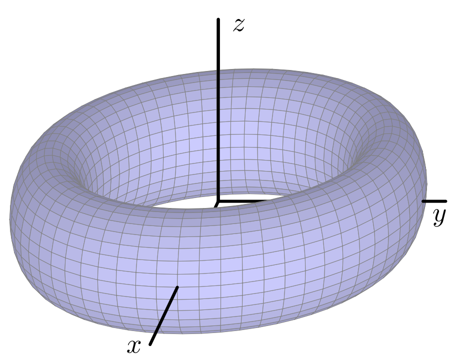


Figure 12.6.2 A torus.

To find a parametrization of this torus, we recall our work in Preview Activity 12.6.1. There, we saw that a circle of radius r that has its center at the point $(0,0,z_0)$ and is contained in the horizontal plane $z = z_0$, as shown in Figure 12.6.3, can be parametrized using the vector-valued function $\vec{r}$ defined by

$$\vec{r}(t) = r\cos(t)\hat{\imath} + r\sin(t)\hat{\jmath} + z_0\hat{k}$$

where $0 \le t \le 2\pi$.

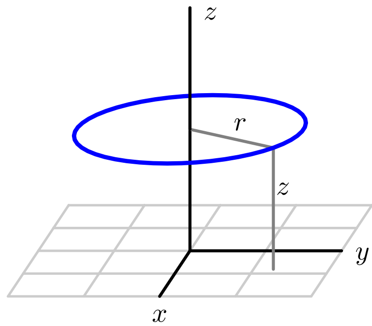


Figure 12.6.3 A circle in a horizontal plane centered at $(0, 0, z_0)$.

To obtain the torus in Figure 12.6.2, we begin with a circle of radius a in the xz-plane centered at $(b, 0)$, as shown on the left of Figure 12.6.4. We may parametrize the points on this circle, using the parameter s, by using the equations

$$x(s) = b + a\cos(s) \text{ and } z(s) = a\sin(s),$$

where $0 \le s \le 2\pi$.

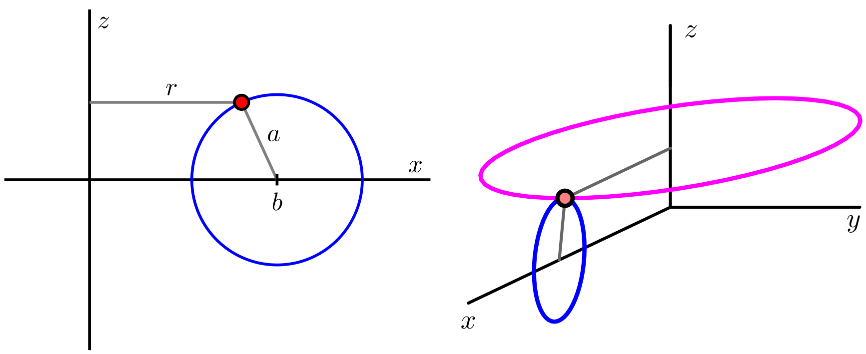


Figure 12.6.4 Revolving a circle to obtain a torus.

Let's focus our attention on one point on this circle, such as the indicated point, which has coordinates $(x(s), 0, z(s))$ for a fixed value of the parameter s. When this point is revolved about the z-axis, we obtain a circle contained in a horizontal plane centered at $(0, 0, z(s))$ and having radius $x(s)$, as shown on the right of Figure 12.6.4. If we let t be the new parameter that generates the circle for the rotation about the z-axis, this circle may be parametrized by

$$\vec{r}(s,t) = x(s)\cos(t)\hat{\imath} + x(s)\sin(t)\hat{\jmath} + z(s)\hat{k}.$$

Now using our earlier parametric equations for $x(s)$ and $z(s)$ for the original smaller circle, we have an overall parameterization of the torus given by

$$\vec{r}(s,t) = (b + a\cos(s))\cos(t)\hat{\imath} + (b + a\cos(s))\sin(t)\hat{\jmath} + a\sin(s)\hat{k}.$$

To trace out the entire torus, we require that the parameters vary through the values $0 \le s \le 2\pi$ and $0 \le t \le 2\pi$. □

Activity 12.6.2 In this activity, we seek a parametrization of the sphere of radius R centered at the origin, as shown on the left in Figure 12.6.5. Notice that this sphere may be obtained by revolving a half-circle contained in the xz-plane about the z-axis, as shown on the right.

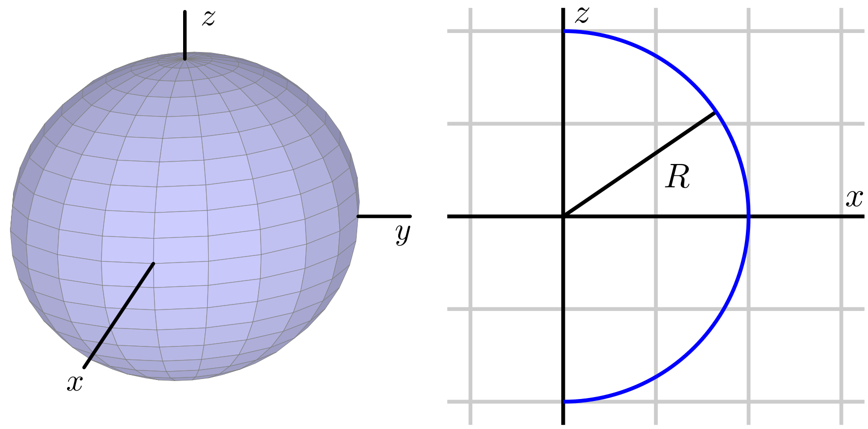


Figure 12.6.5 A sphere obtained by revolving a half-circle.

a. Begin by writing a parametrization of this half-circle using the parameter s:

$$x(s) = \ldots \qquad z(s) = \ldots.$$

Be sure to state the domain of the parameter s.

b. By revolving the points on this half-circle about the z-axis, obtain a parametrization $\vec{r}(s,t)$ of the points on the sphere of radius R. Be sure to include the domain of both parameters s and t. (Hint: What is the radius of the circle obtained when revolving a point on the half-circle around the z axis?)

c. Draw the surface defined by your parameterization with appropriate technology.

12.6.2 The Surface Area of Parametrically Defined Surfaces

Recall that a differentiable function is locally linear — that is, if we zoom in on the surface around a point, the surface looks like its tangent plane. We now exploit this idea in order to determine the surface area generated by a parametrization $\langle x(s,t), y(s,t), z(s,t)\rangle$. The basic idea is a familiar one: we will subdivide the surface into small pieces, in the approximate shape of small parallelograms, and thus estimate the entire the surface area by adding the areas of these approximation parallelograms. Ultimately, we use an integral to sum these approximations and determine the exact surface area.

Let

$$\vec{r}(s,t) = x(s,t)\hat{\imath} + y(s,t)\hat{\jmath} + z(s,t)\hat{k}$$

define a surface over a rectangular domain $a \le s \le b$ and $c \le t \le d$. As a function of two variables, s and t, it is natural to consider the two partial derivatives of the vector-valued function $\vec{r}$, which we define by

$$\vec{r}_s(s,t) = x_s(s,t)\hat{\imath} + y_s(s,t)\hat{\jmath} + z_s(s,t)\hat{k}$$
$$\vec{r}_t(s,t) = x_t(s,t)\hat{\imath} + y_t(s,t)\hat{\jmath} + z_t(s,t)\hat{k}.$$

In the usual way, we slice the domain into small rectangles. In particular, we partition the interval $[a, b]$ into m subintervals of length $\Delta s = \frac{b-a}{n}$ and let $s_0, s_1, \ldots, s_m$ be the endpoints of these subintervals, where $a = s_0 < s_1 < s_2 < \cdots < s_m = b$. Also partition the interval $[c, d]$ into n subintervals of equal length $\Delta t = \frac{d-c}{n}$ and let $t_0, t_1, \ldots, t_n$ be the endpoints of these subintervals, where $c = t_0 < t_1 < t_2 < \cdots < t_n = d$. These two partitions create a partition of the rectangle $R = [a, b] \times [c, d]$ 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 $\Delta A = \Delta s \cdot \Delta t$.

Now we want to think about the small piece of area on the surface itself that lies above one of these small rectangles in the domain. Observe that if we increase s by a small amount Δs from the point (s_{i-1}, t_{j-1}) in the domain, then $\vec{r}$ changes by approximately $\vec{r}_s(s_{i-1}, t_{j-1})\Delta s$. Similarly, if we increase t by a small amount Δt from the point (s_{i-1}, t_{j-1}), then $\vec{r}$ changes by approximately $\vec{r}_t(s_{i-1}, t_{j-1})\Delta t$. So we can approximate the surface defined by $\vec{r}$ on the st-rectangle $[s_{i-1}, s_i] \times [t_{j-1}, t_j]$ with the parallelogram determined by the vectors $\vec{r}_s(s_{i-1}, t_{j-1})\Delta s$ and $\vec{r}_t(s_{i-1}, t_{j-1})\Delta t$, as seen in Figure 12.6.6.

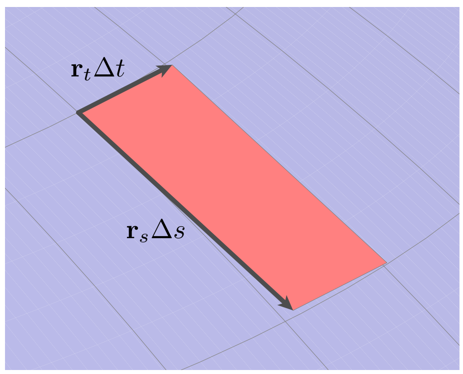


Figure 12.6.6 Approximation surface area with a parallelogram.

Say that the small parallelogram has area S_{ij}. If we can find its area, then all that remains is to sum the areas of all of the generated parallelograms and take a limit. Recall from our earlier work in the course that given two vectors $\vec{u}$ and $\vec{v}$, the area of the parallelogram spanned by $\vec{u}$ and $\vec{v}$ is given by the magnitude of their cross product, $|\vec{u} \times \vec{v}|$. In the present context, it follows that the area, S_{ij}, of the parallelogram determined by the vectors $\vec{r}_s(s_{i-1}, t_{j-1})\Delta s$ and $\vec{r}_t(s_{i-1}, t_{j-1})\Delta t$ is

$$\begin{aligned} S_{ij} &= |(\vec{r}_s(s_{i-1}, t_{j-1})\Delta s) \times (\vec{r}_t(s_{i-1}, t_{j-1})\Delta t)| \\ &= |\vec{r}_s(s_{i-1}, t_{j-1}) \times \vec{r}_t(s_{i-1}, t_{j-1})|\Delta s \Delta t, \end{aligned} \tag{12.6.1}$$

where the latter equality holds from standard properties of the cross product and length.

We sum the surface area approximations from Equation (12.6.1) over all sub-rectangles to obtain an estimate for the total surface area, S, given by

$$S \approx \sum_{i=1}^{m} \sum_{j=1}^{n} |\vec{r}_s(s_{i-1}, t_{j-1}) \times \vec{r}_t(s_{i-1}, t_{j-1})|\Delta s \Delta t.$$

Taking the limit as $m, n \to \infty$ shows that the surface area of the surface defined by $\vec{r}$ over the domain D is given as follows.

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

$$S = \iint_D |\vec{r}_s \times \vec{r}_t| \, dA. \tag{12.6.2}$$

Activity 12.6.3 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 12.6.7.

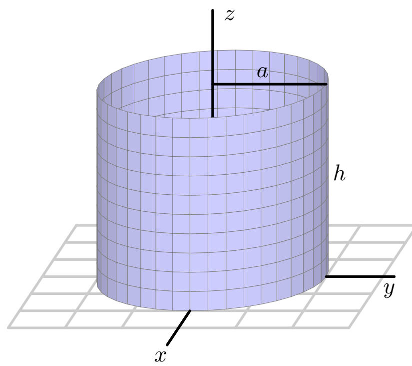


Figure 12.6.7 A cylinder.

a. Set up an iterated integral to determine the surface area of this cylinder.

b. Evaluate the iterated integral.

c. Recall that 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 work in (b).

As we noted earlier, we can take any surface $z = f(x,y)$ and generate a corresponding parameterization for the surface by writing $\langle s, t, f(s,t)\rangle$. Hence, we can use our recent work with parametrically defined surfaces to find the surface area that is generated by a function $f = f(x,y)$ over a given domain.

Activity 12.6.4 Let $z = f(x,y)$ define a smooth surface, and consider the corresponding 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 (12.6.2), show that the area, S, of the surface defined by the graph of f over D is

$$S = \iint_D \sqrt{(f_x(x,y))^2 + (f_y(x,y))^2 + 1} \, dA.$$

b. Use the formula developed in (a) 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 in (b) 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 (b).

12.6.3 Summary

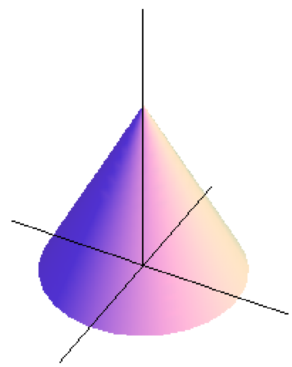
- A parameterization of a curve describes the coordinates of a point on the curve in terms of a single parameter t, while a parameterization of a surface describes the coordinates of points on the surface in terms of two independent parameters.

- 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

$$S = \iint_D |\vec{r}_s \times \vec{r}_t| \, dA.$$

12.6.4 Exercises

1. Consider the cone shown below.



If the height of the cone is 7 and the base radius is 9, 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
__________ $\leq s \leq$ __________ and
__________ $\leq t \leq$ __________.

2. Parameterize the plane through the point $(-5,-2,1)$ with the normal vector $\langle 2,-3,-5\rangle$
$\vec{r}(s,t) =$ ____
(Use s and t for the parameters in your parameterization, and ***enter your vector as a single vector, with angle brackets****: e.g., as* ***\lt 1 + s + t, s - t, 3 - t \gt****.)*

3. Parameterize a vase formed by rotating the curve $z = 10\sqrt{x-1}$, $1 \leq x \leq 6$, around the z-axis. Use s and t for your parameters.
$x(s,t) =$ ____,
$y(s,t) =$ ____, and
$z(s,t) =$ ____, with
__________ $\leq s \leq$ __________ and
__________ $\leq t \leq$ __________

4. Find parametric equations for the sphere centered at the origin and with radius 4. Use the parameters s and t in your answer.
$x(s,t) =$ ____,
$y(s,t) =$ ____, and
$z(s,t) =$ ____, where

__________ $\leq s \leq$ __________ and
__________ $\leq t \leq$ __________.

5. Find the surface area of that part of the plane $5x + 4y + z = 9$ that lies inside the elliptic cylinder $\frac{x^2}{64} + \frac{y^2}{25} = 1$
Surface Area = ____

6. Find the surface area of the part of the circular paraboloid $z = x^2 + y^2$ that lies inside the cylinder $x^2 + y^2 = 9$.

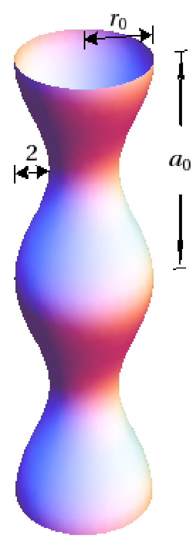
7. Find the surface area of the part of the plane $3x + 3y + z = 3$ that lies inside the cylinder $x^2 + y^2 = 1$.

8. Write down the iterated integral which expresses the surface area of $z = y^6 \cos^3 x$ over the triangle with vertices (-1,1), (1,1), (0,2):

$$\int_a^b \int_{f(y)}^{g(y)} \sqrt{h(x,y)}\, dx dy$$

$a =$ ____________________
$b =$ ____________________
$f(y) =$ ____________________
$g(y) =$ ____________________
$h(x, y) =$ ____

9. 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, $r_0 = 6$ inches and $a_0 = 12$ 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 units[1].)

10. Consider the ellipsoid given by the equation

$$\frac{x^2}{16} + \frac{y^2}{25} + \frac{z^2}{9} = 1.$$

In Activity 12.6.2, we found that a parameterization of the sphere S of radius R centered at the origin is

$$x(r,s) = R\cos(s)\cos(t), \;\; y(s,t) = R\cos(s)\sin(t), \quad \text{and} \quad z(s,t) = R\sin(s)$$

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.

11. 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 12.6.2).

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.

12. 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 (12.6.2) 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?

[1] `/webwork2_files/helpFiles/Units.html`

13. 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 12.6.3.)

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.

12.7 Triple Integrals

Motivating Questions

- How are a triple Riemann sum and the corresponding triple integral of a continuous function $f = f(x, y, z)$ defined?
- What are two things the triple integral of a function can tell us?

We have now learned that we define 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, and that these ideas parallel the single-variable integral of a function $g = g(x)$ on an interval $[a, b]$. Moreover, this double integral has natural interpretations and applications, and can even be considered over non-rectangular regions, D. For instance, given a continuous function f over a region D, the average value of f, $f_{\text{AVG}(D)}$, is given by

$$f_{\text{AVG}(D)} = \frac{1}{A(D)} \iint_D f(x, y)\, dA,$$

where $A(D)$ is the area of D. Likewise, if $\delta(x, y)$ describes a mass density function on a lamina over D, the mass, M, of the lamina is given by

$$M = \iint_D \delta(x, y)\, dA.$$

It is natural to wonder if it is possible to extend these ideas of double Riemann sums and double integrals for functions of two variables to triple Riemann sums and then triple integrals for functions of three variables. We begin investigating in Preview Activity 12.7.1.

Preview Activity 12.7.1 Consider a solid piece of granite in the shape of a box $B = \{(x, y, z) : 0 \le x \le 4, 0 \le y \le 6, 0 \le z \le 8\}$, whose density varies from point to point. Let $\delta(x, y, z)$ represent the mass density of the piece of granite at point (x, y, z) in kilograms per cubic meter (so we are measuring x, y, and z in meters). Our goal is to find the mass of this solid.

Recall that if the density was constant, we could find the mass by multiplying the density and volume; since the density varies from point to point, we will use the approach we did with two-variable lamina problems, and slice the solid into small pieces on which the density is roughly constant.

Partition the interval $[0, 4]$ into 2 subintervals of equal length, the interval $[0, 6]$ into 3 subintervals of equal length, and the interval $[0, 8]$ into 2 subintervals of equal length. This partitions the box B into sub-boxes as shown in Figure 12.7.1.

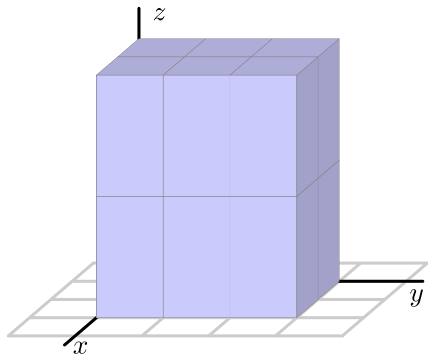


Figure 12.7.1 A partitioned three-dimensional domain.

a. Let $0 = x_0 < x_1 < x_2 = 4$ be the endpoints of the subintervals of $[0,4]$ after partitioning. Draw a picture of Figure 12.7.1 and label these endpoints on your drawing. Do likewise with $0 = y_0 < y_1 < y_2 < y_3 = 6$ and $0 = z_0 < z_1 < z_2 = 8$ What is the length Δx of each subinterval $[x_{i-1}, x_i]$ for i from 1 to 2? the length of Δy? of Δz?

b. The partitions of the intervals $[0,4]$, $[0,6]$ and $[0,8]$ partition the box B into sub-boxes. How many sub-boxes are there? What is volume ΔV of each sub-box?

c. Let B_{ijk} denote the sub-box $[x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$. Say that we choose a point $(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})$ in the i, j, kth sub-box for each possible combination of i, j, k. What is the meaning of $\delta(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})$? What physical quantity will $\delta(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})\Delta V$ approximate?

d. What final step(s) would it take to determine the exact mass of the piece of granite?

12.7.1 Triple Riemann Sums and Triple Integrals

Through the application of a mass density distribution over a three-dimensional solid, Preview Activity 12.7.1 suggests that the generalization from double Riemann sums of functions of two variables to triple Riemann sums of functions of three variables is natural. In the same way, so is the generalization 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 12.7.2 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 interval $[a, b]$ into m subintervals of equal length $\Delta x = \frac{b-a}{m}$. Let $x_0, x_1, \ldots, x_m$ be the endpoints of these subintervals, where $a = x_0 < x_1 < x_2 < \cdots < x_m = b$. Do likewise with the interval $[c, d]$ using n subintervals of equal length $\Delta y = \frac{d-c}{n}$ to generate $c = y_0 < y_1 < y_2 < \cdots < y_n = d$, and with the interval $[r, s]$ using ℓ subintervals of equal length $\Delta z = \frac{s-r}{\ell}$ to have $r = z_0 < z_1 < z_2 < \cdots < z_l = s$.
- Let B_{ijk} be the sub-box of B with opposite vertices $(x_{i-1}, y_{j-1}, z_{k-1})$ and (x_i, y_j, z_k) for i between 1 and m, j between 1 and n, and k between 1 and ℓ. The volume of each B_{ijk} is $\Delta V = \Delta x \cdot \Delta y \cdot \Delta z$.

- Let $(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*)$ be a point in box B_{ijk} for each i, j, and k. The resulting triple Riemann sum for f on B is

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{\ell} f(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*) \cdot \Delta V.$$

◊

If $f(x, y, z)$ represents the mass density of the box B, then, as we saw in Preview Activity 12.7.1, the triple Riemann sum approximates the total mass of the box B. In order to find the exact mass of the box, we need to let the number of sub-boxes increase without bound (in other words, let m, n, and ℓ go to infinity); in this case, the finite sum of the mass approximations becomes the actual mass of the solid B. More generally, we have the following definition of the triple integral.

Definition 12.7.3 With following notation defined as in a triple Riemann sum, the **triple integral of f over B** is

$$\iiint_B f(x, y, z)\, dV = \lim_{m,n,\ell \to \infty} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{\ell} f(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*) \cdot \Delta V.$$

◊

As we noted 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 the following important boxed information. Note that each of these quantities may actually be considered over a general domain S in $\mathbb{R}^3$, not simply a box, B.

- The triple integral
$$V(S) = \iiint_S 1\, dV$$
represents the *volume* of the solid S.

- The *average value* of the function $f = f(x, y, x)$ over a solid domain 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 *center of mass* of the solid S with density $\delta = \delta(x, y, z)$ is $(\overline{x}, \overline{y}, \overline{z})$, where
$$\overline{x} = \frac{\iiint_S x\ \delta(x, y, z)\, dV}{M},$$
$$\overline{y} = \frac{\iiint_S y\ \delta(x, y, z)\, dV}{M},$$
$$\overline{z} = \frac{\iiint_S z\ \delta(x, y, z)\, dV}{M},$$

and $M = \iiint_S \delta(x, y, z)\, dV$ is the mass of the solid S.

In the Cartesian coordinate system, the volume element dV is $dz\,dy\,dx$, and, as a consequence, a triple integral of a function f over a box $B = [a, b] \times [c, d] \times [r, s]$ in Cartesian coordinates can be evaluated as an iterated integral of the form

$$\iiint_B f(x, y, z)\, dV = \int_a^b \int_c^d \int_r^s f(x, y, z)\, dz\, dy\, dx.$$

If we want to evaluate a triple integral as an iterated integral over a solid S that is not a box, then we need to describe the solid in terms of variable limits.

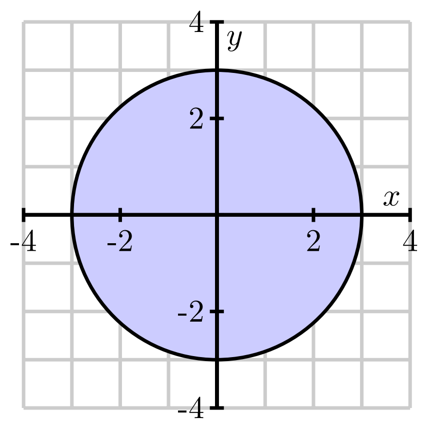
Activity 12.7.2

a. Set up and evaluate the triple integral of $f(x, y, z) = x - y + 2z$ over the box $B = [-2, 3] \times [1, 4] \times [0, 2]$.

b. Let S be the solid cone bounded by $z = \sqrt{x^2 + y^2}$ and $z = 3$. A picture of S is shown at right in Figure 12.7.4. Our goal in what follows is to set up an iterated integral of the form

$$\int_{x=?}^{x=?} \int_{y=?}^{y=?} \int_{z=?}^{z=?} \delta(x, y, z)\, dz\, dy\, dx \tag{12.7.1}$$

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). Our particular task is to find the limits on each of the three integrals.



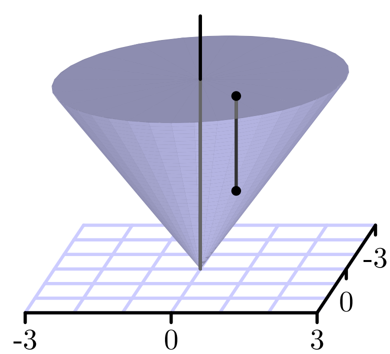


Figure 12.7.4 Left: The cone. Right: Its projection.

i. If we think about slicing up the solid, we can consider slicing the domain of the solid's projection onto the xy-plane (just as we would slice a two-dimensional region in $\mathbb{R}^2$), and then slice in the z-direction as well. The projection of the solid onto the xy-plane is shown at left in Figure 12.7.4. If we decide to first slice the domain of the solid's projection perpendicular to the x-axis, over what range of constant x-values would we have to slice?

ii. If we continue with slicing the domain, what are the limits on y on a typical slice? How do these depend on x? What, therefore, are the limits on the middle integral?

iii. Finally, now that we have thought about slicing up the two-dimensional domain that is the projection of the cone, what are the limits on z in the innermost integral? Note that over any point (x, y) in the plane, a vertical slice in the z direction will involve a range of values from the cone itself to its flat top. In particular, observe that at least one of these limits is not constant but depends on x and y.

iv. In conclusion, write an iterated integral of the form (12.7.1) that represents the mass of the cone S.

Note well: When setting up iterated integrals, the limits on a given variable can be *only* in terms of the remaining variables. 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

- $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\,dz\,dy\,dx$
- $\displaystyle\int_r^s \int_{p_1(z)}^{p_2(z)} \int_{q_1(x,z)}^{q_2(x,z)} f(x,y,z)\,dy\,dx\,dz$

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 (theoretically) be arranged in any order. Of course, 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.

Example 12.7.5 Find the mass of the tetrahedron 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 at left in Figure 12.7.6.

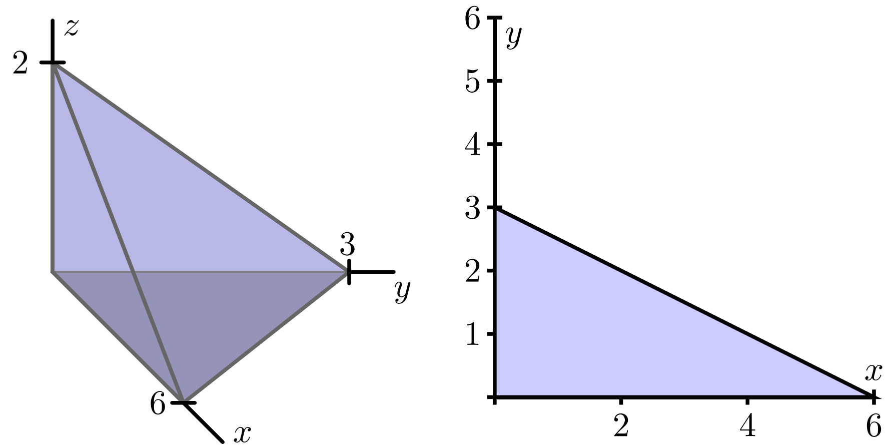


Figure 12.7.6 Left: The tetrahedron. Right: Its projection.

We find the mass, M, of the tetrahedron by the triple integral

$$M = \iiint_S \delta(x, y, z)\,dV,$$

where S is the solid tetrahedron described above. In this example, we choose to integrate with respect to z first for the innermost integral. The top of the tetrahedron is given by the equation

$$x + 2y + 3z = 6;$$

solving for z then yields

$$z = \frac{1}{3}(6 - x - 2y).$$

The bottom of the tetrahedron is the xy-plane, so the limits on z in the iterated integral will be $0 \leq z \leq \frac{1}{3}(6 - x - 2y)$.

To find the bounds on x and y we project the tetrahedron onto the xy-plane; this corresponds to setting $z = 0$ in the equation $z = \frac{1}{3}(6 - x - 2y)$. The resulting relation between x and y is

$$x + 2y = 6.$$

The right image in Figure 12.7.6 shows the projection of the tetrahedron onto the xy-plane.

If we choose to integrate with respect to y for the middle integral in the iterated integral, then the lower limit on y is the x-axis and the upper limit is the hypotenuse of the triangle. Note that the hypotenuse joins the points $(6, 0)$ and $(0, 3)$ and so has equation $y = 3 - \frac{1}{2}x$. Thus, the bounds on y are $0 \leq y \leq 3 - \frac{1}{2}x$. Finally, the x values run from 0 to 6, so the iterated integral that gives the mass of the tetrahedron is

$$M = \int_0^6 \int_0^{3-(1/2)x} \int_0^{(1/3)(6-x-2y)} x + y + z \, dz \, dy \, dx. \tag{12.7.2}$$

Evaluating the triple integral gives us

$$\begin{aligned}
M &= \int_0^6 \int_0^{3-(1/2)x} \int_0^{(1/3)(6-x-2y)} x + y + z \, dz \, dy \, dx \\
&= \int_0^6 \int_0^{3-(1/2)x} \left[xz + yz + \frac{z}{2} \right] \Bigg|_0^{(1/3)(6-x-2y)} dy \, dx \\
&= \int_0^6 \int_0^{3-(1/2)x} \frac{4}{3}x - \frac{5}{18}x^2 - \frac{}{9}xy + \frac{2}{3}y - \frac{4}{9}y^2 + 2 \, dy \, dx \\
&= \int_0^6 \left[\frac{4}{3}xy - \frac{5}{18}x^2 y - \frac{7}{18}xy^2 + \frac{1}{3}y^2 - \frac{4}{27}y^3 + 2y \right] \Bigg|_0^{3-(1/2)x} dx \\
&= \int_0^6 5 + \frac{1}{2}x - \frac{7}{12}x^2 + \frac{13}{216}x^3 \, dx \\
&= \left[5x + \frac{1}{4}x^2 - \frac{7}{36}x^3 + \frac{13}{864}x^4 \right] \Bigg|_0^6 \\
&= \frac{33}{2}.
\end{aligned}$$

□

Setting up limits on iterated integrals can require considerable geometric intuition. It is important to not only create carefully labeled figures, but also to think about how we wish to slice the solid. 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 the preceding example.

Activity 12.7.3 There are several other ways we could have set up the integral to give the mass of the tetrahedron in Example 12.7.5.

a. How many different iterated integrals could be set up that are equal to the integral in Equation (12.7.2)?

b. Set up an iterated integral, integrating first with respect to z, then x, then y that is equivalent to the integral in Equation (12.7.2). Before you write down the integral, think about Figure 12.7.6, and draw an appropriate two-dimensional image of an important projection.

c. Set up an iterated integral, integrating first with respect to y, then z, then x that is equivalent to the integral in Equation (12.7.2). As in (b), think carefully about the geometry first.

d. Set up an iterated integral, integrating first with respect to x, then y, then z that is equivalent to the integral in Equation (12.7.2).

Now that we have begun to understand how to set up iterated triple integrals, we can apply them to determine important quantities, such as those found in the next activity.

Activity 12.7.4 A solid S is bounded below by the square $z = 0$, $-1 \leq x \leq 1$, $-1 \leq y \leq 1$ and above by the surface $z = 2 - x^2 - y^2$. A picture of the solid is shown in Figure 12.7.7.

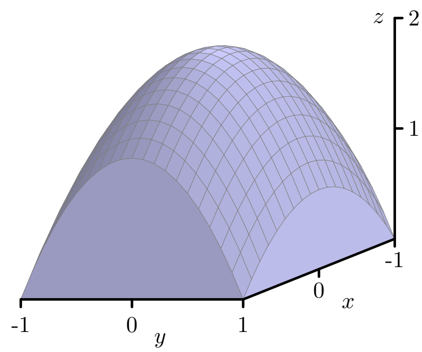


Figure 12.7.7 The solid bounded by the surface $z = 2 - x^2 - y^2$.

a. First, set up an iterated double integral to find the volume of the solid S as a double integral of a solid under a surface. Then set up an iterated triple integral that gives the volume of the solid S. You do not need to evaluate either integral. Compare the two approaches.

b. Set up (but do not evaluate) iterated integral expressions that will tell us the center of mass of S, if the density at point (x, y, z) is $\delta(x, y, z) = x^2 + 1$.

c. Set up (but do not evaluate) an iterated integral to find the average density on S using the density function from part (b).

d. Use technology appropriately to evaluate the iterated integrals you determined in (a), (b), and (c); does the location you determined for the center of mass make sense?

12.7.2 Summary

- Let $f = f(x, y, z)$ be a continuous function on a box $B = [a, b] \times [c, d] \times [r, s]$. The triple integral of f over B is defined as

$$\iiint_B f(x, y, z)\, dV = \lim_{\Delta V \to 0} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} f(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*) \cdot \Delta V,$$

where the triple Riemann sum is defined in the usual way. The definition of the triple integral naturally extends to non-rectangular solid regions S.

- The triple integral $\iiint_S f(x, y, z)\, dV$ can tell us
  - -.
    the volume of the solid S if $f(x, y, z) = 1$,
  - -.
    the mass of the solid S if f represents the density of S at the point (x, y, z).

  Moreover,
  $$f_{\mathrm{AVG}(S)} = \frac{1}{V(S)} \iiint_S f(x, y, z)\, dV,$$
  is the average value of f over S.

12.7.3 Exercises

1. Find the triple integral of the function $f(x, y, z) = x^4 \cos(y + z)$ over the cube $6 \le x \le 8$, $0 \le y \le \pi$, $0 \le z \le \pi$.

2. Evaluate the triple integral
$$\iiint_{\mathbf{E}} xyz\, dV$$
where $\boldsymbol{E}$ is the solid: $0 \le z \le 1, \quad 0 \le y \le z, \quad 0 \le x \le y.$

3. Find the mass of the rectangular prism $0 \le x \le 1, \quad 0 \le y \le 4, \quad 0 \le z \le 3$, with density function $\rho\,(x, y, z) = x$.

4. Find the average value of the function $f\,(x, y, z) = ye^{-xy}$ over the rectangular prism $0 \le x \le 2$, $0 \le y \le 3$, $0 \le z \le 1$

5. Find the volume of the solid bounded by the planes x = 0, y = 0, z = 0, and x + y + z = 4.

6. Find the mass of the solid bounded by the xy-plane, yz-plane, xz-plane, and the plane $(x/3) + (y/4) + (z/12) = 1$, if the density of the solid is given by $\delta(x, y, z) = x + 4y$.
mass = ______________________________

7. 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 54 given by $0 \le x \le 3$, $0 \le y \le 3$, $0 \le z \le 2$.
I_x = ______________________________
I_y = ______________________________
I_z = ______________________________

8. 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 $z = 0, x = 0, z = y - 6x$ and $y = 12$.
1. $\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x, y, z) dz dy dx$
a = ____________________ b = ____________________

$g_1(x) =$ ____________ $g_2(x) =$ ____________
$h_1(x,y) =$ ____________ $h_2(x,y) =$ ____________

2. $\displaystyle\int_a^b \int_{g_1(y)}^{g_2(y)} \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)dzdxdy$

$a =$ ____________ $b =$ ____________
$g_1(y) =$ ____________ $g_2(y) =$ ____________
$h_1(x,y) =$ ____________ $h_2(x,y) =$ ____________

3. $\displaystyle\int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(y,z)}^{h_2(y,z)} f(x,y,z)dxdydz$

$a =$ ____________ $b =$ ____________
$g_1(z) =$ ____________ $g_2(z) =$ ____________
$h_1(y,z) =$ ____________ $h_2(y,z) =$ ____________

4. $\displaystyle\int_a^b \int_{g_1(y)}^{g_2(y)} \int_{h_1(y,z)}^{h_2(y,z)} f(x,y,z)dxdzdy$

$a =$ ____________ $b =$ ____________
$g_1(y) =$ ____________ $g_2(y) =$ ____________
$h_1(y,z) =$ ____________ $h_2(y,z) =$ ____________

5. $\displaystyle\int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,z)}^{h_2(x,z)} f(x,y,z)dydzdx$

$a =$ ____________ $b =$ ____________
$g_1(x) =$ ____________ $g_2(x) =$ ____________
$h_1(x,z) =$ ____________ $h_2(x,z) =$ ____________

6. $\displaystyle\int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(x,z)}^{h_2(x,z)} f(x,y,z)dydxdz$

$a =$ ____________ $b =$ ____________
$g_1(z) =$ ____________ $g_2(z) =$ ____________
$h_1(x,z) =$ ____________ $h_2(x,z) =$ ____________

9. Calculate the volume under the elliptic paraboloid $z = 4x^2 + 7y^2$ and over the rectangle $R = [-1,1] \times [-1,1]$.

10. 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 5 and contained between the surfaces $z = \sqrt{x^2+y^2}$ and $z = 5$. If C has constant mass density of 2 g/cm^3, find the z-coordinate of C's center of mass.

$\overline{z} =$ ______________________

(Include units[1].)

11. 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) $\displaystyle\iiint\limits_W (z-2)\ dV$

[1] /webwork2_files/helpFiles/Units.html

(b) $\displaystyle\iiint_W e^{-xyz}\,dV$

(c) $\displaystyle\iiint_W \left(z - \sqrt{x^2+y^2}\right)\,dV$

12. 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 $x/3+y/2+z/6=1$, 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.

$$\iiint_R f(x,y,z)\,dV = \int_A^B\int_C^D\int_E^F \underline{\hspace{6cm}}\,dz\,dy\,dx$$

with limits of integration

A = ____________________

B = ____________________

C = ____________________

D = ____________________

E = ____________________

F = ____________________

13. 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 12.7.8.

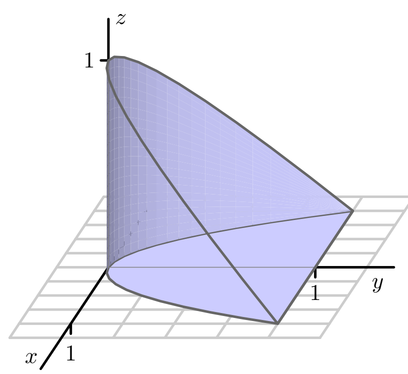


Figure 12.7.8 The solid bounded by $y=x^2$ and the planes $z=0$ and $z=1-y$.

Assume the density of S is given by $\delta(x,y,z)=z$

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 12.7.9.

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 12.7.9.

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 12.7.9.

d. Which of these three orders of integration is the most natural to you? Why?

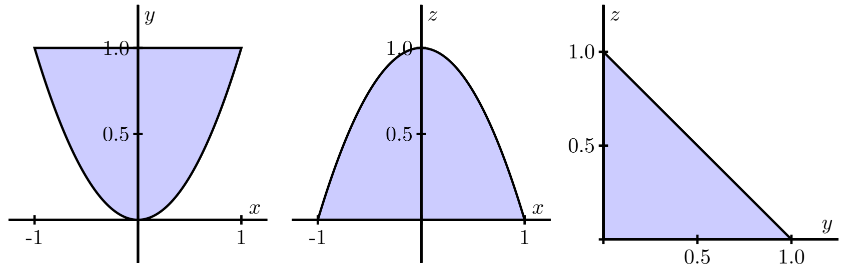


Figure 12.7.9 Projections of S onto the xy, xz, and yz-planes.

14. 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?

15. 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?

16. 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

Note well: 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 are the cylindrical coordinates of a point, and how are they related to Cartesian coordinates?
- 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 are the spherical coordinates of a point, and how are they related to Cartesian 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?

Just as with rectangular coordinates, where we usually write z as a function of x and y to plot the resulting surface, in cylindrical coordinates, we often express z as a function of r and θ. In the following activity, we explore several basic equations in cylindrical coordinates and the corresponding surface each generates.

Activity 12.8.1 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 your answer 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?

As the name and Activity 12.8.1 suggests, cylindrical coordinates are useful for describing surfaces that are cylindrical in nature.

12.8.1 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.

Activity 12.8.2 A picture of a cylindrical box, $B = \{(r, \theta, z) : r_1 \leq r \leq r_2, \theta_1 \leq \theta \leq \theta_2, z_1 \leq z \leq z_2\}$, is shown in Figure 12.8.1. Let $\Delta r = r_2 - r_1$, $\Delta\theta = \theta_2 - \theta_1$, and $\Delta z = z_2 - z_1$. We want to determine the volume ΔV of B in terms of Δr, $\Delta\theta$, Δz, r, θ, and z.

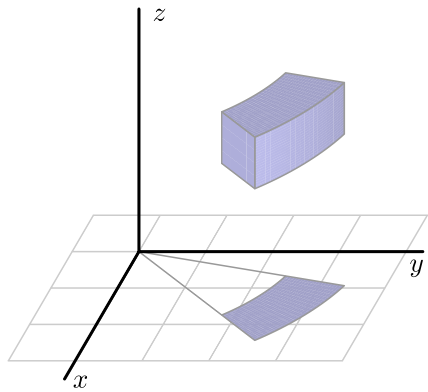


Figure 12.8.1 A cylindrical box.

a. Appropriately label Δr, $\Delta\theta$, and Δz in Figure 12.8.1.

b. Let ΔA be the area of the projection of the box, B, onto the xy-plane, which is shaded blue in Figure 12.8.1. Recall that we previously determined the area ΔA in polar coordinates in terms of r, Δr, and $\Delta\theta$. In light of the fact that we know ΔA and that z is the standard z coordinate from Cartesian coordinates, what is the volume ΔV in cylindrical coordinates?

Activity 12.8.2 demonstrates that the volume element dV in cylindrical coordinates is given by $dV = r\,dz\,dr\,d\theta$, and hence the following rule holds in general.

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.*

Of course, to complete the task of writing an iterated integral in cylindrical coordinates, we need to determine the limits on the three integrals: 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 12.8.3 In this activity we work with triple integrals in cylindrical coordinates.

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 polar coordinates.

   ii. Now describe the surfaces bounding the solid S using cylindrical coordinates.

   iii. Determine an iterated triple integral expression in cylindrical coordinates that gives 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 at left in Figure 12.8.2, and the projection of the cone onto the xy-plane in given at right 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.

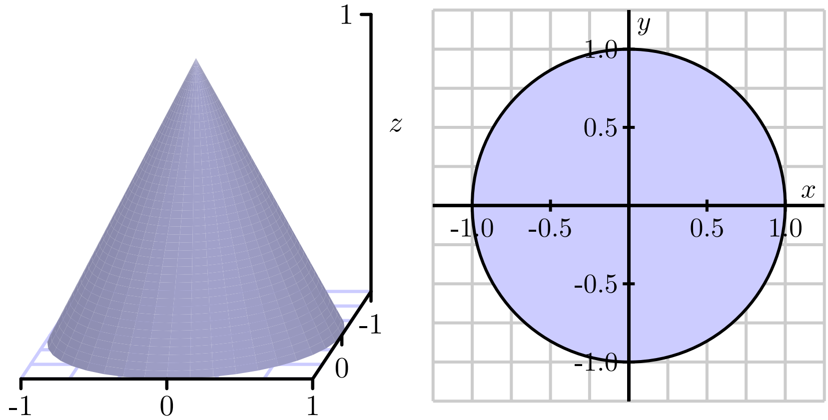


Figure 12.8.2 The cylindrical cone $r = 1 - z$ and its projection onto the xy-plane.

c. Determine 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.

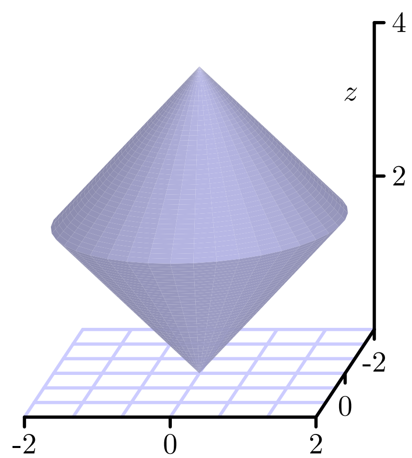


Figure 12.8.3 A solid bounded by the cones $z = \sqrt{x^2 + y^2}$ and $z = 4 - \sqrt{x^2 + y^2}$.

12.8.2 Spherical Coordinates

When it comes to thinking about particular surfaces in spherical coordinates, similar to our work with cylindrical and Cartesian 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.

In the following activity, we explore several basic equations in spherical coordinates and the surfaces they generate.

Activity 12.8.4 In this activity, we graph some surfaces using spherical 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 spherical coordinates with $\rho = 1$, $0 \le \theta \le 2\pi$, and $0 \le \phi \le \pi$? How does this particular example demonstrate the reason for the name of this coordinate system? What if we restrict ϕ to $0 \le \phi \le \frac{\pi}{2}$?

b. What familiar surface is described by the points in spherical coordinates with $\phi = \frac{\pi}{3}$, $0 \le \rho \le 1$, and $0 \le \theta \le 2\pi$?

c. What familiar shape is described by the points in spherical coordinates with $\theta = \frac{\pi}{6}$, $0 \le \rho \le 1$, and $0 \le \phi \le \pi$?

d. Plot the graph of $\rho = \theta$, for $0 \le \phi \le \pi$ and $0 \le \theta \le 2\pi$. How does the resulting surface appear?

As the name and Activity 12.8.4 indicate, spherical coordinates are particularly useful for describing surfaces that are spherical in nature; they are also convenient for working with certain conical surfaces.

12.8.3 Triple Integrals in Spherical Coordinates

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.

Activity 12.8.5 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 $\rho_1 \le \rho \le \rho_2$ (with $\Delta\rho = \rho_2 - \rho_1$), $\phi_1 \le \phi \le \phi_2$ (with $\Delta\phi = \phi_2 - \phi_1$), and $\theta_1 \le \theta \le \theta_2$ (with $\Delta\theta = \theta_2 - \theta_1$). An illustration of such a box is given at left 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. In other words,

$$\Delta V \approx |PS|\,|\overparen{PR}|\,|\overparen{PQ}|,$$

where $|\overparen{PR}|$ denotes the length of the circular arc from P to R.

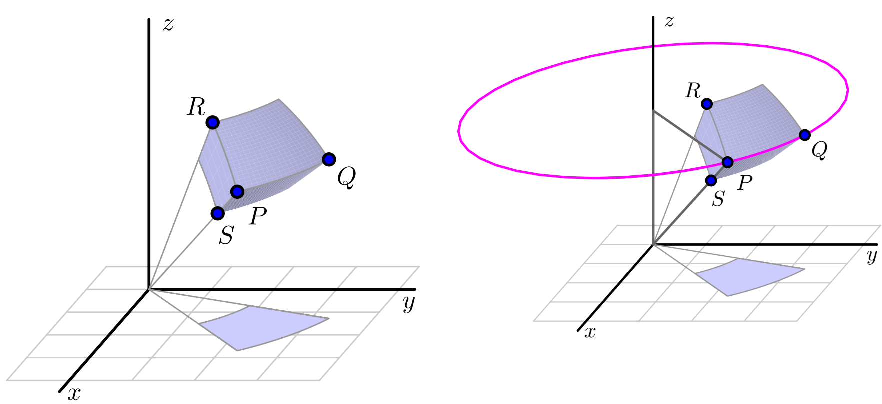


Figure 12.8.4 Left: A spherical box. Right: A spherical volume element.

a. What is the length $|PS|$ in terms of ρ?

b. What is the length of the arc $\overset{\frown}{PR}$? (Hint: The arc $\overset{\frown}{PR}$ is an arc of a circle of radius ρ_2, and arc length along a circle is the product of the angle measure (in radians) and the circle's radius.)

c. What is the length of the arc $\overset{\frown}{PQ}$? (Hint: The arc $\overset{\frown}{PQ}$ lies on a horizontal circle as illustrated at right in Figure 12.8.4. What is the radius of this circle?)

d. Use your work in (a), (b), and (c) to determine an approximation for ΔV in spherical coordinates.

Letting $\Delta\rho$, $\Delta\theta$, and $\Delta\phi$ go to 0, it follows from the final result in Activity 12.8.5 that $dV = \rho^2 \sin(\phi)\, d\rho\, d\theta\, d\phi$ in spherical coordinates, and thus allows us to state the following general rule.

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*.

Finally, in order to actually evaluate an iterated integral in spherical coordinates, we must of course determine the limits of integration in ϕ, θ, and ρ. The process is similar to our earlier work in the other two coordinate systems.

Activity 12.8.6 We can 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 of 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 cone $\phi = \frac{\pi}{4}$ from the sphere $\rho = 2$ if the density δ at the point (x, y, z) is $\delta(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. An illustration of the solid is shown in Figure 12.8.5.

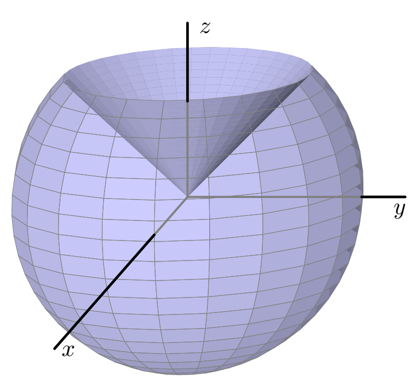


Figure 12.8.5 The solid cut from the sphere $\rho = 2$ by the cone $\phi = \frac{\pi}{4}$.

12.8.4 Summary

- 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 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

$$\iiint_S f(r\cos(\theta), r\sin(\theta), z)\,r\,dz\,dr\,d\theta.$$

- 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 Cartesian 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, ϕ, θ), its rectangular coordinates are

$$x = \rho\sin(\phi)\cos(\theta), \qquad y = \rho\sin(\phi)\sin(\theta), \qquad z = \rho\cos(\phi).$$

- 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

$$\iiint_S f(\rho\sin(\phi)\cos(\theta), \rho\sin(\phi)\sin(\theta), \rho\cos(\phi))\,\rho^2\sin(\phi)\,d\rho\,d\theta\,d\phi.$$

12.8.5 Exercises

1. Match the given equation with the verbal description of the surface:

A. Half plane

B. Elliptic or Circular Paraboloid

C. Cone

D. Sphere

E. Circular Cylinder

F. Plane

(a) $r = 4$

(b) $\rho\cos(\phi) = 4$

(c) $z = r^2$

(d) $\rho = 4$

(e) $\rho = 2\cos(\phi)$

(f) $r = 2\cos(\theta)$

(g) $\theta = \dfrac{\pi}{3}$

(h) $r^2 + z^2 = 16$

(i) $\phi = \dfrac{\pi}{3}$

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) $\displaystyle\iint_D \frac{1}{x^2+y^2}\, dA$ where D is: $x^2 + y^2 \le 4$

(b) $\displaystyle\iiint_E z^2\, dV$ where E is: $-2 \le z \le 2, \quad 1 \le x^2 + y^2 \le 2$

(c) $\displaystyle\iiint_E z\, dV$ where E is: $1 \le x \le 2, \quad 3 \le y \le 4, \quad 5 \le z \le 6$

(d) $\displaystyle\int_0^1 \int_0^{y^2} \frac{1}{x}\, dx\, dy$

(e) $\displaystyle\iiint_E dV$ 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. cylindrical coordinates

C. polar coordinates

D. cartesian coordinates

3. Evaluate the integral.

$$\int_0^5 \int_{-\sqrt{25-x^2}}^{\sqrt{25-x^2}} \int_{-\sqrt{25-x^2-z^2}}^{\sqrt{25-x^2-z^2}} \frac{1}{(x^2+y^2+z^2)^{1/2}}\, dy\, dz\, dx =$$

____

4. Use cylindrical coordinates to evaluate the triple integral $\displaystyle\iiint_{\mathbf{E}} \sqrt{x^2+y^2}\, dV$, where $\boldsymbol{E}$ is the solid bounded by the circular paraboloid $z = 4 - 9\left(x^2 + y^2\right)$ and the xy -plane.

5. Use spherical coordinates to evaluate the triple integral $\displaystyle\iiint_{\mathbf{E}} x^2 + y^2 + z^2\, dV$, where $\boldsymbol{E}$ is the ball: $x^2 + y^2 + z^2 \le 49$.

6. Find the volume of the solid enclosed by the paraboloids $z = 16\left(x^2 + y^2\right)$ and $z = 18 - 16\left(x^2 + y^2\right)$.

7. FInd the volume of the ellipsoid $x^2 + y^2 + 10z^2 = 25$.

8. The density, δ, of the cylinder $x^2 + y^2 \leq 1$, $0 \leq z \leq 5$ varies with the distance, r, from the z-axis:

$$\delta = 4 + r \text{ g/cm}^3.$$

Find the mass of the cylinder, assuming x, y, z are in cm.

mass = ______

(Include units[1].)

9. 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 2. Enter ρ as *rho*, ϕ as *phi*, and θ as *theta*.

(a) As an iterated integral,

$$\iiint_W f\, dV = \int_A^B \int_C^D \int_E^F \text{\_\_\_\_\_\_} \; d\rho\, d\phi\, d\theta$$

with limits of integration

A = ______

B = ______

C = ______

D = ______

E = ______

F = ______

(b) Evaluate the integral. ______

10. 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 $x^2 + y^2 + z^2 = 2z$.

11. 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.

[1] /webwork2_files/helpFiles/Units.html

a. $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{2-x^2-y^2}} xy\,dz\,dy\,dx$

b. $\displaystyle\int_0^{\pi/2} \int_0^{\pi} \int_0^1 \rho^2 \sin(\phi)\,d\rho\,d\phi\,d\theta$

c. $\displaystyle\int_0^{2\pi} \int_0^1 \int_r^1 r^2 \cos(\theta)\,dz\,dr\,d\theta$

12. 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 (c). 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?
- What is the Jacobian, and how is it related to a change of variables?

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 implicitly seen some of the issues that arise in using a change of variables with two or three variables present. In what follows, we seek to understand the general ideas behind any change of variables in a multiple integral.

Preview Activity 12.9.1 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.

a. 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.

b. When we write the double integral (12.9.1) as an iterated integral in polar coordinates we make a change of variables, namely

$$x = r\cos(\theta) \qquad \text{and} \qquad y = r\sin(\theta). \tag{12.9.2}$$

We also then have to change dA to $r\,dr\,d\theta$. This process also identifies a "polar rectangle" $[r_1, r_2] \times [\theta_1, \theta_2]$ with the original Cartesian rectangle, under the transformation[1] in Equation (12.9.2). The vertices of the polar rectangle are transformed into the vertices of a closed and bounded region in rectangular coordinates.

To work with a numerical example, let's now consider the polar rectangle P given by $[1, 2] \times [\frac{\pi}{6}, \frac{\pi}{4}]$, so that $r_1 = 1$, $r_2 = 2$, $\theta_1 = \frac{\pi}{6}$, and $\theta_2 = \frac{\pi}{4}$.

   i. Use the transformation determined by the equations in (12.9.2) to find the rectangular vertices that correspond to the polar vertices in the polar rectangle P. In other words, by substituting appropriate values of r and θ into the two equations in (12.9.2), find the values of the corresponding x and y coordinates for the vertices of the polar rectangle P. Label the point that corresponds to the polar vertex (r_1, θ_1) as (x_1, y_1), the point corresponding to the polar vertex (r_2, θ_1) as (x_2, y_2), the point corresponding to the polar vertex (r_1, θ_2) as (x_3, y_3), and the point corresponding to the polar vertex (r_2, θ_2) as (x_4, y_4).

   ii. Draw a picture of the figure in rectangular coordinates that has the points (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) as vertices. (Note carefully that because of the trigonometric functions in the transformation, this region will not look like a Cartesian rectangle.) What is the area of this region in rectangular coordinates? How does this area compare to the area of the original polar rectangle?

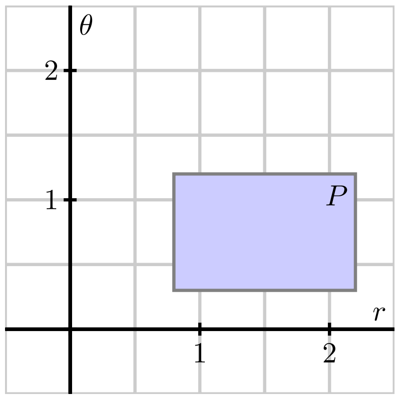
12.9.1 Change of Variables in Polar Coordinates

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

$$x = r\cos(\theta) \qquad \text{and} \qquad y = r\sin(\theta).$$

[1]A *transformation* is another name for function: here, the equations $x = r\cos(\theta)$ and $y = r\sin(\theta)$ define a function T by $T(r, \theta) = (r\cos(\theta), r\sin(\theta))$ so that T is a function (transformation) from $\mathbb{R}^2$ to $\mathbb{R}^2$. We view this transformation as mapping a version of the xy-plane where the axes are viewed as representing r and θ (the $r\theta$-plane) to the familiar xy-plane.

The vertices of the polar rectangle P 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 12.9.1. The image P' of the polar rectangle P under the transformation given by (12.9.2) is shown at right in Figure 12.9.1. 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 $T(r, \theta) = (r\cos(\theta), r\sin(\theta))$.



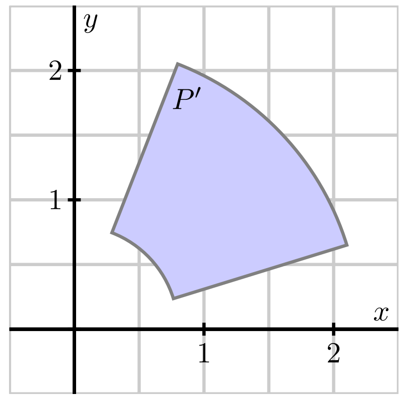


Figure 12.9.1 A rectangle P and its image P'.

Furthermore, as Preview Activity 12.9.1 suggests, 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 proceed to working with other transformations for different changes in coordinates, we have to understand how the transformation affects area so that we may use the correct area element in the new system of variables.

12.9.2 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

$$x = x(s, t) \qquad \text{and} \qquad y = y(s, t)$$

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 to the new s, t variables, we need to understand the area element, dA, in this new system. The following activity helps to illustrate the idea.

Activity 12.9.2 Consider the change of variables

$$x = s + 2t \qquad \text{and} \qquad y = 2s + \sqrt{t}.$$

Let's see what happens to the rectangle $T = [0, 1] \times [1, 4]$ in the st-plane under this change of variable.

a. Draw a labeled picture of T in the st-plane.

b. Find the image of the st-vertex $(0,1)$ in the xy-plane. Likewise, find the respective images of the other three vertices of the rectangle T: $(0,4)$, $(1,1)$, and $(1,4)$.

c. In the xy-plane, draw a labeled picture of the image, T', of the original st-rectangle T. What appears to be the shape of the image, T'?

d. To transform an integral with a change of variables, we need to determine the area element dA for image of the transformed rectangle. Note that T' is not exactly a parallelogram since the equations that define the transformation are not linear. But we can approximate the area of T' with the area of a parallelogram. How would we find the area of a parallelogram that approximates the area of the xy-figure T'? (Hint: Remember what the cross product of two vectors tells us.)

Activity 12.9.2 presents the general idea of how a change of variables works. 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 12.9.2.

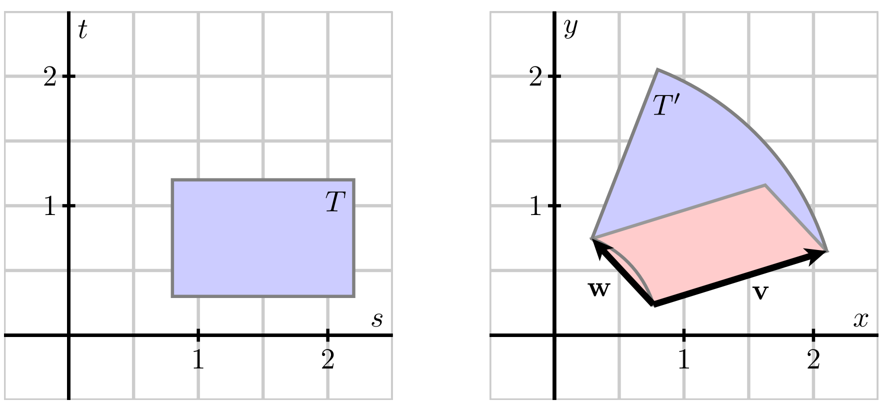


Figure 12.9.2 Approximating an area of an image resulting from a transformation.

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 similarly those for $\vec{w}$ 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

$$dA = \left| \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s} \right| ds\, dt. \tag{12.9.3}$$

Equation (12.9.3) hence determines the general change of variable formula in a double integral, and we can now say that

$$\iint_T f(x, y)\, dy\, dx = \iint_{T'} f(x(s, t), y(s, t)) \left| \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial y}{\partial s} \right| ds\, dt.$$

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 9.4 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 9.4, 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.

To summarize, the preceding change of variable formula that we have derived now follows.

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 Find the Jacobian when changing from rectangular to polar coordinates. That is, for the transformation given by $x = r\cos(\theta)$, $y = r\sin(\theta)$, determine a simplified expression for the quantity

$$\frac{\partial x}{\partial r}\frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta}\frac{\partial y}{\partial r}.$$

What do you observe about your result? How is this connected to our earlier work with 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 $x^2 + \frac{y^2}{4} = 1$. Here we will make a change of variables so that the pre-image of the domain is a circle.

- 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.
- 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.
- 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.

12.9.3 Change of Variables in a Triple Integral

The argument for the change of variable formula for triple integrals is complicated, and we will not go into the details. The general process, though, is the same as the two-dimensional case. 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 $f = f(x(s,t,u), y(s,y,u), z(s,t,u))$ 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 9.4.) That is, $\frac{\partial(x,y,z)}{\partial(s,t,u)}$ is given by

$$\frac{\partial x}{\partial s}\left[\frac{\partial y}{\partial t}\frac{\partial z}{\partial u} - \frac{\partial y}{\partial u}\frac{\partial z}{\partial t}\right] - \frac{\partial x}{\partial t}\left[\frac{\partial y}{\partial s}\frac{\partial z}{\partial u} - \frac{\partial y}{\partial u}\frac{\partial z}{\partial s}\right] + \frac{\partial x}{\partial u}\left[\frac{\partial y}{\partial s}\frac{\partial z}{\partial t} - \frac{\partial y}{\partial t}\frac{\partial z}{\partial s}\right].$$

To summarize,

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 $0 \leq x \leq 2$, $\frac{x}{2} \leq y \leq \frac{x}{2}+1$, and $0 \leq z \leq 6$. 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-coordinatte descriptions of the box S.

b. Find 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

$$\iiint_S f(x(s,t,u), y(s,t,u), z(s,t,u)) \left|\frac{\partial(x,y,z)}{\partial(s,t,u)}\right| ds\,dt\,du.$$

12.9.4 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. Find the absolute value of the Jacobian, $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$, for the change of variables given by $x = 2s + 2t, y = 5s + 5t$

$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ ______________________________

2. Find the Jacobian. $\frac{\partial(x,y,z)}{\partial(s,t,u)}$, where $x = 4t - 2s + 5u, y = -(3s + 2t + 4u), z = 3s - 5t - 2u$.

$\frac{\partial(x,y,z)}{\partial(s,t,u)} =$ ______________________________

3. 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 $S : -40 \leq u \leq 40, -40 \leq v \leq 40$ 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 $\iint_{T(S)} x^2 + y^2\, dA$

4. 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 $y = 0$, $y = 4$, $y = x^2$, and $y = x^2 - 6$.

$\int\int_R x\,dx\,dy =$ ______________________________

5. 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 = ______________________________

6. Use the change of variables $s = xy$, $t = xy^2$ to compute $\int_R xy^2\, dA$, where R is the region bounded by $xy = 4,\ xy = 7,\ xy^2 = 4,\ xy^2 = 7$.
$\int_R xy^2\, dA =$ ______________________________

7. 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 $0 \le x \le 70$, $0 \le y \le 15$ and the region T, the square $0 \le s, t \le 1$.
$a =$ ______________________________
$b =$ ______________________________
What is $\left|\frac{\partial(x,y)}{\partial(s,t)}\right|$ in this case?
$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ ______________________________

8. 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 $(0,0)$, $(10,0)$, $(-18,8)$, $(-8,8)$ 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?
$\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ ______________________________

9. In this problem we use the change of variables $x = 2s + t$, $y = s - 3t$ to compute the integral $\int_R (x + y)\, dA$, where R is the parallelogram with vertices $(x, y) = (0, 0)$, $(2, 1)$, $(3, -2)$, and $(1, -3)$.
First find the magnitude of the Jacobian, $\left|\frac{\partial(x,y)}{\partial(s,t)}\right| =$ ____________.
Then, with $a =$ ____________, $b =$ ____________,
$c =$ ____________, and $d =$ ____________,
$\int_R (x + y)\, dA = \int_a^b \int_c^d ($ _____ $s+$ _____ $t+$ _____ $)\ dt\, ds =$
____

10. 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.

11. 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?

12. 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 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?

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 10.1.2 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 13.1.1 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 $\vec{G}(x,y) = -x\hat{\jmath}$ 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 13.1.1, 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 13.1.1, plot the vectors $\vec{G}(x,y)$ for the eight points in the table in part b.

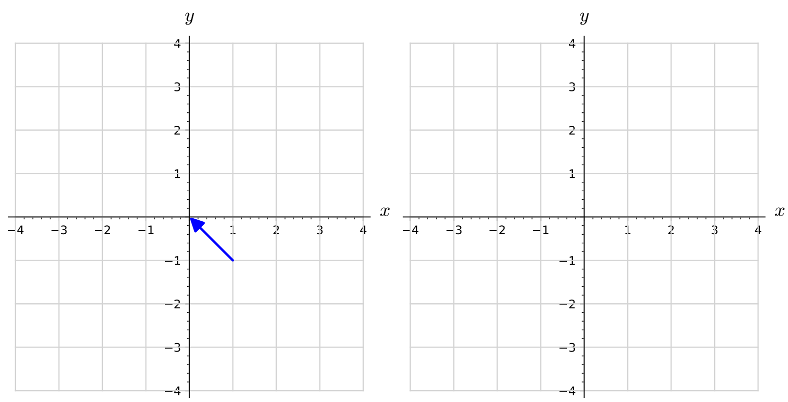


Figure 13.1.1 Axes for plotting some vectors from $\vec{F}(x,y)$ and $\vec{G}(x,y)$.

13.1.1 Examples of Vector Fields

As Preview Activity 13.1.1 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 13.1.2 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.

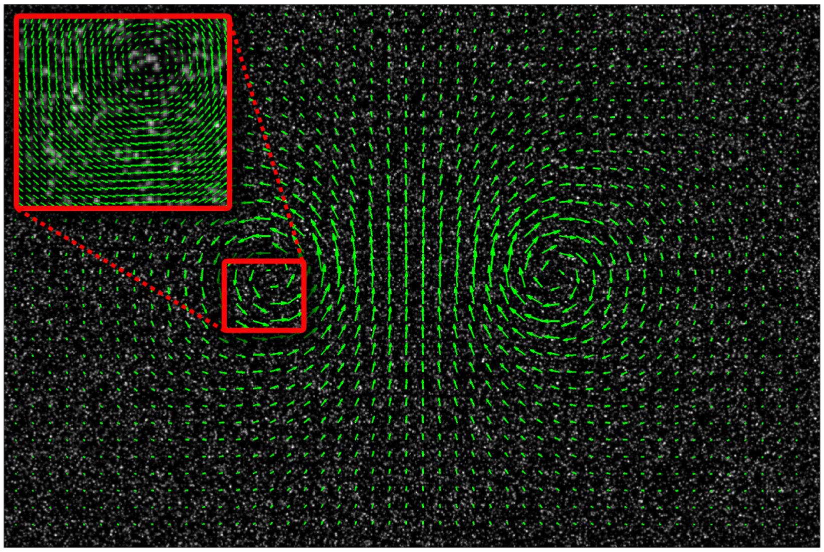
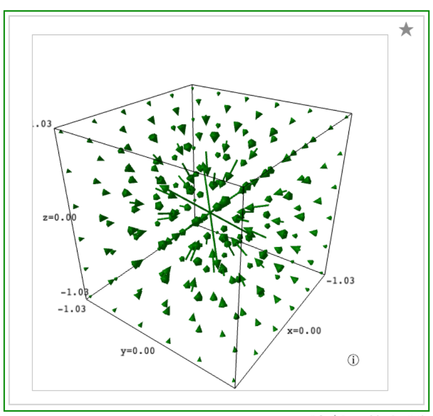


Figure 13.1.2 An illustration of some of the vectors in a fluid velocity vector field. "PIVlab multipass" by Willa[1] Licensed under CC-BY-SA 3.0 via Wikimedia Commons.

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 13.1.3, 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
Embed

Figure 13.1.3 Gravitational vector field.

13.1.2 Mathematical Vector Fields

As suggested in the introduction and Preview Activity 13.1.1, vector fields can be specified using the notation of functions and vectors.

[1] commons.wikimedia.org/wiki/File:PIVlab_multipass.jpg#/media/File:PIVlab_multipass.jpg

Definition 13.1.4 A **vector field** in 2-space is 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.)

13.1.3 Plotting Vector Fields

Preview Activity 13.1.1 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 13.1.5.

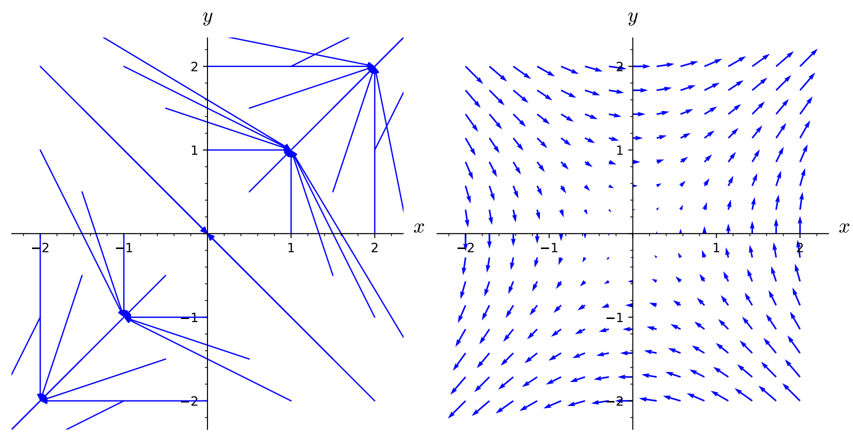


Figure 13.1.5 Two plots of $\vec{F}(x, y) = y\hat{\imath} + x\hat{\jmath}$ from Sage

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 13.1.2 The plot in Figure 13.1.6 illustrates the vector field $\vec{F}(x,y) = y\hat{\imath} - x\hat{\jmath}$.

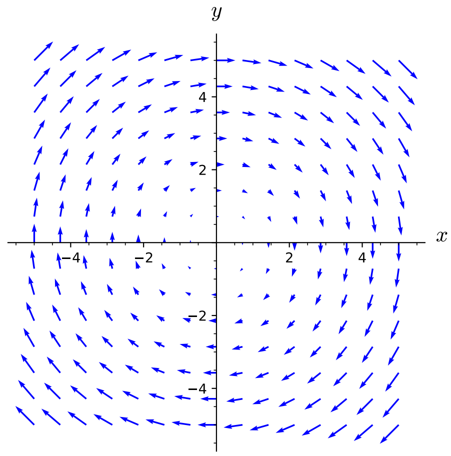


Figure 13.1.6 The vector field $y\hat{\imath} - x\hat{\jmath}$

(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 $|\vec{F}(x,y)|$ and the radius of that circle?

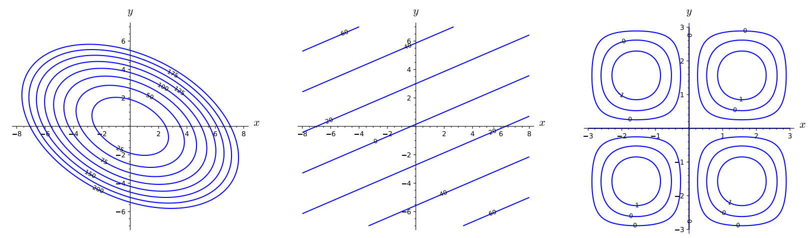
In the previous activity, we looked a special type of curve in a vector field, namely the curve that flows with the output of the vector field. Geometrically, the output vectors of the vector field will be tangent to the flow curves. These flow curves come up in several physical situations, including 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.

13.1.4 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 11.7.3 for a review).

Activity 13.1.3

(a) In Figure 13.1.7 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 13.1.7 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?

13.1.5 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.

13.1.6 Exercises

1. Compute and sketch the vector assigned to the points $P = (0, -6, -9)$ and $Q = (4, 1, 0)$ by the vector field $\mathbf{F} = \langle xy, z^2, x \rangle$.

$\mathbf{F}(P) =$ ____________________

$\mathbf{F}(Q) =$ ____________________

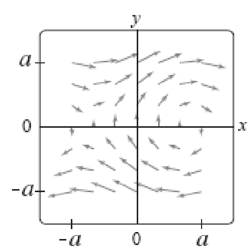
2. Suppose $\vec{F} = \langle -3y, 3x \rangle$. Complete the following table of values of $\vec{F}$.

Values of $\vec{F}$

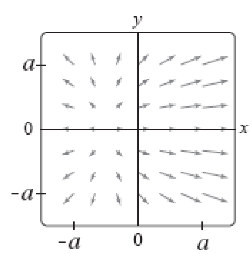
$y = 1$	________	________	________
$y = 0$	________	________	________
$y = -1$	________	________	________
	$x = -1$	$x = 0$	$x = 1$

Using your table of values as a starting point, sketch this vector field on a piece of paper for $-2 \le x \le 2$ and $-2 \le y \le 2$.

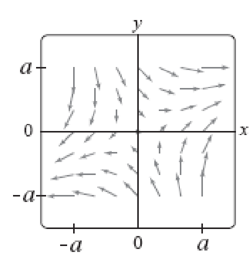
3. Match the planar vector field $\mathbf{F} = \langle 3x + 3, y \rangle$ with the corresponding plot in the Figures below.



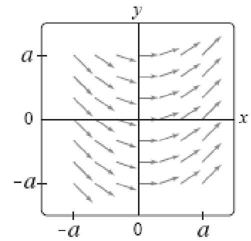
$\Leftarrow$ ***Plot 1***



$\Leftarrow$ ***Plot 2***



$\Leftarrow$ ***Plot 3***



⇐ ***Plot 4***
With $a = 3$
Answer :

- ⊙ Plot 1
- ⊙ Plot 4
- ⊙ Plot 3
- ⊙ Plot 2

4. 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

13.1.7 Notes to the Instructor

This section uses tools from the chapter on multivariable functions and thier derivatives, with specific references to gradients. Additionally, vector calculations and geometry are used throughout to understand the output of the vector field.

13.2 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?

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 fuild 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 13.1, 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 9.3 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 compass rose. The compass rose given in Figure 13.2.1 is an example of a sixteen point rose. Note that directions like ESE are read as "east-southeast" half way between east and southeast.

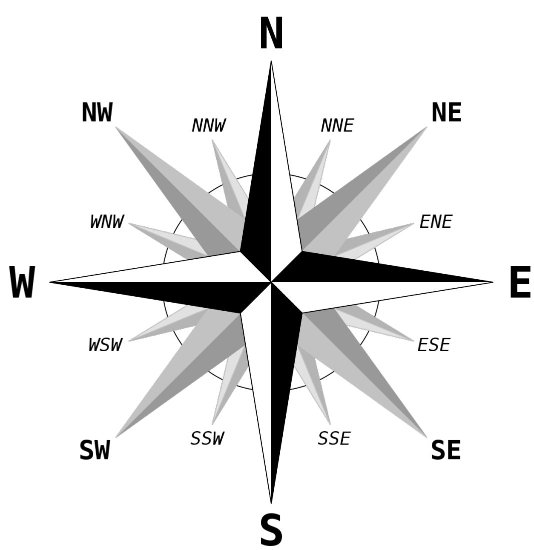


Figure 13.2.1 An example of a sixteen point compass rose (from By Brosen~commonswiki - Own work, CC BY 2.5[1])

Preview Activity 13.2.1 Recall from Section 9.3 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.

[1] commons.wikimedia.org/w/index.php?curid=667281

(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 thier 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 13.2.1.c connect to the answers to part 13.2.1.a and part 13.2.1.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.

13.2.1 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 13.2.2) than if the object moves along the line segment from $(0,1)$ directly to $(4,3)$(illustrated by C_2 in Figure 13.2.2). 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 13.2.2). 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.

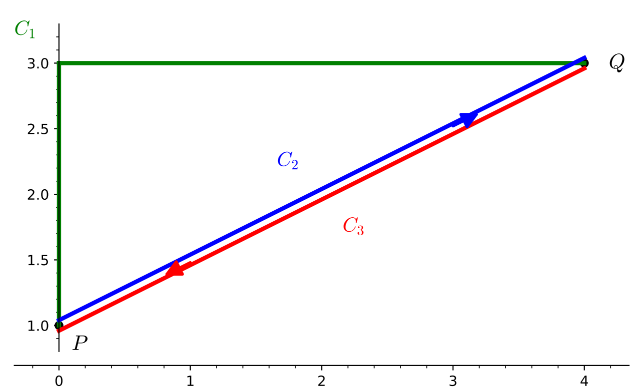


Figure 13.2.2 Curves C_1, C_2, and C_3 demonstrating different paths and orientations between P and Q

Activity 13.2.2 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 $(3,-1,2)$.

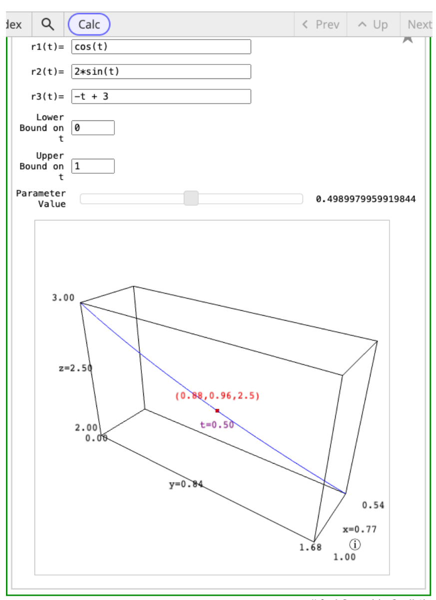
(b) The line segment in $\mathbb{R}^3$ from $(3,-1,2)$ to $(0,1,-2)$.

(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 coefficents)
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
Embed

Figure 13.2.3 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 $r_3(t)$.

13.2.2 Line Integrals

Preview Activity 13.2.1 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.

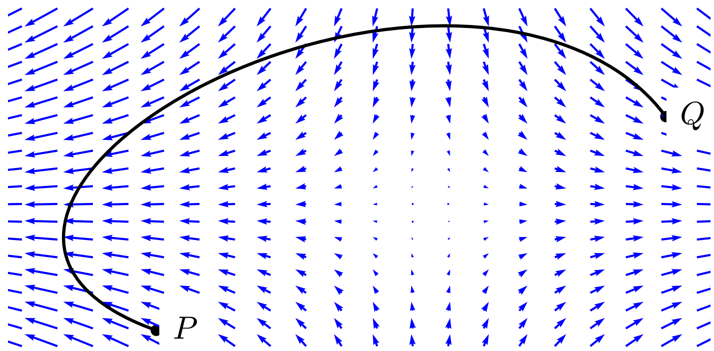


Figure 13.2.4 A curve C oriented from the point P to the point Q with the plot of a vector field $\vec{F}$

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 13.2.5 the actual output vectors are plotted for some points on C.

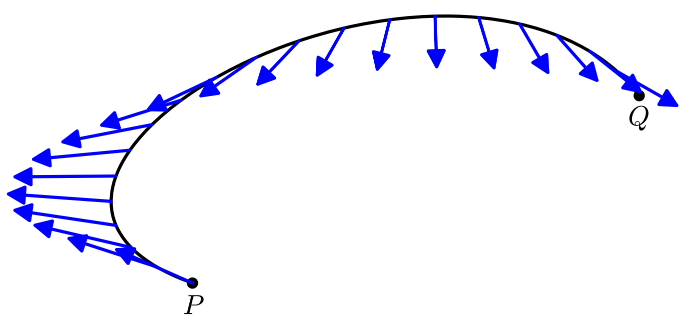


Figure 13.2.5 A curve C oriented from the point P to the point Q with the vector field, $\vec{F}$, plotted at points along the curve

Since the output of $\vec{F}$ is changing as you move along the path, we will use a type of argument that has come up repeatedly throughout our study of integral calculus. Here we will break our region up into smaller pieces to approximate the work done on each piece. Figure 13.2.6 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 $\mathbf{r}_i$.

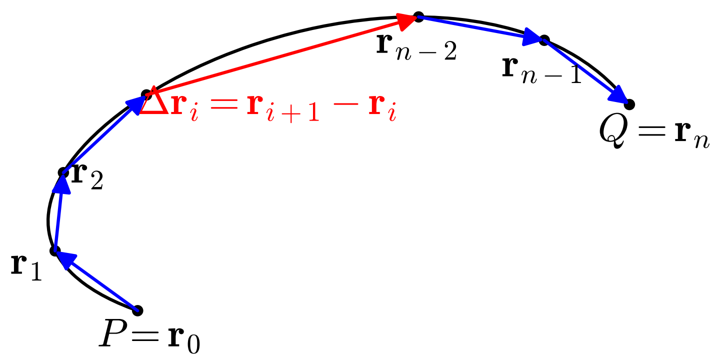


Figure 13.2.6 A curve C oriented from the point P to the point Q broken into n pieces

If we look at one of these smaller pieces (as show in Figure 13.2.7), 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 9.3; 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.

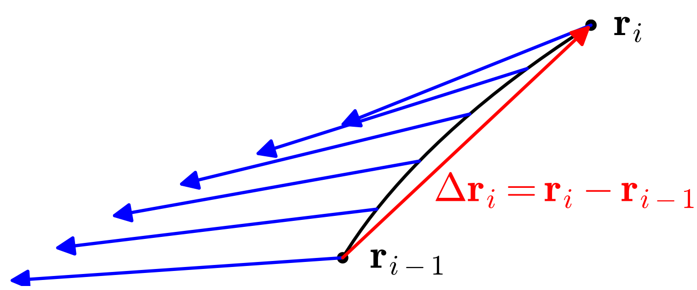


Figure 13.2.7 A plot of the piece C_i with the output of our vector field plotted at six poins on C_i

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 means that our approximation with n pieces of C would be

$$\sum_{i=1}^{n} \vec{F}(\mathbf{r}_i) \cdot \Delta\mathbf{r}_i.$$

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.

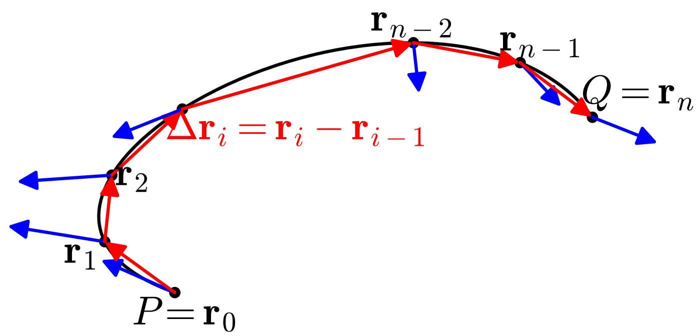


Figure 13.2.8 A curve C oriented from the point P to the point Q broken into n pieces with the output of $\vec{F}$ plotted at each intermediate point

Definition 13.2.9 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. ◊

We are able to say a bit here about conditions that guarantee the limit in Definition 13.2.9 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 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).

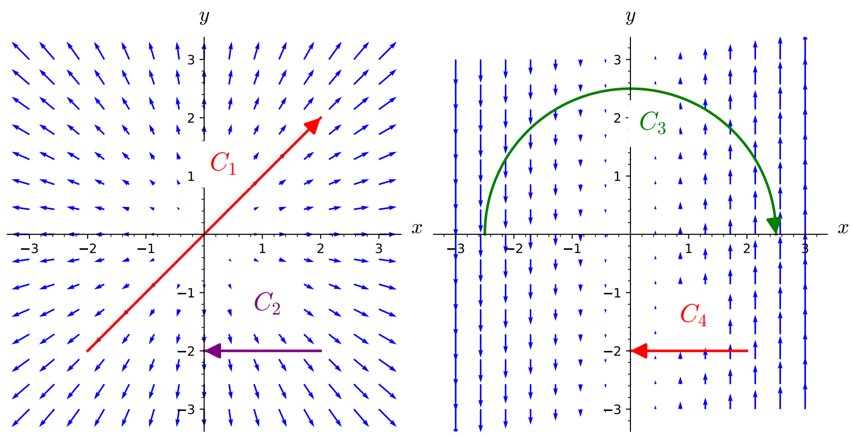
Alternative notation for line integrals. Our notation for line integrals is one of several common notations. This notation's strength is that it emphasizes the role of a vector field and dot product. Another common notation for the line integral of a vector field $\langle P, Q, R\rangle$ along a curve 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 13.3.2.

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. [2] 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

[2]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 $\vec{r} = \langle r_1, r_2\rangle$.

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 13.2.3 Shown in Figure 13.2.10 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.



(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 13.2.10 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 13.2.11 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 13.2.12 shows $\vec{F}$ and C, including the orientation. We will calculate $\displaystyle\int_C \vec{F} \cdot d\vec{r}$.

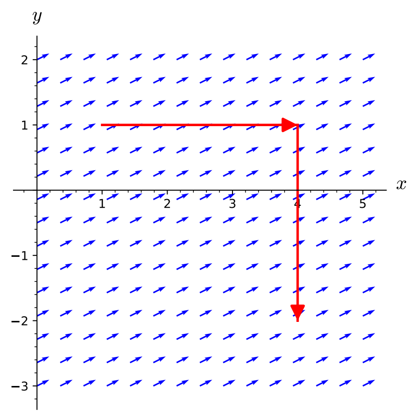


Figure 13.2.12 An oriented curve from $(1,1)$ to $(4,-2)$ in a vector field $\vec{F}$.

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 $\int_C \vec{F}\cdot d\vec{r} = 6 - 3 = 3$. □

13.2.3 Properties of Line Integrals

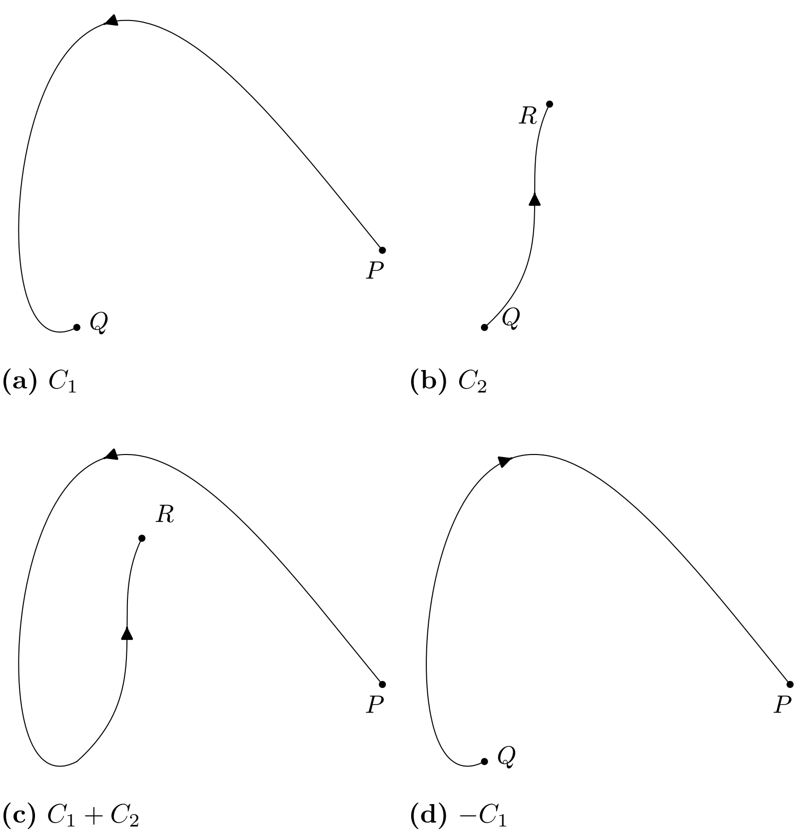
In Example 13.2.11, we implicitly made use of the idea that if C can be broken up into two curves C_1 and C_2 such that the terminal point of C_1 is the initial point of C_2, then the line integral of $\vec{F}$ along C is the sum of the line integrals of $\vec{F}$ along C_1 and along C_2. This is a generalization of the property for definite integrals[3] 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 13.2.13 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.

[3] activecalculus.org/single/sec-4-3-definite-integral.html#zac



(a) C_1 **(b)** C_2

(c) $C_1 + C_2$ **(d)** $-C_1$

Figure 13.2.14 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 13.2.4 Figure 13.2.15 shows a vector field $\vec{F}$ as well as six oriented curves, as labeled in the plot.

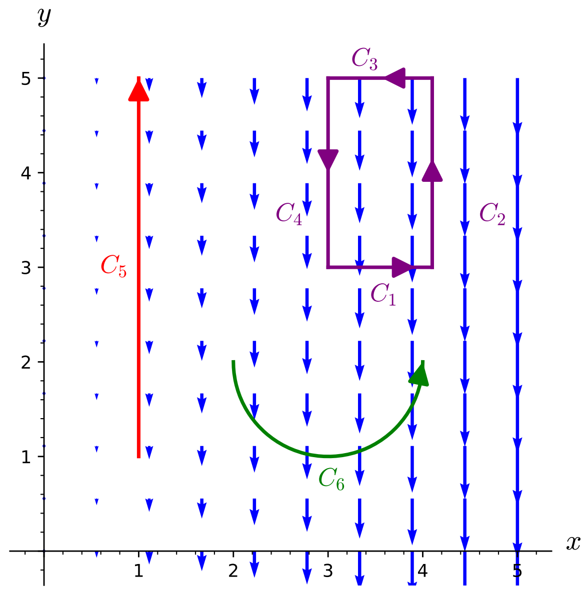


Figure 13.2.15 A vector field $\vec{F}$ and six oriented curves.

(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.

$$\int_{C_1} \vec{F} \cdot d\vec{r} \quad \int_{C_2} \vec{F} \cdot d\vec{r} \quad \int_{C_3} \vec{F} \cdot d\vec{r} \quad \int_{C_4} \vec{F} \cdot d\vec{r} \quad \int_{C_5} \vec{F} \cdot d\vec{r}$$

13.2.4 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 13.2.5 Determine if the circulation of the vector field around each of the closed curves shown in Figure 13.2.16 is positive, negative, or zero.

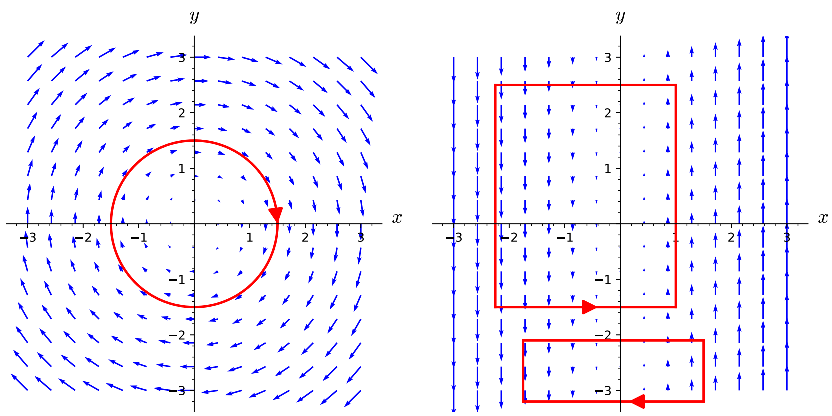


Figure 13.2.16 Vector fields and closed curves

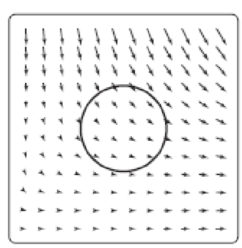
13.2.5 Summary

- 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.

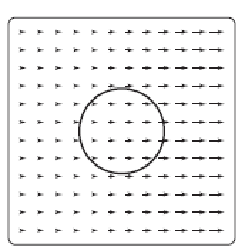
13.2.6 Exercises

1. 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.

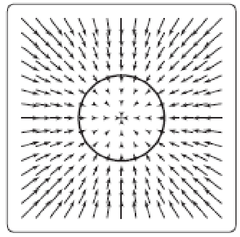
Note: Use "0" for zero, "P" for positive, and "N" for negative.



(A) __



(B)__



(C)__

2. 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 **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.

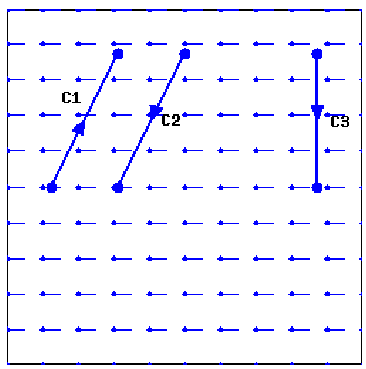
A. $\mathbf{F}$ = the radial vector field = $x\mathbf{i} + y\mathbf{j}$: __________

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 = $\mathbf{i} + \mathbf{j}$: __________

3. 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) .

4. 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 $\int_{C_3} \vec{F} \cdot d\vec{r} = 7$.

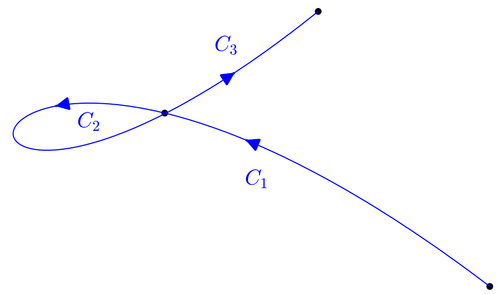


Figure 13.2.17 An oriented path broken into three pieces

Find the following:

(a) $\displaystyle\int_{-C_3} \vec{F} \cdot d\vec{r}$

(b) $\displaystyle\int_{C_2} \vec{F} \cdot d\vec{r}$

(c) $\displaystyle\int_{-C_1-C_3} \vec{F} \cdot d\vec{r}$

13.2.7 Notes to Instructors and Dependencies

This section relies heavily on understanding vector fields from Section 13.1, understanding curves in space (from Section 10.1), and the work interpretation of the dot product from Section 9.3.

13.3 Using Parametrizations 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?

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 13.3.1 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 13.3.1.

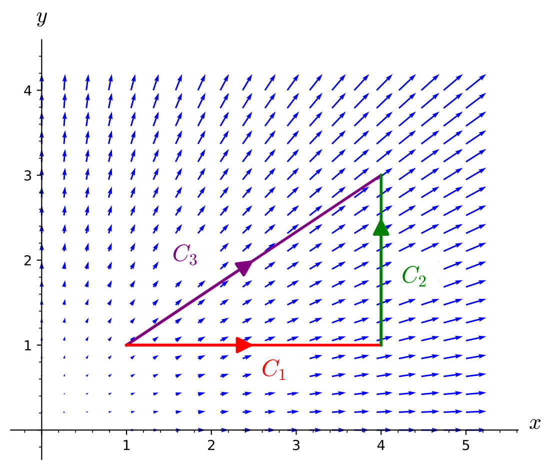


Figure 13.3.1 A vector field $\vec{F}$ and three oriented curves

(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.

13.3.1 Parametrizations in the Definition of $\int_C \vec{F} \cdot d\vec{r}$

In Preview Activity 13.3.1, 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 13.3.1.

Example 13.3.2 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

$$\begin{aligned}\int_{C_3} \vec{F} \cdot d\vec{r} &= \int_1^4 \left\langle \frac{2}{3}x^2 + \frac{1}{3}x, (\frac{2}{3}x + \frac{1}{3})^2 \right\rangle \cdot \left\langle 1, \frac{2}{3} \right\rangle \, dx \\ &= \int_1^4 \frac{2}{3}x^2 + \frac{1}{3}x + \frac{2}{3}\left(\frac{2}{3}x + \frac{1}{3}\right)^2 \, dx = \frac{151}{6}.\end{aligned}$$

Notice that this result is different than what you obtained for $\int_C \vec{F} \cdot d\vec{r}$ in Preview Activity 13.3.1, 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 13.3.2 to curves that are not line segments, it is worth examining this question for the vector field we have been investigating.

Activity 13.3.2 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 13.3.1 and Example 13.3.2 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 13.3.1 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, parametrizing 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 13.3.3 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 13.2.9. 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 $i = 1, \ldots, n$.

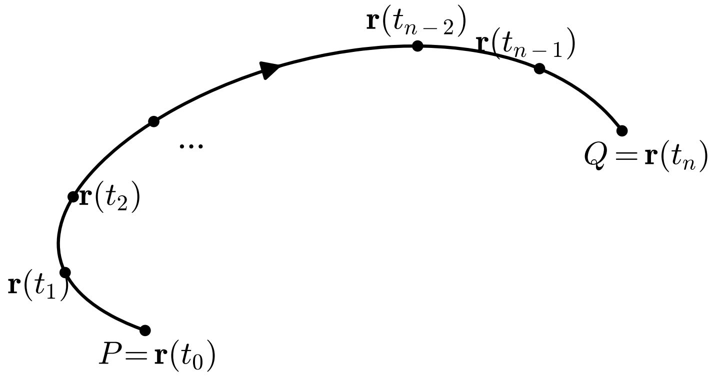


Figure 13.3.4 A curve C oriented from the point P to the point Q broken into n pieces evenly spaced in terms of the paramter t

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 13.2.6. Now notice that

$$\Delta\vec{r}_i = \vec{r}(t_i) - \vec{r}(t_{i-1}) = \vec{r}(t_i) - \vec{r}(t_i - \Delta t) = \left(\frac{\vec{r}(t_i) - \vec{r}(t_i - \Delta t)}{\Delta t}\right)\Delta t.$$

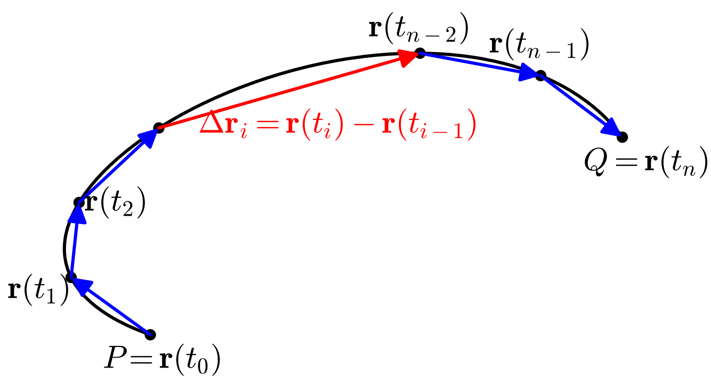


Figure 13.3.5 A curve C oriented from the point P to the point Q broken into n pieces evenly spaced in terms of the paramter t with displacement vectors Δr_i

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 13.2.9 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 13.3.6 *Let C be a smooth, oriented curve traced out by the vector-valued function $\vec{r}(t)$ for $a \leq t \leq b$. 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 13.3.6 is for evaluating line integrals, consider the following example.

Example 13.3.7 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 13.3.8. 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 $0 \leq t \leq \pi/2$.

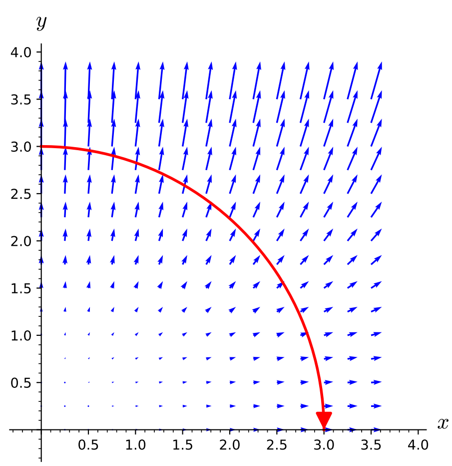


Figure 13.3.8 The vector field $\vec{F} = x\hat{\imath} + y^2\hat{\jmath}$ and an oriented curve C

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 13.3.7 shows, Theorem 13.3.6 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. The 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 13.3.3

(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 13.3.4 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 13.3.4 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 counterclockwise. 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?

13.3.2 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.

13.3.3 Independence of Parametrization for a Fixed Curve

Up to this point, we have chosen whatever parametrization of an oriented curve C was most convienent, and our argument for how we can use parametrizations to calculate line integrals did not depend on the specific choice of parametrization. However, it is not immediately obvious that different parametrizations don't result in different values of the line integral. Our next example explores this question.

Example 13.3.9 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 13.3.10, we can break C up into two oriented curves C_1 (the horizontal portion) and C_2 (the vertical portion) so that $C = C_1 + C_2$.

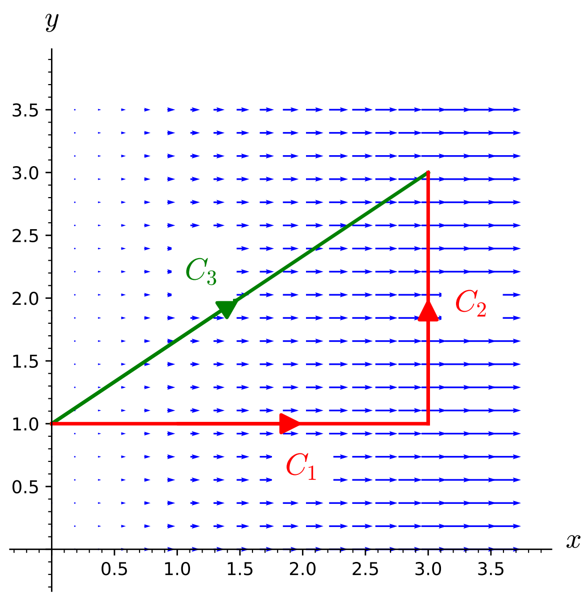


Figure 13.3.10 The vector field $\vec{F} = x\hat{\imath}$ and some oriented curves.

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 $0 \leq t \leq 3$ (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 $0 \leq t \leq \frac{\pi}{2}$. (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 13.3.5) 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 13.3.4.

Activity 13.3.5

(a) The typical parametrization of the line segment from $(0, 1)$ to $(3, 3)$ (the oriented curve C_3 in Example 13.3.9) is $\vec{r}(t) = \langle 3t, 1+2t\rangle$ where $0 \leq t \leq 1$. Use this parametrization to calculate $\int_{C_3} \vec{F}\cdot d\vec{r}$ for the vector field $\vec{F} = x\hat{\imath}$ and compare your answer to the result of Example 13.3.9.

(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 13.3.4?

Although we have not given a proof or even an intuitive argument, the phenomenon you observed in part a of Activity 13.3.5 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 13.3.6.

13.3.4 Summary

- Line integrals of vector fields along oriented curves can be evaluated by parametrizing 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 $\int_{C_2} \vec{F}\cdot d\vec{r}$.

13.3.5 Exercises

1. 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 $(9,0)$ centered at the origin oriented clockwise.

(a) Find a vector parametric equation $\vec{r}(t)$ for the portion of the ellipse described above for $0 \leq t \leq \pi/2$.

$\vec{r}(t) =$ ______________________

(b) Using your parametrization in part (a), set up an integral for calculating the circulation of $\vec{F}$ around C.

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt = \int_a^b \underline{\qquad\qquad\qquad}\, dt$$

with limits of integration $a =$ __________ and $b =$ __________

(c) Find the circulation of $\vec{F}$ around C.

Circulation = ______________________

2. Evaluate the line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$, where $\mathbf{F}(x, y, z) = 5x\mathbf{i} + 2y\mathbf{j} - 3z\mathbf{k}$ and C is given by the vector function $\mathbf{r}(t) = \langle \sin t, \cos t, t\rangle$, $\quad 0 \leq t \leq 3\pi/2$.

3. Evaluate the line integral $\int_{\mathcal{C}} \mathbf{F}\cdot d\,\mathbf{r}$ where $\mathbf{F} = \langle -\sin x, 4\cos y, xz\rangle$ and $\mathcal{C}$ is the path given by $\mathbf{r}(t) = (-t^3, t^2, -3t)$ for $0 \leq t \leq 1$

$\int_{\mathcal{C}} \mathbf{F}\cdot d\,\mathbf{r} =$ ______________________

4.

(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 $(0, 0)$.

5. 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.

6. Let C_3 be the circle of radius 7 centered at the origin traveled counter-clockwise. 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) $\langle M, N \rangle = \langle 3, x \rangle$

7. 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.

13.3.6 Notes to Instructors and Dependencies

This section relies heavily on the idea of line integrals developed in Section 13.2, understanding curves in space (from Section 10.1), and the work interpretation of the dot product from Section 9.3.

Preview Activity 13.4.1 compares its result with the first two parts of Activity 13.3.4, so prioritize getting those done in class.

13.4 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 $\vec{F} = \nabla f$?

In Activity 13.3.4, Example 13.3.9, and Activity 13.3.5, 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 13.4.1 In Activity 13.3.4, 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 13.3.4?

(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 13.4.1.b.

13.4.1 Path-Independent Vector Fields

Hopefully Preview Activity 13.4.1 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 13.3.4 and Example 13.3.9, 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[1] 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.

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[2], which can be written as

$$\int_a^b \frac{df}{dx}(x)\, dx = f(b) - f(a).$$

Theorem 13.4.1 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 13.4.2

(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 13.3.6 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 13.4.1 allows us to be more efficient in how we approach this.

In order to use Theorem 13.4.1, 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 13.4.3 for finding a protential 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 13.4.1,

$$\int_C \vec{F} \cdot d\vec{r} = f(4,-2) - f(2,0) = 3e^{-2} - 41$$

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

[1] activecalculus.org/single/sec-4-4-FTC.html#QmI
[2] https://activecalculus.org/single/sec-4-4-FTC.html#QmI

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 13.4.1 when we are given the potential function.

Activity 13.4.2 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 13.3.2, 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 13.4.1, 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 variables. However, in that case there are six mixed partials to calculate, and thus it can be rather tedious. Activity 13.4.3 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 13.4.3 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 13.4.3.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 13.4.3 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 13.4.4 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 $(0,0,0)$.

13.4.2 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 13.4.5 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 $\int_{C_2} \vec{G} \cdot d\vec{r}$?

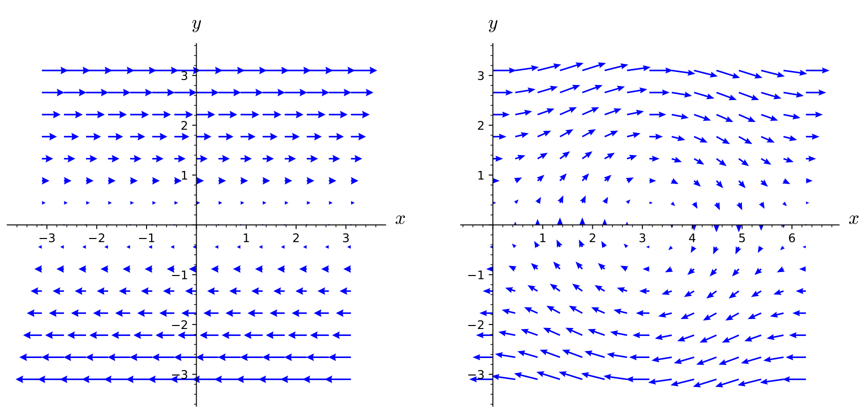
(h) Explain why this shows that $\vec{G}$ is path-independent.

We summarize the result of Activity 13.4.5 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 13.4.3 *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 13.4.6 Explain why neither of the vector fields in Figure 13.4.4 is path-independent.



(a) $\vec{F}$ **(b)** $\vec{G}$

Figure 13.4.4 Two vector fields that are not path-independent.

13.4.3 What other vector fields are path-independent?

Recall that in single variable calculus, The Second Fundamental Theorem of Calculus[3] 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 13.4.5). 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 13.4.5). 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.

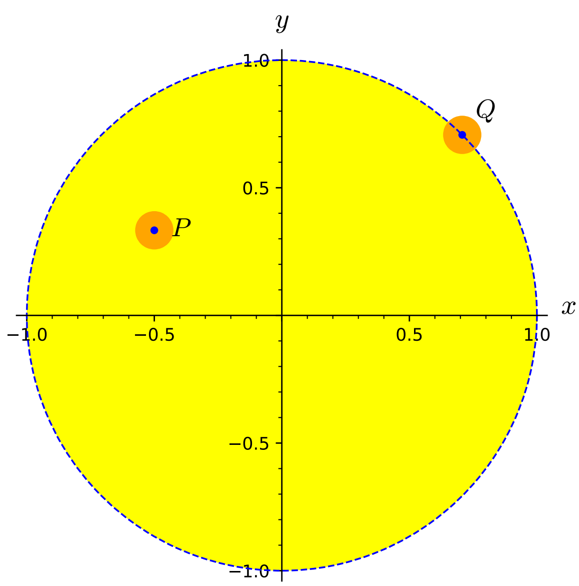


Figure 13.4.5 The open set $\{(x, y)|x^2 + y^2 < 1\}$ is plotted in yellow with a point Q on the boundary that shows why $\{(x, y)|x^2 + y^2 \leq 1\}$ is not open

Activity 13.4.7 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,

[3]activecalculus.org/single/sec-5-2-FTC2.html#qxB

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 13.4.6.) Rewrite $f(x, y)$ as a sum of $f(x_0, y_0)$ and line integrals along C_1, C_y, and C_x.

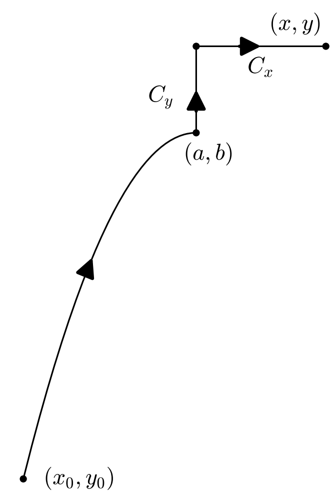


Figure 13.4.6 A piecewise smooth oriented curve from (x_0, y_0) to (x, y).

(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 13.3. 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 13.4.7 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 definnition of the potential function is in terms of a line integral.

Theorem 13.4.7 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. 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 $\vec{F}$.

13.4.4 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.

13.4.5 Exercises

1. 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 ____________________.

2. Consider the vector field $\mathbf{F}(x, y, z) = (3z + 3y)\,\mathbf{i} + (2z + 3x)\,\mathbf{j} + (2y + 3x)\,\mathbf{k}$.
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 $\int_C \mathbf{F} \cdot d\mathbf{r}$.

3. 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 13.4.1 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 $-\pi \le t \le \pi$.

(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 $(4, 16)$.

4. 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.

5. 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.

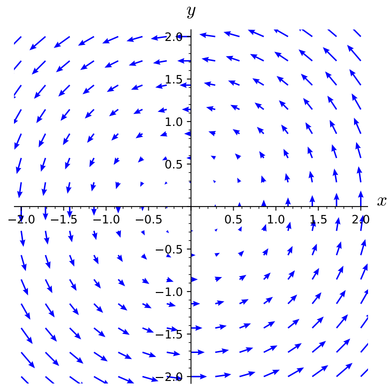


Figure 13.4.8 A vector field

(b) Explain why you think the following vector field is path independent or not.

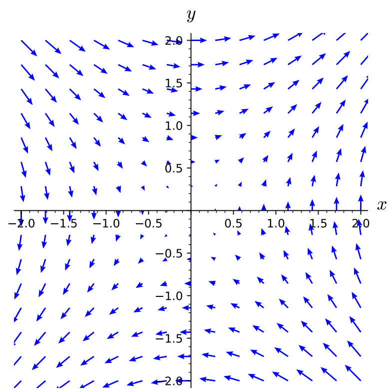


Figure 13.4.9 Another vector field

6. 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 $1 \le t \le 2$.

13.4.6 Notes to Instructors and Dependencies

This section is long, but important. To do all the activities, you will need multpile 50-minute class periods. One way to facilitate this would be to treat Activity 13.4.4 as if it were a preview activity and have students work on it before the second class meeting. You could also opt to de-emphasize finding potential functions and omit that activity.

Activity 13.4.7 is also rather long. We have arranged the subsection by placing Theorem 13.4.7 after the activity so as to not spoil the discovery to which the activity builds. This activity can be skipped without adversely impacting the remainder of the chapter. However, if choosing to omit Activity 13.4.7, you may wish to specifically point out the culminating theorem to students.

13.5 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?

In Section 13.2, 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 13.2.9 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 13.3.6 shows how the line integral $\int_C \vec{F} \cdot d\vec{r}$ can be efficently calculated in terms of the vector field and the derivative of *any* paramaterization of C. Philosophically, the Riemann sum in Definition 13.2.9 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 efficent way to calculate this accumulation?

Preview Activity 13.5.1 In order to pay for tutition, 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.

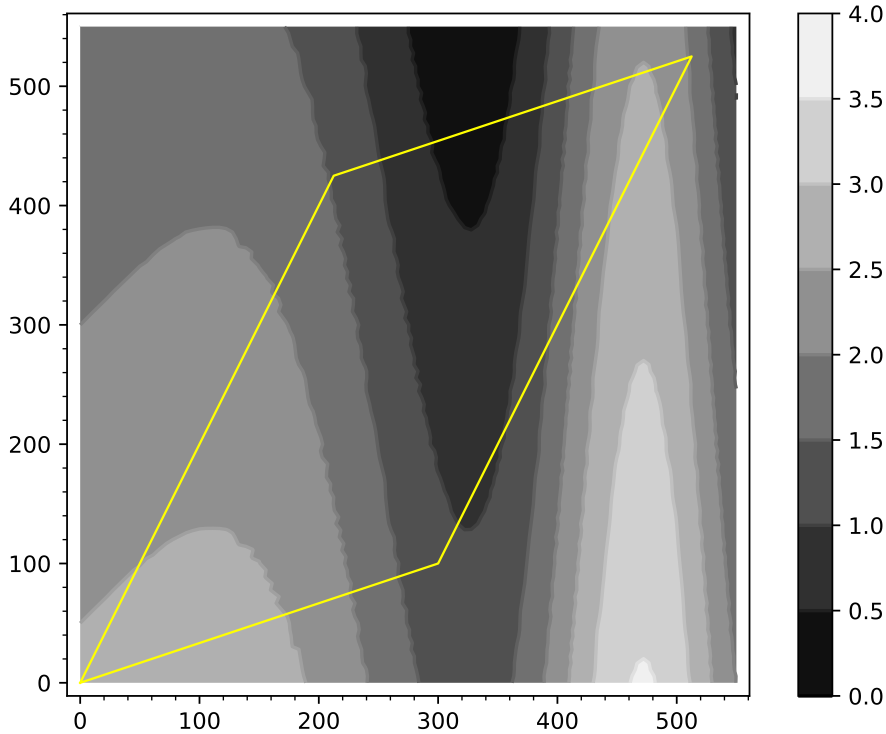


Figure 13.5.1 A plot of land with density of copium deposits

(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 below. 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.

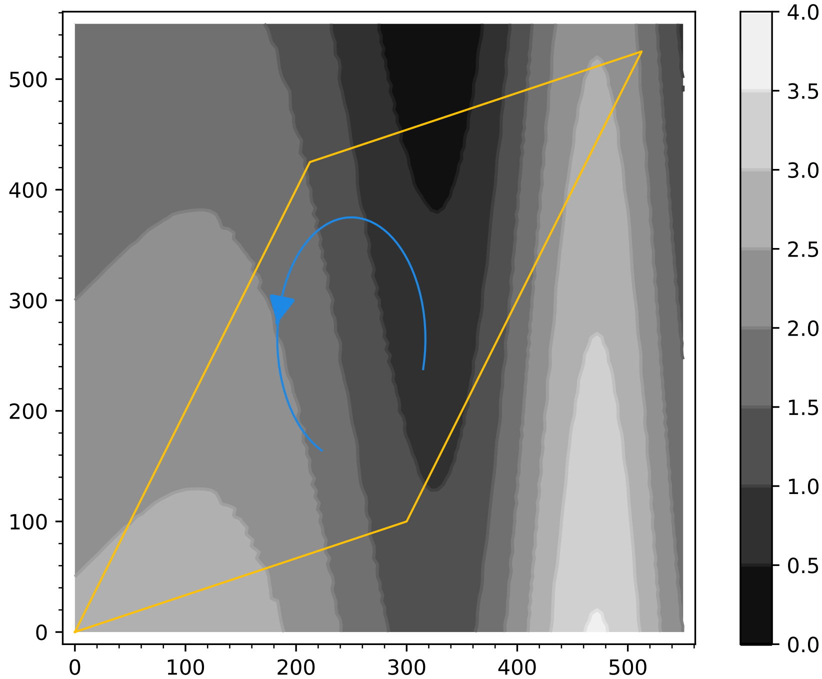


Figure 13.5.2 A plot of land with copium Density and a path plotted in blue

(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.

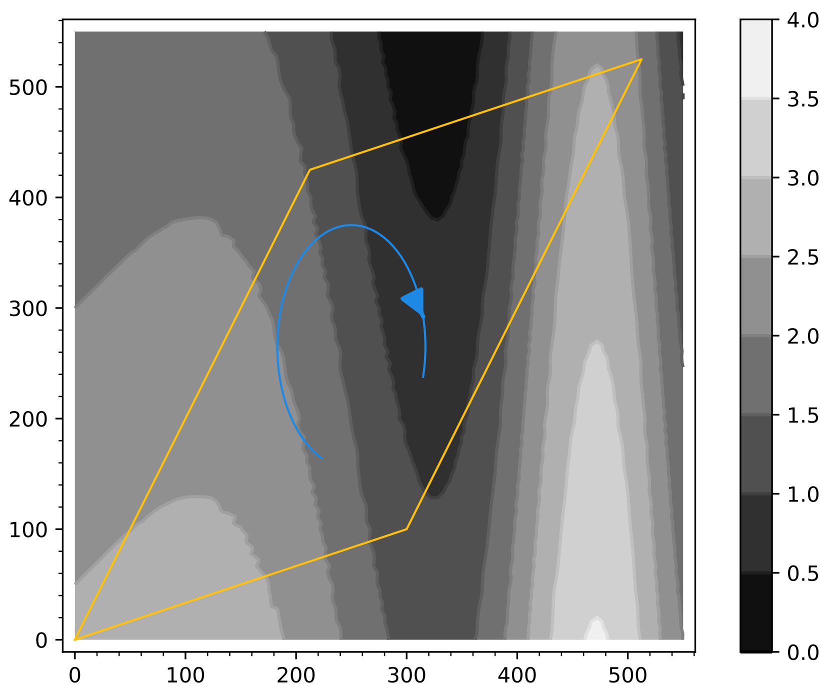


Figure 13.5.3 A plot of land with copium density and a path plotted in blue

13.5.1 Defining line integrals of scalar functions

In Preview Activity 13.5.1, 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 acculumation of a scalar function's output over a path in space.

Let f be a contiunuous 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 acculumlation 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 $\vec{r}_i$

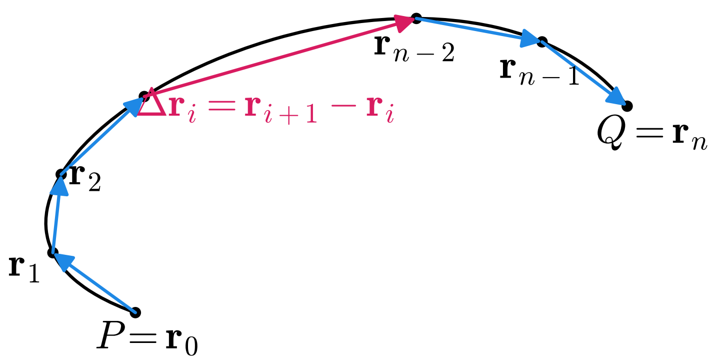


Figure 13.5.4 A curve in space with segments given by $\mathbf{r}_i$

We can approximate the accumulation of f over C with the following sum

$$\sum_{i=1}^{n} f(\vec{r}_i^{*})|\vec{r}_i - \vec{r}_{i-1}|,$$

where $f(r_i^*)$ is the output of f for some $\vec{r}_i^* \in C_i$. 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 acculumation 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 choosen 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 13.5.5 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 contiunuous function of x, y, and z for some open set around C. The **line integral of f over C** is defined as

$$\int_C f ds = \lim_{n\to\infty} \sum_{i=1}^{n} f(\vec{r}_i^{*})|\vec{r}_i - \vec{r}_{i-1}|$$

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 functiton of multiple variables. The differential ds may seem unusual to you. If you remember from Section 10.3, 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 13.5.6 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 $-1 \le x \le 2$. We also will define the following three scalar-valued functions on $\mathbb{R}^2$.

- $f(x, y) = x$
- $g(x, y) = -y$
- $h(x, y) = y - x^2$

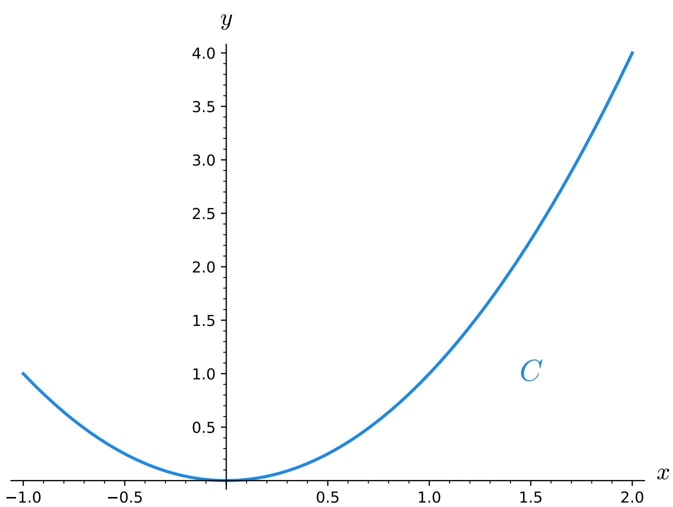


Figure 13.5.7 The curve C used in the line integral for Example 13.5.6

(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 13.5.8 shows that the blue branch of the curve C (for $0 \le x \le 2$) will yield positive output values for f, while the magenta part, which has $-1 \le x \le 0$, 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.

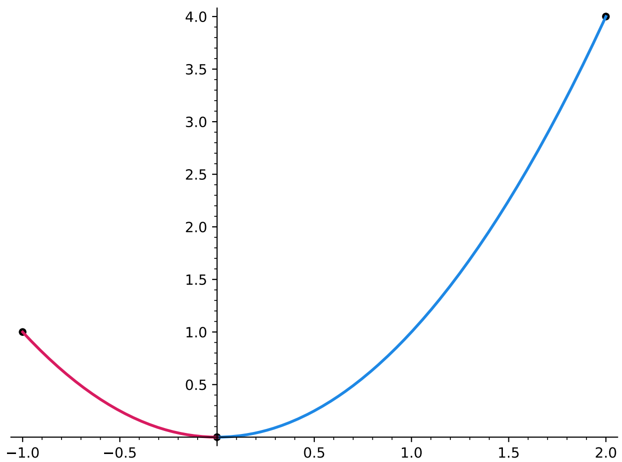


Figure 13.5.8 The curve C split into regions with positive (blue) and negative (magenta) outputs of f

If we break our curve C into three pieces by dividing at $x = 0$ and $x = 1$ as shown in Figure 13.5.9, we can make a rigourous argument about why the result of $\int_C f\,ds$ will be positive. We will call the magenta, green, and blue sections of Figure 13.5.9 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 13.5.5 we can do the symmetric version on curve C_2. The only difference between the Riemann sums and thier 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

$$\begin{aligned}\int_C f\,ds &= \int_{C_1} f\,ds + \int_{C_2} f\,ds + \int_{C_3} f\,ds\\ &= -\int_{C_2} f\,ds + \int_{C_2} f\,ds + \int_{C_3} f\,ds\\ &= \int_{C_3} f\,ds.\end{aligned}$$

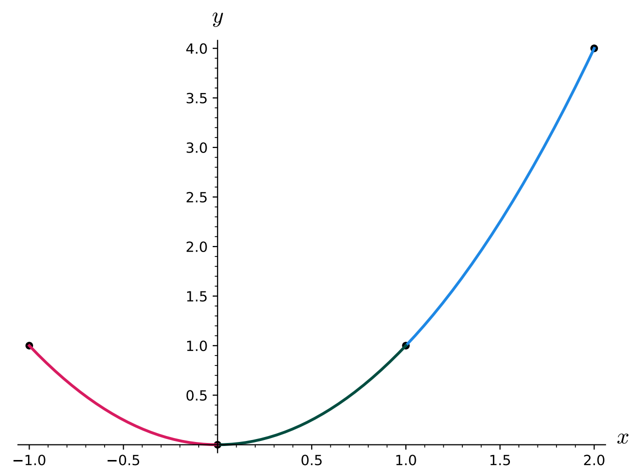


Figure 13.5.9 The curve C split into three regions with positive (blue and green) and negative (magenta) outputs of f

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 13.5.3 you will see examples where having only one part of this symmetry is not sufficent 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 13.5.5 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 (execept 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 Reimann 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 13.5.6.

Activity 13.5.2 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.

$$\int_{C_1} f_4 \, ds$$
$$\int_{C_2} f_4 \, ds$$
$$\int_{C_3} f_4 \, ds,$$

13.5.2 Using Parameterizations to Calculate Scalar Line Integrals

Definition 13.5.5 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 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.

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 13.5.5. 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.

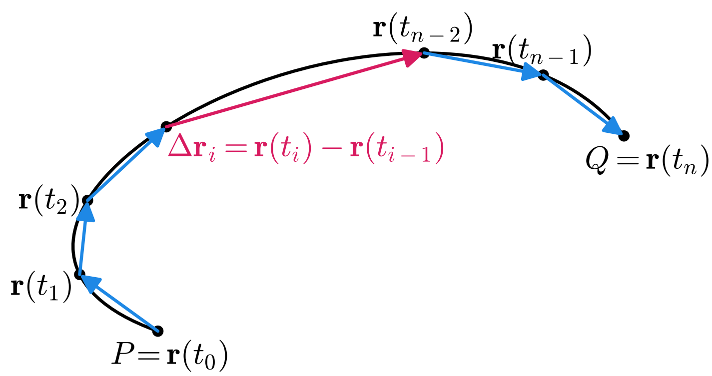


Figure 13.5.10 The curve C split into segments defined by equally spaced parameter values

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 13.5.5 gives

$$\int_C f\,ds = \lim_{n\to\infty}\sum_{i=1}^{n} f(\vec{r}_i^*)|\vec{r}_i - \vec{r}_{i-1}|$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i^*) \frac{|r(t_{i-1}+\Delta t) - r(t_{i-1})|}{\Delta t} \Delta t.$$

This Riemann sum correponds 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 $v(t)$ is the speed of the parameterization $\vec{r}$.

Theorem 13.5.11 Calculating Scalar Line Integrals with Parameterizations. *Let C be a curve in space parameterized by $r(t)$ for $a \leq t \leq b$. 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 13.5.11 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 13.5.2.

Example 13.5.12 Let us return to some of the problems from Activity 13.5.2. 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 $-\frac{\pi}{4} \leq t \leq \frac{3\pi}{4}$. 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 13.5.11, 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 13.5.2.b. Specifically, we can break the line integral of f_1 along C_1 into three parts according to the plot shown in Figure 13.5.13.

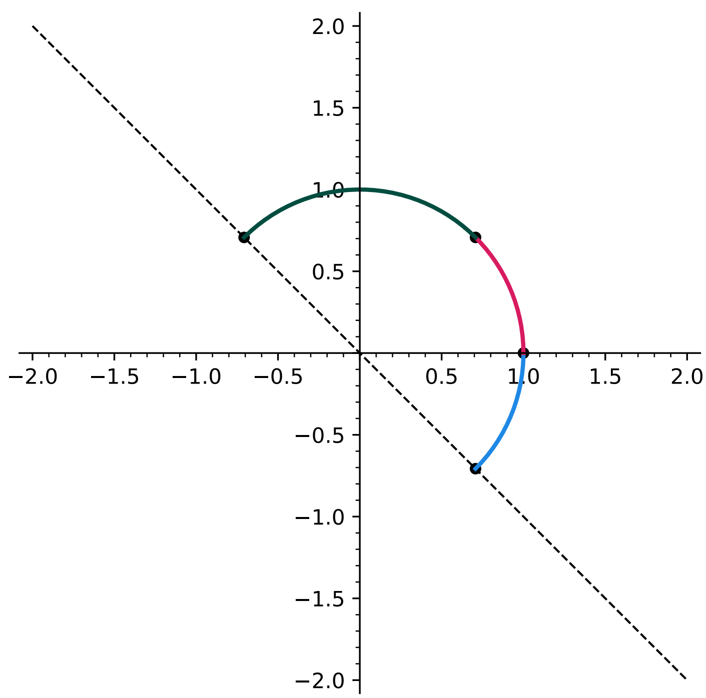


Figure 13.5.13 A subdivided plot of C_1

The scalar line integal 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 13.5.2, 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 13.5.14 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 $-\frac{\pi}{4} \le t \le \frac{3\pi}{4}]$. 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 13.5.11, we have

$$\begin{aligned}\int_{C_1} f_4\, ds &= \int_{-\frac{\pi}{4}}^{\frac{3\pi}{4}} (\cos(t) - \sin(t))\,(1) dt \\ &= \sin(\frac{3\pi}{4}) - \cos(\frac{3\pi}{4}) - \Big(\sin(-\frac{\pi}{4}) - \cos(-\frac{\pi}{4})\Big) = 0\end{aligned}$$

This result should make sense from your earlier work in part 13.5.2.b. Specifically, we can break the line integral of f_4 into two parts according to the plot in Figure 13.5.15.

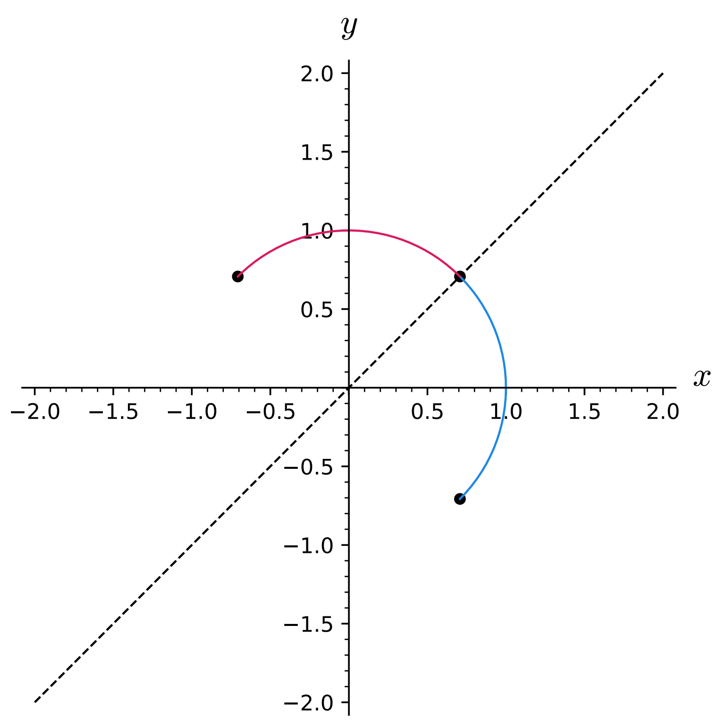


Figure 13.5.15 A subdivided plot of C_1

The scalar line integal 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 13.5.3 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 13.5.16. We will consider the function $f(x, y) = x$ for this activity.

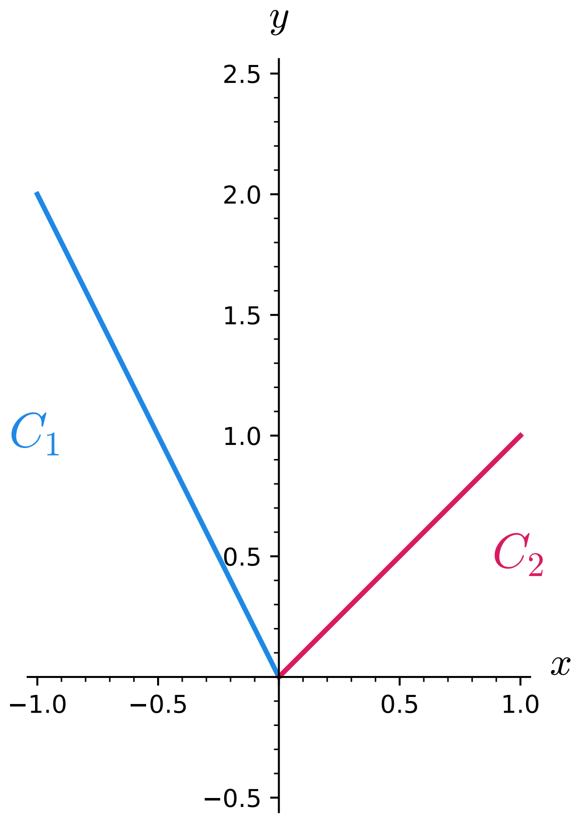


Figure 13.5.16 A plot of paths C_1 and C_2

(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 13.5.11 to compute $\displaystyle\int_{C_1} f\, ds$ and $\displaystyle\int_{C_2} f\, ds$.

(c) As with line integrals of vector fields, we have that $\displaystyle\int_{C_1+C_2} f\, ds = \int_{C_1} f\, ds + \int_{C_2} f\, ds$. Use this property to compute $\displaystyle\int_{C_1+C_2} f\, ds$.

In Activity 13.5.3, 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 cancelation argument as in part 13.5.6.a.

13.5.3 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 bewteen 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 13.5.4 will have you make sense of these properties for scalar line integrals.

Activity 13.5.4 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 13.5.1. Remember that the curve in our scalar line integral coresponds 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.

(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.

13.5.4 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 13.5.1. In particular, we can look the left side of the mining area.

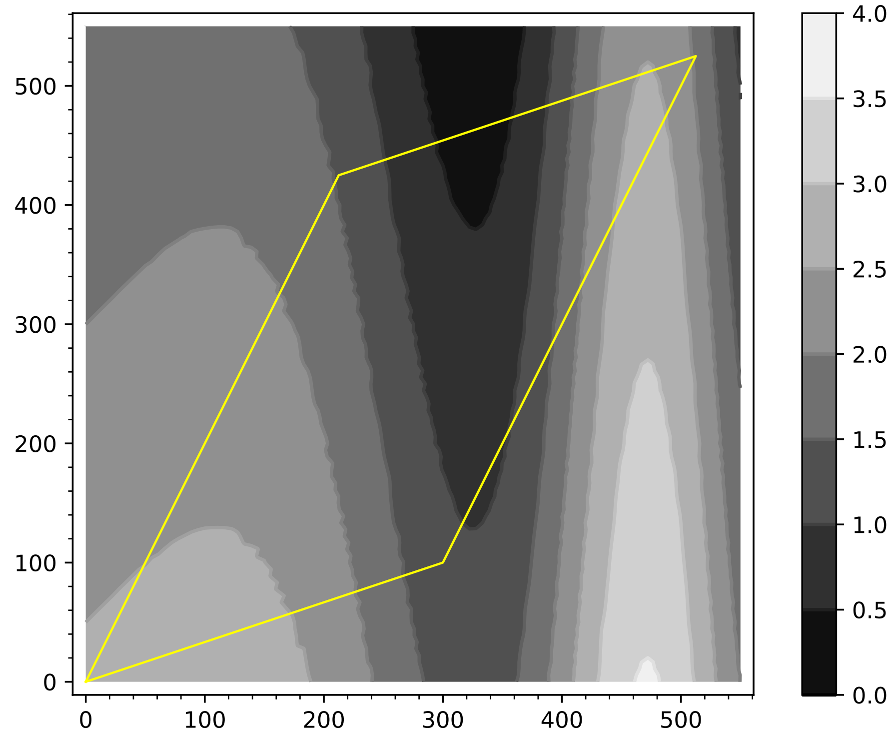


Figure 13.5.17 A plot of land with density of copium deposits

We could visualize the linear density of copium along the left side of the area using a plot like Figure 13.5.18.

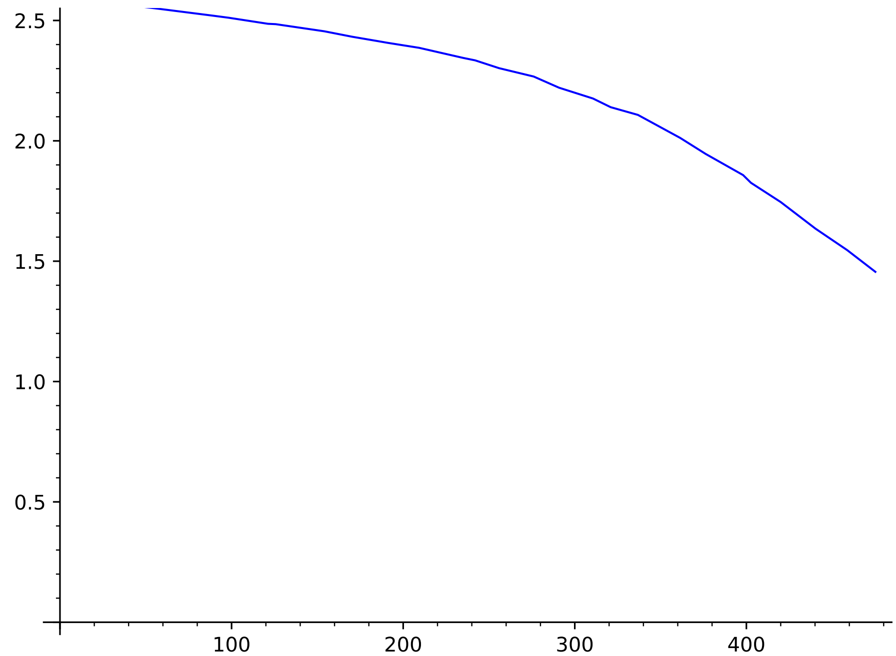
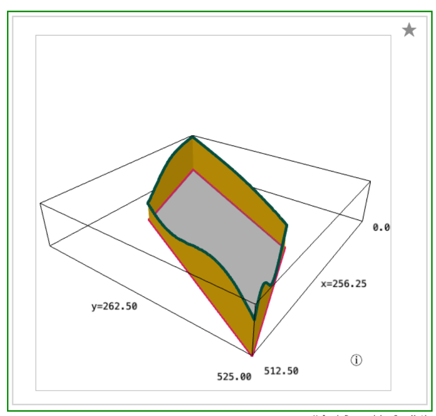


Figure 13.5.18 A plot of the copium density on the left side of the mine shown above

In Figure 13.5.18, the horizontal axis gives the distance traveled along the left side of Figure 13.5.1. 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.

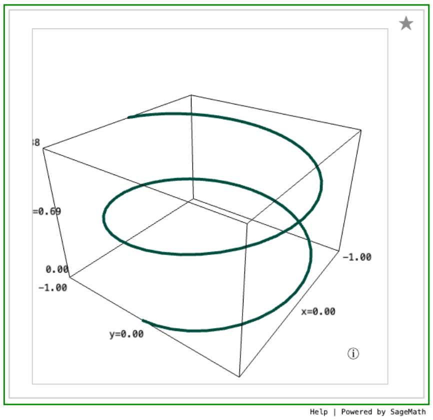


Standalone
Embed

Figure 13.5.19 A three-dimensional plot of the copium Density plotted for the edges of the mining area

Figure 13.5.19 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.

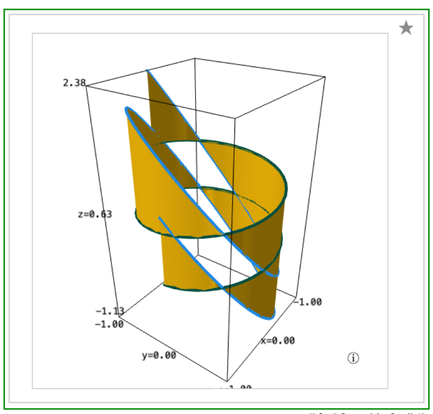
Because the curve we are looking at in Figure 13.5.19 involves straight lines and simple heights, there is no confusion when looking at this plot and using an area under the curve analogy. However, what if we looked at the scalar line integral of a function like $f(x, y, z) = x - y$ along the helix given below?



Standalone
Embed

Figure 13.5.20 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 13.5.19, 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
Embed

Figure 13.5.21 A three-dimensional plot of a helix with height given by $f(x, y, z) = x - y$

13.5.5 Summary

- The scalar line integral, denoted by $\int_C f\,ds$, measures the acculumation 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.

13.5.6 Exercises

1. Find the line integral with respect to arc length $\int_C (9x+2y)ds$, where C is the line segment in the xy-plane with endpoints $P=(5,0)$ and $Q=(0,3)$.

(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

$$\int_C (9x+2y)ds = \int_a^b \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \, dt$$

with limits of integration $a =$ __________ and $b =$ __________

(c) Evaluate the line integral with respect to arc length in part (b).

$\int_C (9x+2y)ds =$ ______________________

2. If C is the part of the circle $\left(\frac{x}{4}\right)^2 + \left(\frac{y}{4}\right)^2 = 1$ in the first quadrant, find the following line integral with respect to arc length.

$\int_C (9x-7y)ds =$ ______________

3. 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 $1 \le x \le 2$.

4. 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 $\int_{C_3} f\,ds = 9$.

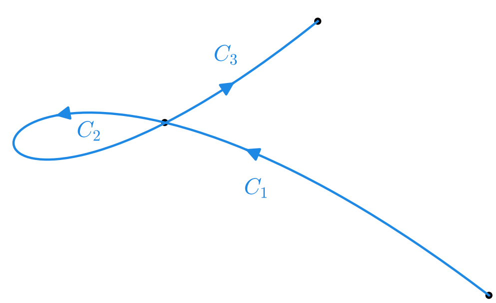


Figure 13.5.22 An oriented path broken into three parts

Find the following:

(a) $\displaystyle\int_{-C_3} f\,ds$

(b) $\displaystyle\int_{C_2} f\,ds$

(c) $\displaystyle\int_{-C_1-C_3} f\,ds$

5. Calculate the following line integral where C is the path on $x = y$ with $-1 \geq y \geq 2$:

$$\int_C (y^2 - 2x)\, ds$$

13.6 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?

As we saw in Section 13.1, 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 13.6.1, 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.

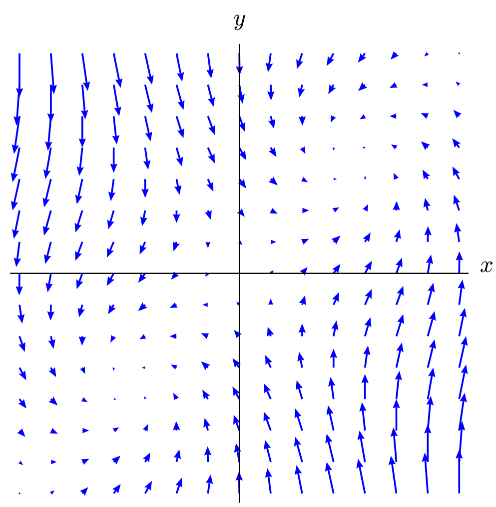


Figure 13.6.1 A vector field with changing strength

Preview Activity 13.6.1 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.

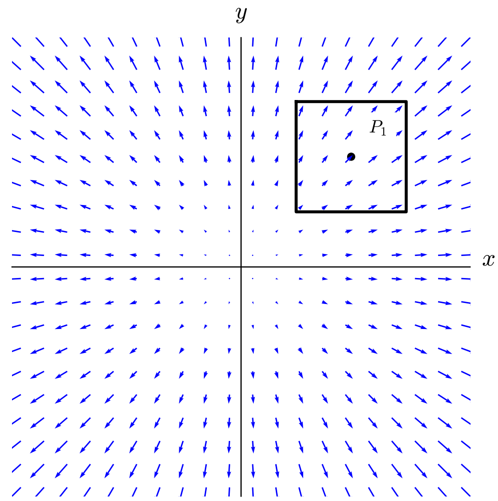


Figure 13.6.2 Vector Field $\vec{F}$

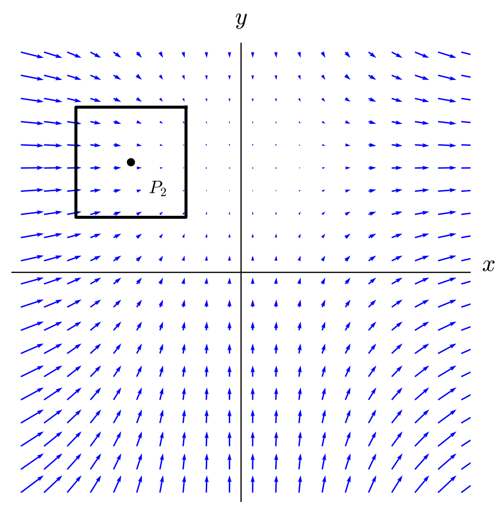


Figure 13.6.3 Vector Field $\vec{G}$

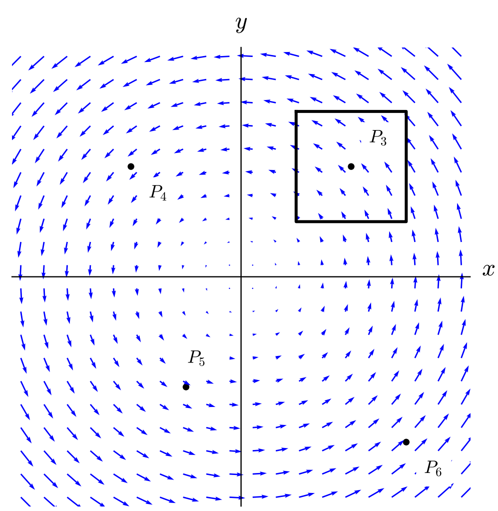


Figure 13.6.4 Vector Field $\vec{H}$

(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.

13.6.1 Definition of the Divergence of a Vector Field

We begin this subsection by stating a definition that captures analytically the ideas you reasoned aboug geometrically in Preview Activity 13.6.1. After the statement of the definition, we discuss what it means.

Definition 13.6.5 The **divergence of a vector field**

$$\vec{F}(x,y) = \langle F_1(x,y), F_2(x,y), F_3(x,y)\rangle$$

is given by

$$\operatorname{div}(\vec{F}) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}.$$

In two dimensions, the definition is analogous by omiting the third term. ◊

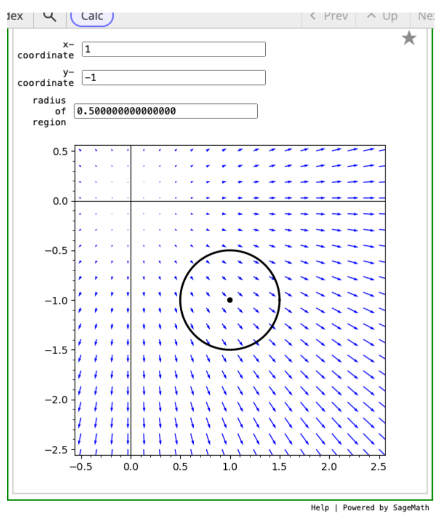
Alternative Notation for Divergence. In other sources you may see the divergence written using a dot product as $\operatorname{div}(\vec{F}) = \nabla \cdot \vec{F}$. This notation is very compact and works well with the understanding that the del operator $\nabla = \langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\rangle$ 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 13.6.2 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 13.6.1). When looking at a small neighborhood of the point (as shown by the square in Figure 13.6.15), 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 13.6.6 Let $\vec{F} = \langle x, y \rangle$. Then

$$\operatorname{div}(\vec{F}) = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) = 1 + 1 = 2.$$

From our conceptual description above, $\operatorname{div}(\vec{F})$ 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
Embed

Figure 13.6.7 An interactive plot of $\vec{F} = \langle x, y \rangle$ with point (a, b) and region plotted

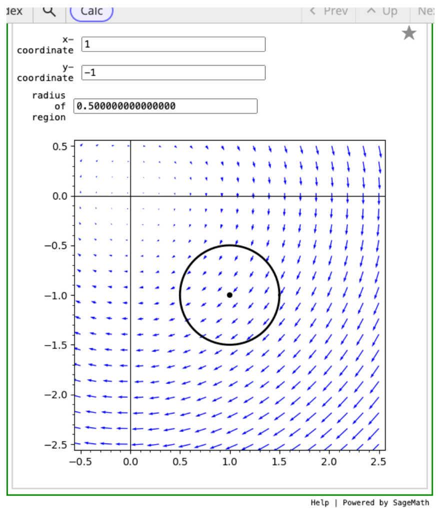
□

Example 13.6.8 Let $\vec{G} = \langle y, -x \rangle$. We can compute that

$$\operatorname{div}(\vec{G}) = \frac{\partial}{\partial x}(y) + \frac{\partial}{\partial y}(-x) = 0.$$

From our conceptual description of divergence, $\operatorname{div}(\vec{G})$ 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
Embed

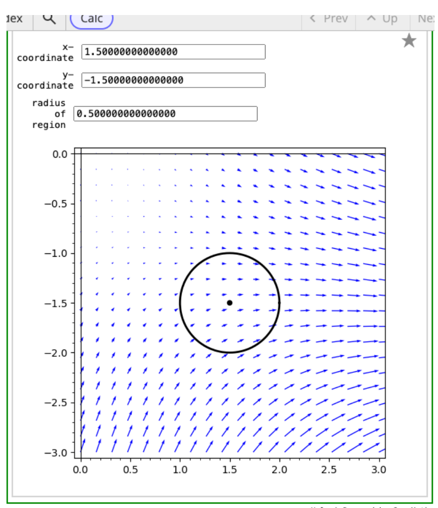
Figure 13.6.9 An interactive plot of $\vec{G} = \langle y, -x\rangle$ with point (a, b) and region plotted

□

Example 13.6.10 Let $\vec{H} = \langle x^2 - y, y^2 - x\rangle$, which means that

$$\operatorname{div}(\vec{H}) = \frac{\partial}{\partial x}(x^2 - y) + \frac{\partial}{\partial y}(y^2 - x) = 2x + 2y.$$

In constrast to the previous examples, you can see that the value of divergence of $\vec{H}$ will depend on the input point choosen. For instance, $\operatorname{div}(\vec{H})$ 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
Embed

Figure 13.6.11 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 13.6.11, 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 13.6.2 Graphical Representations of Divergence.

(a) For this part of the activity, consider the vector field $\vec{F}$ shown in Figure 13.6.12.

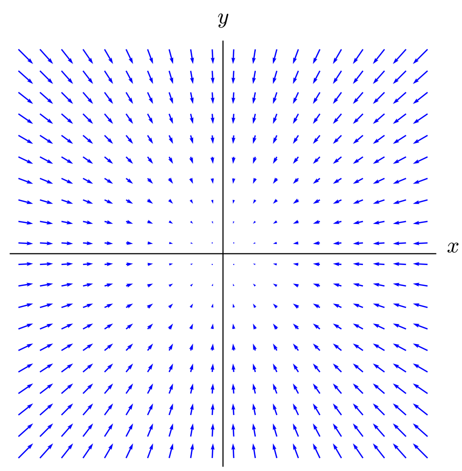


Figure 13.6.12 Vector field $\vec{F}$

(i) Draw a circle in the first quadrant of the vector field $\vec{F}$ depicted in Figure 13.6.12. 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 13.6.12, 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 13.6.12, 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 13.6.13 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.

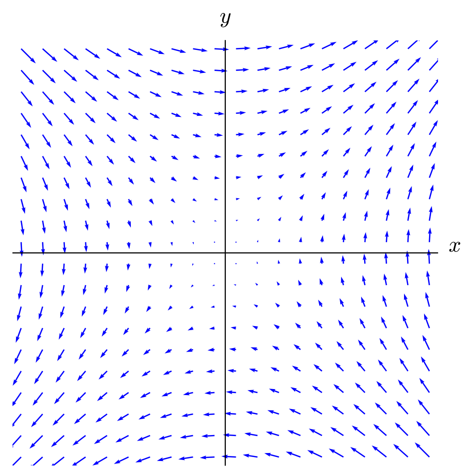


Figure 13.6.13 Vector field $\vec{G}$

(c) Look at the plot of the vector field $\vec{H}$ in Figure 13.6.14 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.

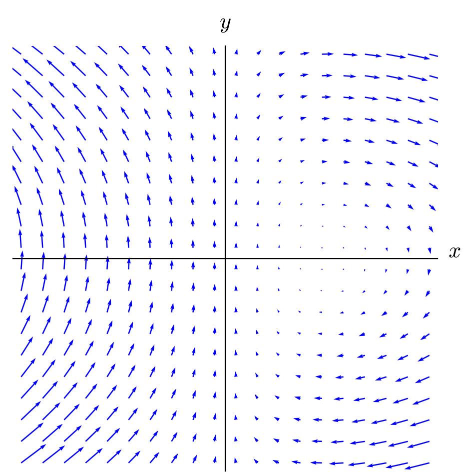


Figure 13.6.14 Vector field $\vec{H}$

The next activity of this section asks you to do some algebraic calculations of divergence using Definition 13.6.5.

Activity 13.6.3

(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 13.6.2 can be explained by using your results from part a of this activity.

13.6.2 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 in a classic calculus fashion. 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.

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 13.6.1. 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 $\vec{F}(x, y) = \langle F_1(x, y), F_2(x, y) \rangle$.

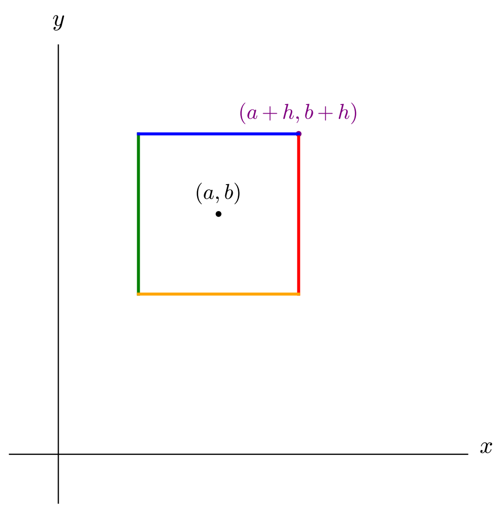


Figure 13.6.15 A square around the point (a, b)

We can parametrize the top edge of the box by $\vec{r}_{\text{top}}(t) = \langle a + t, b + h \rangle$ with $-h \leq t \leq h$. 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 $-h \leq t \leq h$.

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

$$\begin{aligned} N = &\int_{-h}^{h} F_2(a+t, b+h)dt - \int_{-h}^{h} F_2(a+t, b-h)dt \\ &+ \int_{-h}^{h} F_1(a+h, b+t)dt - \int_{-h}^{h} F_1(a-h, b+t)dt. \end{aligned}$$

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

$$\begin{aligned} \frac{1}{(2h)^2} \Bigg(&\int_{-h}^{h} F_2(a+t, b+h)dt - \int_{-h}^{h} F_2(a+t, b-h)dt \\ &+ \int_{-h}^{h} F_1(a+h, b+t)dt - \int_{-h}^{h} F_1(a-h, b+t)dt \Bigg) \end{aligned}$$

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[1] 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

$$\int_{-h}^{h} F_2(a+t, b+h)dt = (2h) \;\; F_2(t_1^*, b+h),$$

[1] activecalculus.org/single/sec-4-3-definite-integral.html#XCM

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.

$$\text{Flow Density}(a,b) = \lim_{h\to 0} \frac{\text{net flow}}{\text{area}}$$
$$= \lim_{h\to 0} \frac{(2h)\left(F_2(t_1^*, b+h) - F_2(t_2^*, b-h) + F_1(a+h, t_3^*) - F_1(a-h, t_1^*)\right)}{4h^2}.$$

To simplify the limit further, we will reorganize our limit expression. Specifically, we collect the F_2 and F_1 terms separately. This gives

$$\text{Flow Density}(a,b) =$$
$$\lim_{h\to 0}\left[\frac{F_1(a+h, t_3^*) - F_1(a-h, t_1^*)}{2h} + \frac{F_2(t_1^*, b+h) - F_2(t_2^*, b-h)}{2h}\right].$$

Recall the central difference method of estimating derivatives from Section 1.5.2[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 13.9.

13.6.3 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
$$\text{div}(\vec{G}) = \frac{\partial G_1}{\partial x} + \frac{\partial G_2}{\partial y} + \frac{\partial G_3}{\partial z}$$

[2] activecalculus.org/single/sec-1-5-units.html#VjW

- 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.

13.6.4 Exercises

1. Find the divergence of each of the following vector fields at all points where they are defined.

(a) $\operatorname{div}\left(\left(4x^2 - \sin(xz)\right)\boldsymbol{i} + 4\boldsymbol{j} - \sin(xz)\,\boldsymbol{k} \right) =$ ______________________

(b) $\operatorname{div}\left(4e^{xy}\boldsymbol{i} + 3\cos(xy)\,\boldsymbol{j} + 2e^{\ln\left(x^2+y^2+4\right)}\boldsymbol{k} \right) =$ ______________________

(c) $\operatorname{div}\left(\frac{x}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{i} + \frac{y}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{j} + \frac{z}{(x^2+y^2+z^2)^{1.5}}\boldsymbol{k} \right) =$ ______________________

2. Consider $\operatorname{div}\left(\frac{-y\vec{i}+x\vec{j}}{(x^2+y^2)^4}\right)$.

(a) Is this a vector or a scalar? (☐ vector ☐ scalar)
(b) Calculate it:
$\operatorname{div}\left(\frac{-y\vec{i}+x\vec{j}}{(x^2+y^2)^4}\right) =$ ______________ + ______________ = ______________

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.

(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 $\operatorname{div}(\vec{G}) = \vec{0}$.

(b) Vector fields with a zero divergence everywhere in their domain are called **divergence-free** 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?

13.6.5 Notes to Instructors and Dependencies

This section relies heavily on understanding vector fields from Section 13.1. We have separated the details of how the flux density of a region leads to the definition and understanding of the divergence (Subsection 13.6.2). The second subsection is optional, but is worthwhile reading for students who are interested in a good conceptual understanding of how divergence is developed.

13.7 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?

In Section 13.6, 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 13.7.1 shows a vector field which flows in a counterclockwise fashion around the origin. The vector field on the right of Figure 13.7.1 shows a vector field which does not have a rotational aspect to its flow.

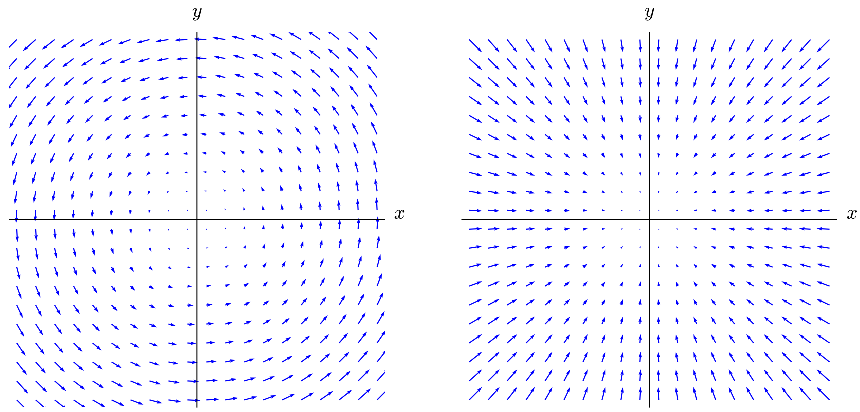


Figure 13.7.1 Examples of vector fields with global rotation properties (left) and without rotation in either the global or local sense (right)

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 13.7.2 seems to have rotation around at least a couple of different points. In each quadrant, the vector field on the right of Figure 13.7.2 has different rotational patterns.

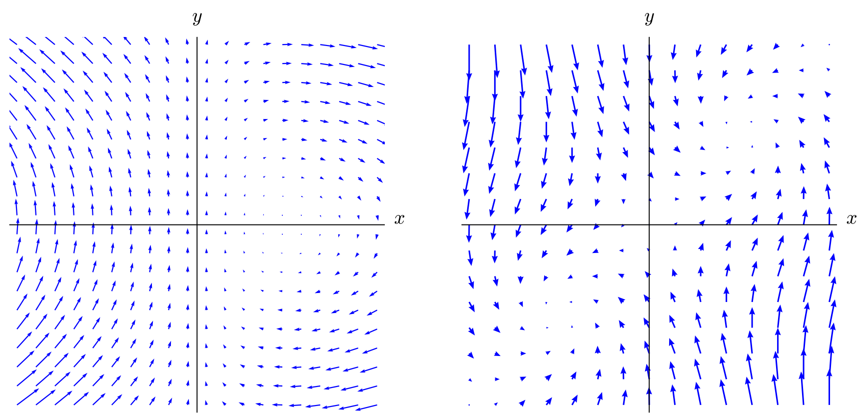
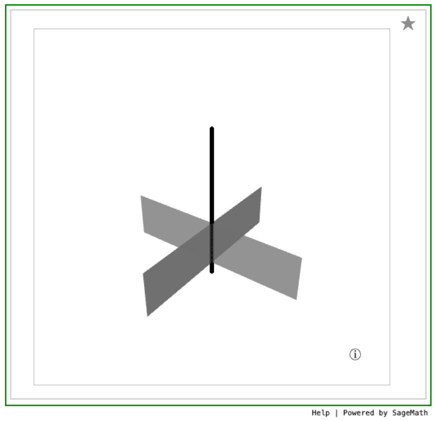


Figure 13.7.2 Examples of vector fields with varying rotation properties

Preview Activity 13.7.1 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 13.7.3. 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
Embed

Figure 13.7.3 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 13.7.4. 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.

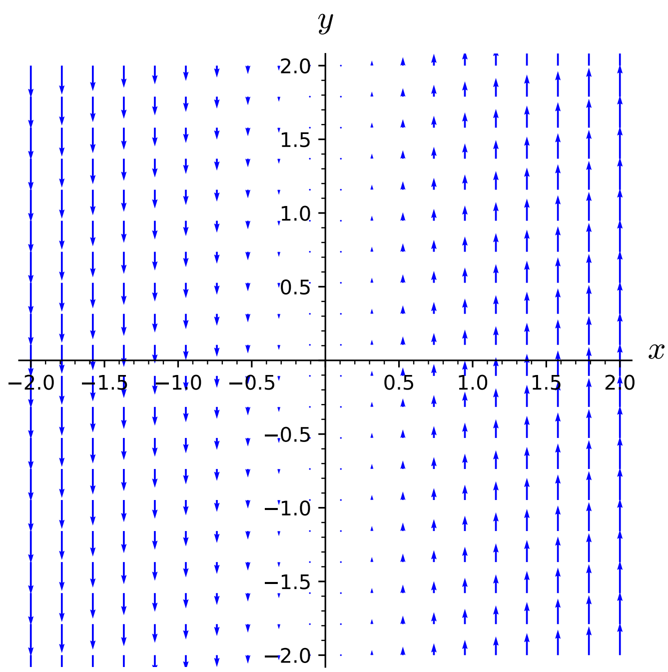


Figure 13.7.4 The vector field $\vec{F}$

(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?

13.7.1 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 13.6.2, 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 13.7.5.

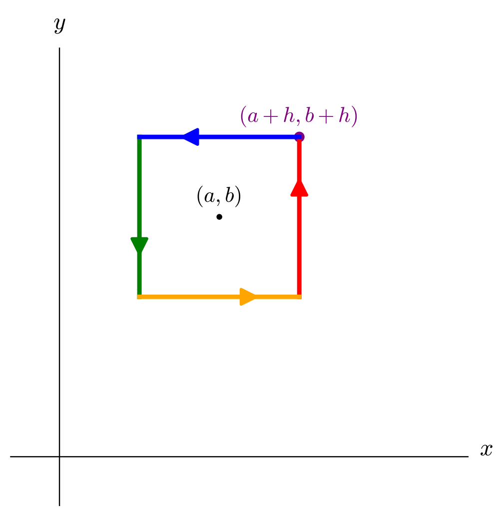


Figure 13.7.5 A square centered at the point (a, b) with side lengths $2h$ traveled counterclockwise

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 13.3.6. 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 13.7.2.

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

$$\begin{aligned} \text{Total Circulation} &= \int_{\text{Square}} \vec{F} \cdot d\vec{r} \\ &= \int_{-h}^{h} -F_1(a - t, b + h) \, dt + \int_{-h}^{h} -F_2(a - h, b - t) \, dt \end{aligned}$$

$$
\begin{aligned}
&+\int_{-h}^{h} F_1(a+t,b-h)\,dt+\int_{-h}^{h} F_2(a+h,b+t)\,dt\\
=&-2hF_1(t^*_{\text{top}},b+h)-2hF_2(a-h,t^*_{\text{left}})\\
&+2hF_1(t^*_{\text{bottom}},b-h)+2hF_2(a+h,t^*_{\text{right}})
\end{aligned}
$$

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[1] 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 13.6. 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.

[1] activecalculus.org/single/sec-1-5-units.html#VjW

Activity 13.7.2 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:

1. What is the formula for the circulation density of the vector field?
2. For what points will you have a positive circulation density?
3. For what points will you have a negative circulation density?
4. For what points will you have a zero circulation density?

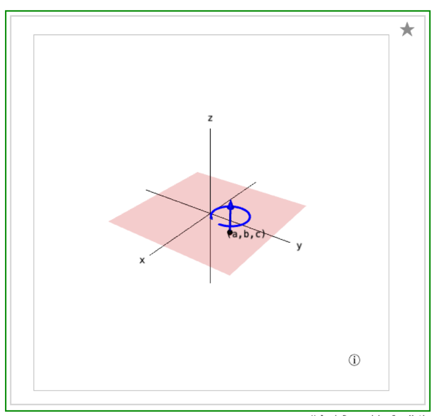
(a) $\langle -y, x \rangle$

(b) $\langle x, y \rangle$

(c) $\langle 1 - x, xy \rangle$

13.7.2 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.



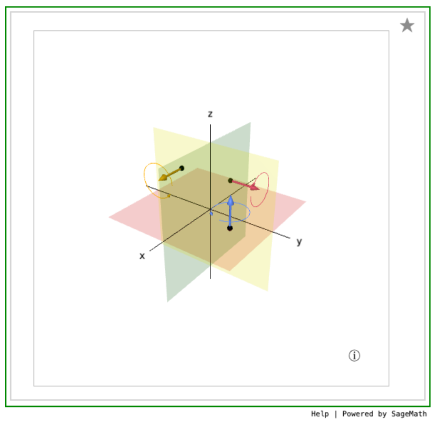
Standalone
Embed

Figure 13.7.6 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 13.7.6. 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 13.7.7, you can see that each vector corresponds to the rotation displayed and is consistent with the conventions described above.



Standalone
Embed

Figure 13.7.7 Rotation as a vector on each coordinate plane

If you look carefully at Figure 13.7.7, 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[2] (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.[3] 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:

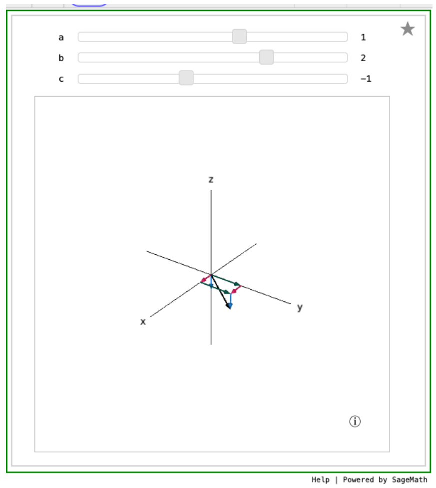
$$\hat{\imath} \times \hat{\jmath} = \hat{k} \qquad \hat{\imath} \times \hat{k} = -\hat{\jmath} \qquad \hat{\jmath} \times \hat{k} = \hat{\imath}$$

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 $a\hat{\imath} + b\hat{\jmath} + c\hat{k}$.

[2]When placing the x-axis horizontally for the blue vector and placing the y-axis horizontally for the yellow vector

[3]When placing the x-axis horizontally.



Standalone
Embed

Figure 13.7.8 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 13.7.8, 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 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.

13.7.3 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 13.7.9 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

$$\operatorname{curl}(\vec{F}) = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\hat{\imath} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z}\right)\hat{\jmath} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\hat{k}.$$

$\Diamond$

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 $\vec{F}$, $\operatorname{curl}(\vec{F}) = \nabla \times \vec{F}$.

The curl is exactly the rotational description described at the end of Subsection 13.7.2. 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 13.7.10 In this example, we will look at $\vec{F} = \langle 0, x, 0\rangle$. Applying Definition 13.7.9,

$$\operatorname{curl}(\vec{F}) = \left\langle \frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(x), -\left(\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}(0)\right), \frac{\partial}{\partial x}(x) - \frac{\partial}{\partial y}(0) \right\rangle$$

Thus, $\operatorname{curl}(\vec{F}) = \langle 0, 0, 1\rangle$. This means that $\vec{F}$ will have a rotational aspect only with an axis parallel to the z-axis. This should not be suprising since the plot of $\vec{F}$ will be the same as Figure 13.7.4 copied on each plane $z = c$. As you saw in Preview Activity 13.7.1, 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 13.7.11 In this example, we will look at $\vec{G} = \langle x, y, z\rangle$. If we apply Definition 13.7.9, we obtain

$$\operatorname{curl}(\vec{G}) = \left\langle \frac{\partial}{\partial y}(z) - \frac{\partial}{\partial z}(y), -\left(\frac{\partial}{\partial x}(z) - \frac{\partial}{\partial z}(x)\right), \frac{\partial}{\partial x}(y) - \frac{\partial}{\partial y}(x) \right\rangle$$

Therefore, we have $\operatorname{curl}(\vec{G}) = \langle 0, 0, 0\rangle$. Thus, $\vec{G}$ will have no rotational strength anywhere. In Figure 13.7.12 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.

Help | Powered by SageMath

Standalone
Embed

Figure 13.7.12 The black vector, $\langle a, b, c\rangle$, is decomposed into parts parallel to $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

□

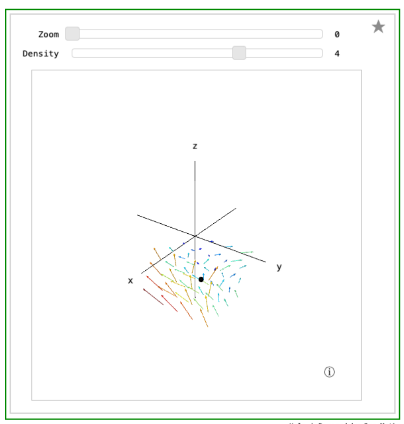
Example 13.7.13 In this example, we will look at $\vec{H} = \langle x-y, y+2z, x^2 \rangle$. By Definition 13.7.9,

$$\operatorname{curl}(\vec{H}) = \left\langle \frac{\partial}{\partial y}(x^2) - \frac{\partial}{\partial z}(y+2z), -\left(\frac{\partial}{\partial x}(x^2) - \frac{\partial}{\partial z}(x-y)\right), \frac{\partial}{\partial x}(y+2z) - \frac{\partial}{\partial y}(x-y) \right\rangle$$

Therefore, $\operatorname{curl}(\vec{H}) = \langle 0, 2x, -1 \rangle$. 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 13.7.3 Estimating Curl in Three Dimensions.

(a) Consider the vector field $\vec{F}$ plotted in Figure 13.7.14. 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
Embed

Figure 13.7.14 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 13.7.15. 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 $z = -2$.

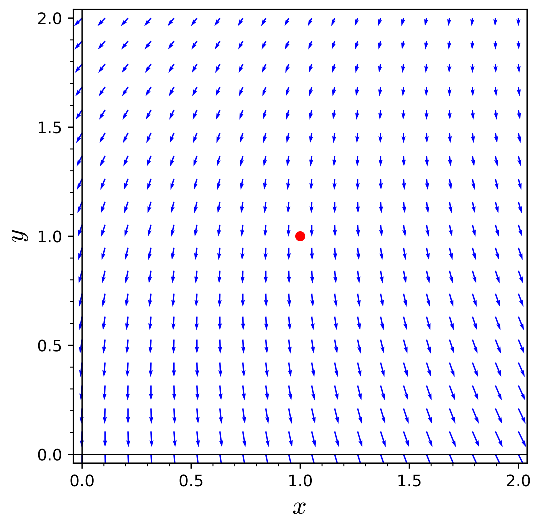


Figure 13.7.15 The trace of $\vec{F}$ in the plane $z = -2$

(c) Next we will look at the vector field in the plane $y = 1$, as displayed in Figure 13.7.16. 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 13.7.15?

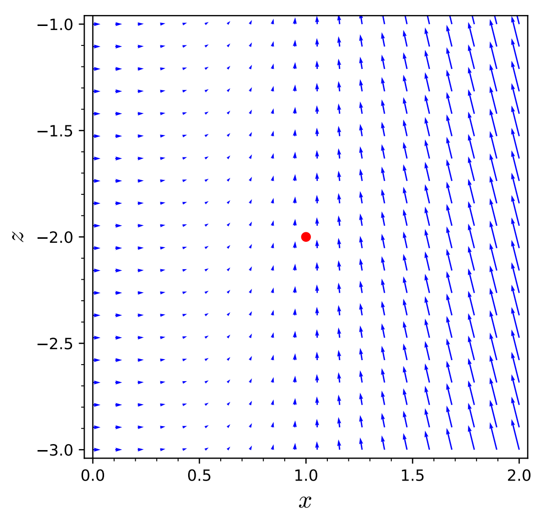


Figure 13.7.16 The trace of $\vec{F}$ in the plane $y = 1$

(d) Finally, consider the trace of $\vec{F}$ in the plane $x = 1$, as displayed in Figure 13.7.17. 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 13.7.15 and Figure 13.7.15?

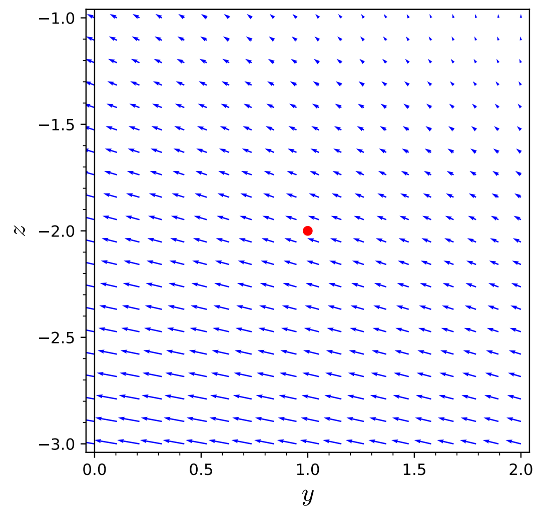


Figure 13.7.17 The trace of $\vec{F}$ in the plane $x = 1$

(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 $\text{curl}(\vec{F})$ 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.

13.7.4 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 13.7.18 *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 $\left(\text{curl}(\vec{F})(a, b, c)\right) \cdot \vec{v}$.*

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 13.7.4 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 13.7.19. 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 $\text{curl}(\langle x, y, z\rangle) = \vec{0}$.

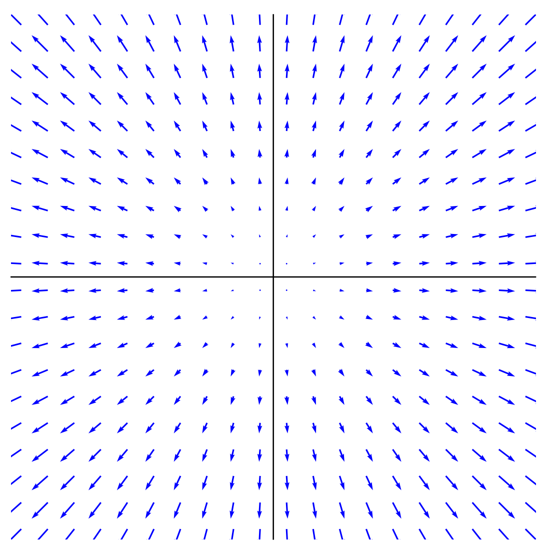
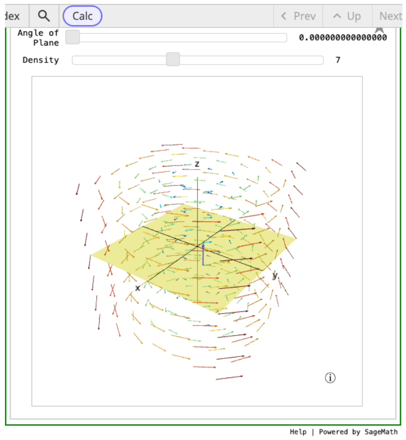


Figure 13.7.19 The vector field $\langle x, y, z\rangle$ on every plane through the origin

(b) Now consider the vector field $\vec{F} = \langle -y, x, 0\rangle$, which is shown in Figure 13.7.20. 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
Embed

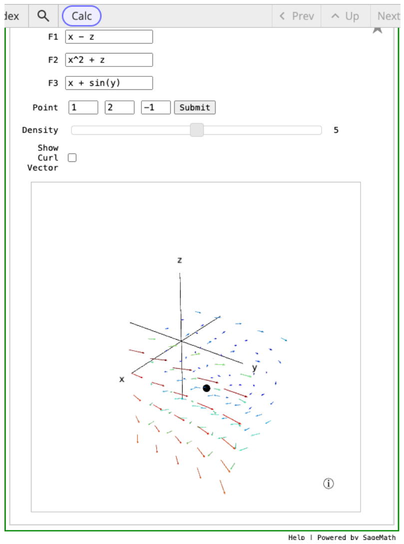
Figure 13.7.20 Two views of the vector field $\langle -y, x, 0\rangle$

(i) Based on the figure, would you estimate the components of $\text{curl}(\vec{F})$ to be positive, negative, or zero at the origin? Does your answer change if you pick a different point?

(ii) Calculate $\text{curl}(\vec{F})$ 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 13.7.21 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 13.7.21 will show the curl vector at the base point specified so you can make sense of your vector field and its curl.



Standalone
Embed

Figure 13.7.21 A plot of the vector field $\langle F_1, F_2, F_3\rangle$

Use the figure to estimate the direction of $\text{curl}(\langle x - z, x^2 + z, x + \sin(y)\rangle)$ 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.

13.7.5 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 $\text{curl}(\vec{F})(a, b, c)$ 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 $\left(\text{curl}(\vec{F})(a, b, c)\right) \cdot \vec{v}$

13.7.6 Exercises

1. Compute the curl of the vector field $\vec{F} = 8z\,\tilde{i} + 8y\,\tilde{j} + 2x\,\tilde{k}$.
curl = ______________________________

2. Compute the curl of the vector field $\vec{F} = \langle x^2, y^2, z^3 \rangle$.

curl$(\vec{F}(x, y, z))$ = ________________

What is the curl at the point $(-2, -5, 0)$?

curl$(\vec{F}(-2, -5, 0))$ = ________________

Is this vector field irrotational or not? (☐ Choose ☐ irrotational ☐ not irrotational ☐ cannot be determined)

3. A)

Consider the vector field $F(x, y, z) = (-5yz, 6xz, 6xy)$.

Find the divergence and curl of F.

div$(F) = \nabla \cdot F$ = ________________.

curl$(F) = \nabla \times F$ = (________________, ________________, ________________).

B)

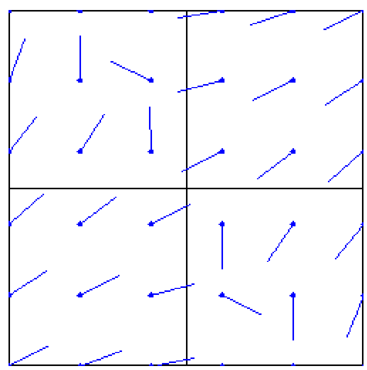
Consider the vector field $F(x, y, z) = (-9x^2, 9(x+y)^2, -9(x+y+z)^2)$.

Find the divergence and curl of F.

div$(F) = \nabla \cdot F$ = ________________.

curl$(F) = \nabla \times F$ = (________________, ________________, ________________).

4. 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)

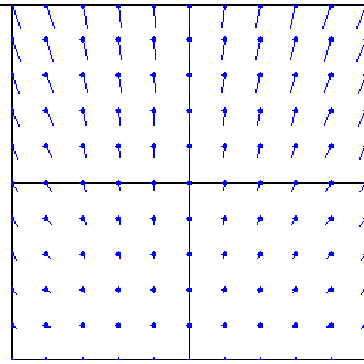
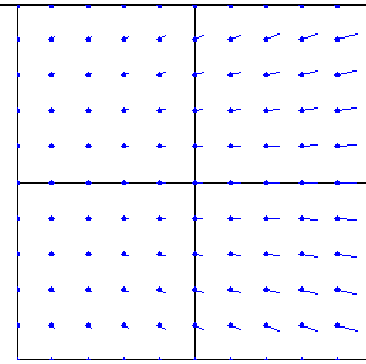
(c) If we put a paddle wheel at the origin, instead of a twig, in what direction would you expect that to rotate?

(☐ clockwise ☐ counterclockwise ☐ it won't rotate)

(d) Compute curl $\vec{F}$.

curl$\vec{F}$ = ________________

5. For each of the vector fields below, decide if they have a nonzero or zero curl at the origin. Each vector field is shown in the xy-plane; they have no z-component and are independent of z. Note that the vector fields are shown with a dot at the *tail* of each vector.

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

Vector field ***(a)*** has (☐ nonzero ☐ zero) curl at the origin.
Vector field ***(b)*** has (☐ nonzero ☐ zero) curl at the origin.
Vector field ***(c)*** has (☐ nonzero ☐ zero) curl at the origin.

6. For each of the following vector fields, find its curl and determine if it is a gradient field.

(a) $\vec{F} = \left(6xy + 4x^3\right)\,\tilde{i} + \left(3x^2 + z^2\right)\,\tilde{j} + (2yz - z)\,\tilde{k}$:
curl $\vec{F} =$ ____________________
$\vec{F}$ (☐ is a gradient field ☐ is not a gradient field)

(b) $\vec{G} = (3xy + yz)\,\tilde{i} + \left(4x^2 + z^2\right)\,\tilde{j} + xz\,\tilde{k}$:
curl $\vec{G} =$ ____________________
$\vec{G}$ (☐ is a gradient field ☐ is not a gradient field)

(c) $\vec{H} = 3yz\,\tilde{i} + \left(z^2 - 3xz\right)\,\tilde{j} + (3xy + 2yz)\,\tilde{k}$:
curl $\vec{H} =$ ____________________
$\vec{H}$ (☐ is a gradient field ☐ is not a gradient field)

7. Let $\mathbf{F} = (7xy, 5y, 6z)$.
The curl of $\mathbf{F} =$ (__________, __________, __________).
Is there a function f such that $\mathbf{F} = \nabla f$? ______ (y/n)

8. Three small squares, S_1, S_2, and S_3, each with side 0.1 and centered at the point $(1, 8, 6)$, 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 5, around S_2 of -0.01, and around S_3 of -2. Estimate curl $\vec{G}$ at the point $(1, 8, 6)$
curl $\vec{G} \approx$ ____________________

13.7.7 Notes to Instructors and Dependencies

This section relies heavily on understanding vector fields from Section 13.1. If the calculation of the curl is your primary purpose for using this section, then you can skip many of the details of Subsection 13.7.1 and Subsection 13.7.2. The details have been made as bite-sized as possible, but the particulars of curl typically aren't easily understood without some of the calculations of circulation density. Additionally, the visualization of vector fields in 3D makes geometric interpretations harder to demonstrate.

13.8 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?

We know from Section 13.4 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 13.8.1 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. 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 13.7 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 $\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$ 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.

13.8.1 Circulation

In Preview Activity 13.8.1, 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 13.8.1. 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.

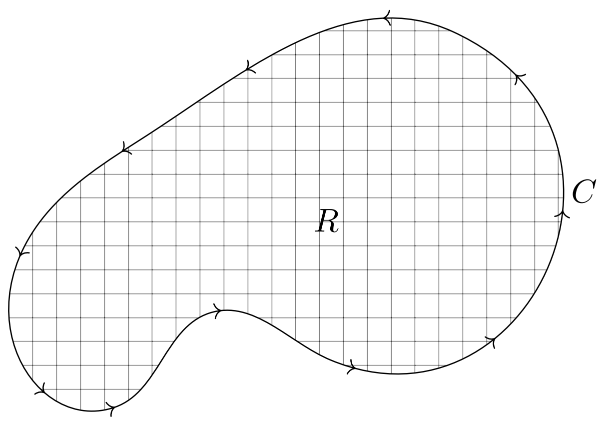


Figure 13.8.1 A region bounded by a curve

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 13.8.2 and think of these two square regions as being two of the square regions inside R in Figure 13.8.1.

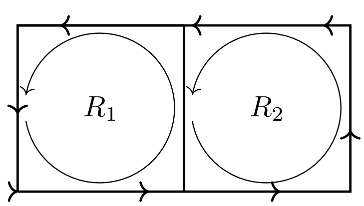


Figure 13.8.2 A square region divided into two smaller squares

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 13.8.1, 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 subsititute 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

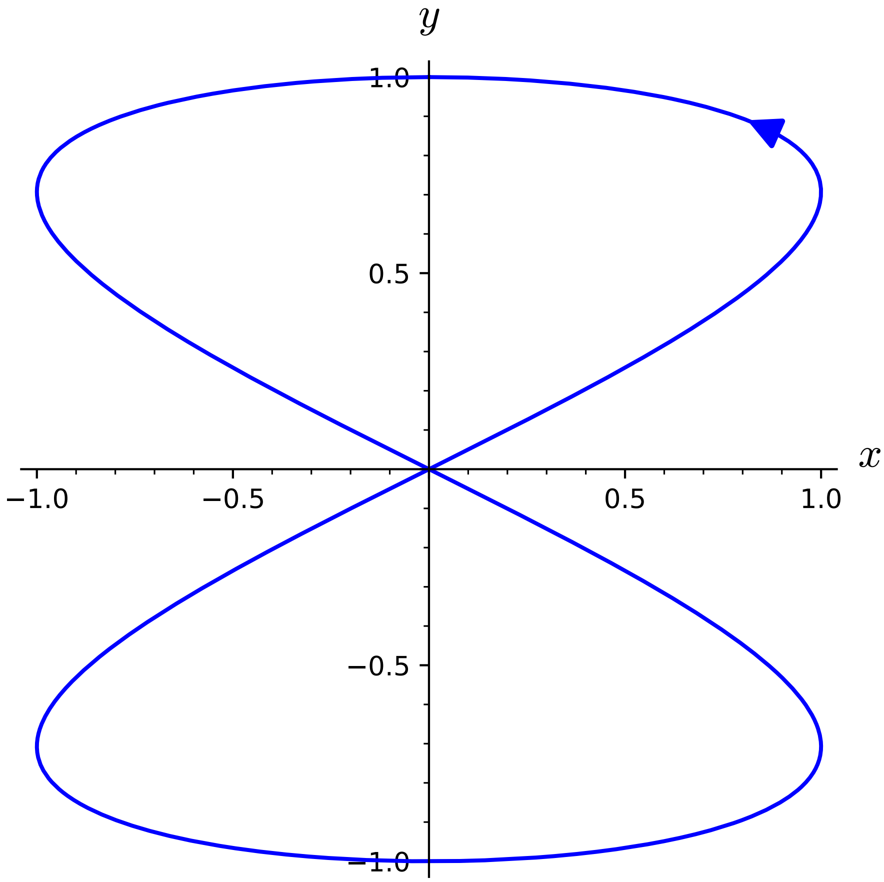
$$\begin{aligned}\oint_C \vec{F} \cdot d\vec{r} &= \lim_{\Delta R_i \to 0} \sum_{R_i} \oint_{C_i} \vec{F} \cdot d\vec{r} \\ &= \lim_{\Delta R_i \to 0} \sum_{R_i} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \Delta R_i.\end{aligned}$$

This expression should look familiar as the Reimann 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).

13.8.2 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 13.8.3 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 parametrizations for several pieces of the curve, such as with the rectangular curve pictured in Figure 13.8.4. 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 13.8.4 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.

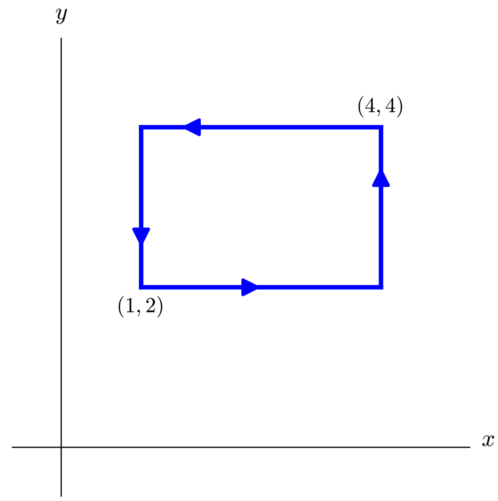


Figure 13.8.4 A rectangular curve

Example 13.8.5 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 13.8.4.

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 itegral 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 segements 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 paramterization uses $0 \le t \le 1$. These parameterizations will yeild 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 13.8.2 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.

13.8.3 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 7, 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 13.8.3 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 13.8.6, 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 $(0, 0)$.

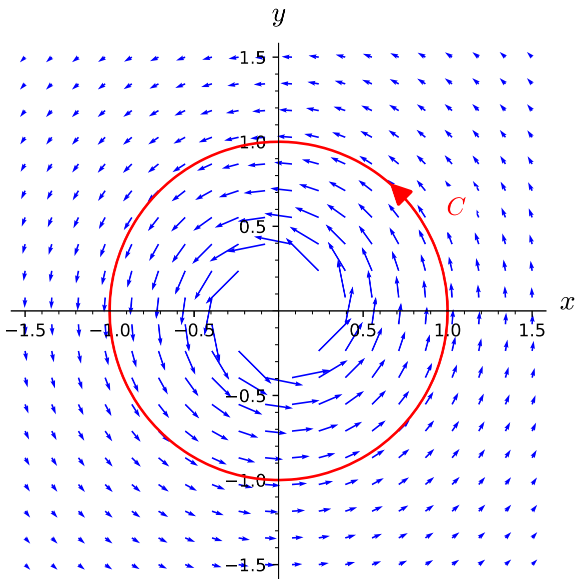


Figure 13.8.6 The vector field $\vec{F}$

(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 13.8.7, path-independent? Why or why not?

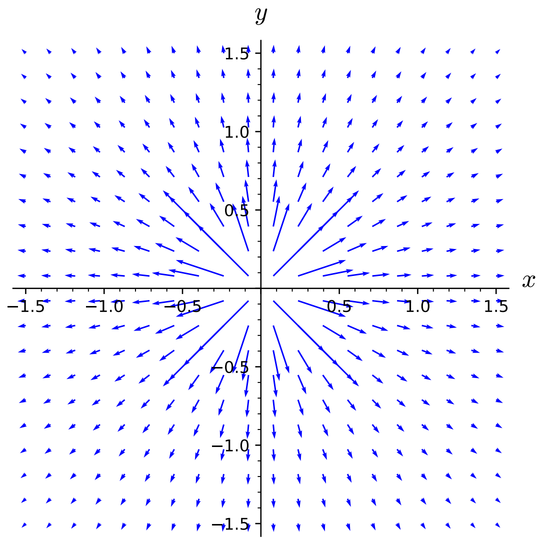


Figure 13.8.7 The vector field $\vec{G}$

(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.

13.8.4 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.

13.8.5 Exercises

1. Calculate $\int_C \left(8(x^2-y)\vec{i}+7(y^2+x)\vec{j}\right) \cdot d\vec{r}$ if:

(a) C is the circle $(x-3)^2+(y-4)^2=25$ oriented counterclockwise.

$\int_C \left(8(x^2-y)\vec{i}+7(y^2+x)\vec{j}\right) \cdot d\vec{r} =$ ____________________

(b) C is the circle $(x-a)^2+(y-b)^2=R^2$ in the xy-plane oriented counterclockwise.

$\int_C \left(8(x^2-y)\vec{i}+7(y^2+x)\vec{j}\right) \cdot d\vec{r} =$ ____________________

2. Let $\vec{F} = 4xe^y\,\vec{i} + 2x^2e^y\,\vec{j}$ and $\vec{G} = 4(x-y)\,\vec{i} + 2(x+y)\,\vec{j}$. Let C be the path consisting of lines from $(0,0)$ to $(8,0)$ to $(8,3)$ to $(0,0)$. Find each of the following integrals exactly:

(a) $\int_C \vec{F} \cdot d\vec{r} =$ ____________________

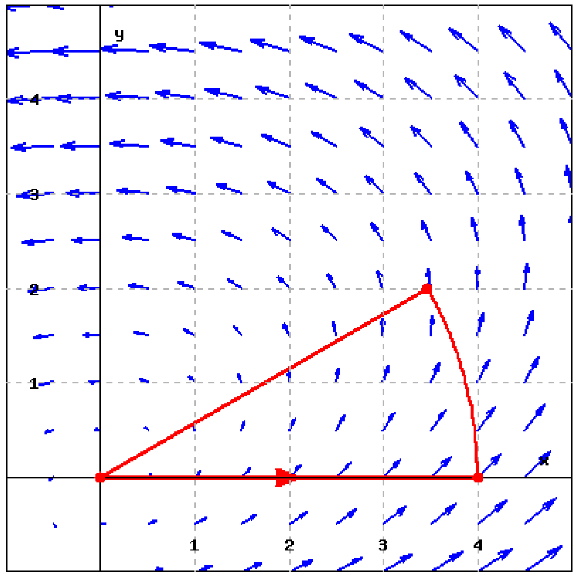
(b) $\int_C \vec{G} \cdot d\vec{r} =$ ______________________________

3. Suppose

$$\vec{F}(x, y) = (2x - 4y)\vec{i} + 2x\vec{j}$$

and C is the counter-clockwise oriented sector of a circle centered at the origin with radius 4 and central angle $\pi/6$. Use Green's theorem to calculate the circulation of $\vec{F}$ around C.

Circulation = ______________________________



(Click on graph to enlarge)

4. Let $\vec{F} = (9x^2y + 3y^3 + 6e^x)\,\vec{i} + (5e^{y^2} + 81x)\,\vec{j}$. 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.)

5. ***(a)*** Show that each of the vector fields $\vec{F} = 2y\,\vec{i} + 2x\,\vec{j}$, $\vec{G} = \frac{2y}{x^2+y^2}\,\vec{i} + \frac{-2x}{x^2+y^2}\,\vec{j}$, and $\vec{H} = \frac{x}{\sqrt{x^2+y^2}}\,\vec{i} + \frac{y}{\sqrt{x^2+y^2}}\,\vec{j}$ 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.)

6. 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 $0 \leq t \leq 2\pi$. 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?

7. 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.

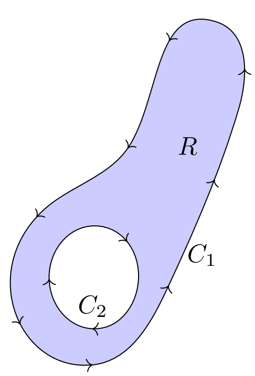


Figure 13.8.8 A region that is not simply connected

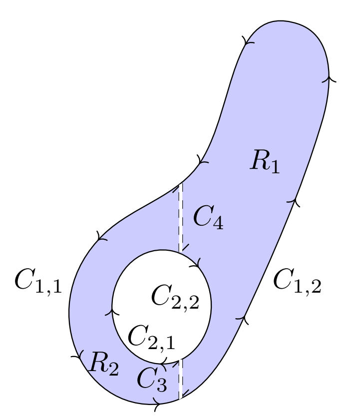


Figure 13.8.9 Breaking up a region

(a) Explain why the region R in Figure 13.8.8 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 13.8.9 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 13.8.3, 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 13.8.8 using the curve C in place of C_1.

8. 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 12.4 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 $a,b>0$.

13.8.6 Notes to Instructors and Dependencies

This section relies on the calculation of circulation around a closed curve from Section 13.2 and Section 13.3 as well as the idea of circulation density that was described in Section 13.6.

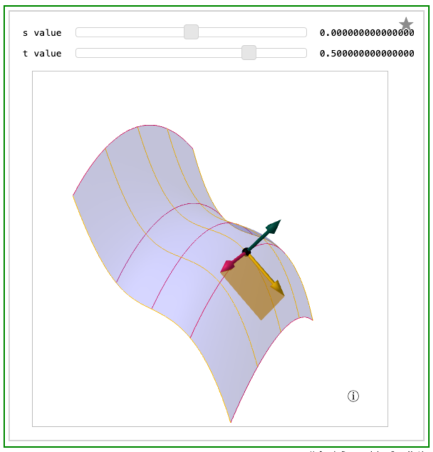
13.9 Flux Integrals

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 $z = f(x,y)$?

Section 12.6 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 12.6 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 12.6.2, we set up a Riemann sum based on a parameterization that would measure the surface area of our curved surfaces in space.

In Figure 13.9.1, 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
Embed

Figure 13.9.1 A three-dimensional plot of Figure 12.6.6

From Section 9.4, 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 13.9.1 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 12.6 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:

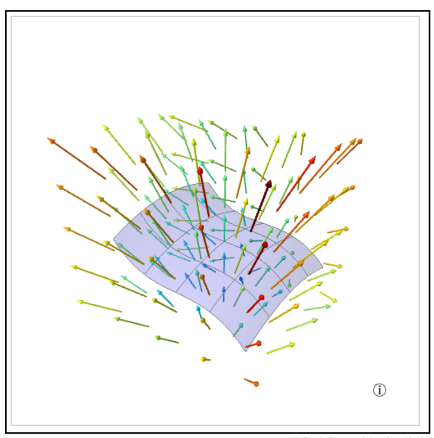
(i) S_2: the sphere centered at the origin of radius 3

(ii) S_3: the first octant portion of the plane $x + 2y + 3z = 6$

As we saw in Section 12.6, we can set up a Riemann sum of the areas for the parallelograms in Figure 13.9.1 to approximate the surface area of the region plotted by our parametrization. Equation (12.6.2) 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.

13.9.1 The Idea of the Flux of a Vector Field through a Surface

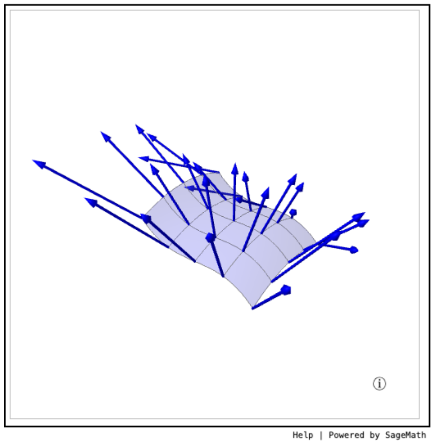
In this section we will look at how to measure the amount of a vector field that flows through a surface in space. Figure 13.9.2 illustrates a plot that demonstrates this idea. Our definition of divergence in Section 13.6 looked at measuring the amount of vector field flowing out of a small region on 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
Embed

Figure 13.9.2 A three-dimensional vector field and a surface

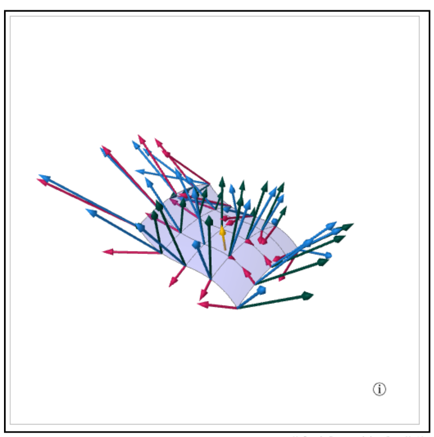
As with understanding line integrals of vector-valued functions in Section 13.2, we don't care about the output of the vector field at points away from the surface. We would really would like to examine the output vectors for the points on our surface. To do this, we will look at Figure 13.9.3, which plots the output of our vector field at an array of points on our surface.



Standalone
Embed

Figure 13.9.3 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 13.9.4, 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
Embed

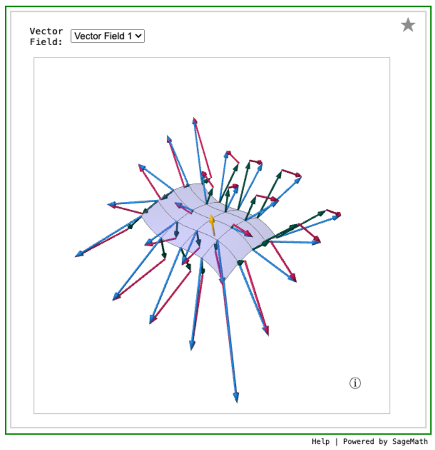
Figure 13.9.4 The decomposition of three-dimensional vector field evaluated along a surface into normal and tangent components

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 13.9.4. 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 13.9.4, 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 13.9.2 Visualizing flux through a surface. In this activity, you will compare the net flow of different vector fields through our sample surface. In Figure 13.9.5 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.



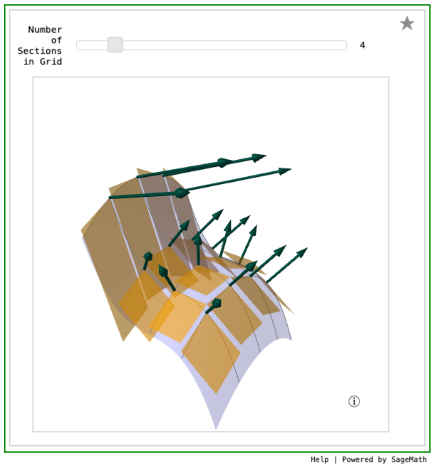
Standalone
Embed

Figure 13.9.5 Five vector fields flowing through a surface, with decomposition into normal and tangential components

13.9.2 The Details of Measuring 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 the corresponding Riemann sum to measure the flux of a vector field through a section of a surface. 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 12.6 (specifically (12.6.1)) 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
Embed

Figure 13.9.6 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

$$\text{Flux} = \sum_{i=1}^{n}\sum_{j=1}^{m} |\vec{F}_{\perp Q_{i,j}}| \cdot S_{i,j},$$

where $|\vec{F}_{\perp Q_{i,j}}|$ 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 13.9.6, 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

$$\vec{F}_{\perp Q_{i,j}} = |\text{proj}_{\vec{w}_{i,j}} \vec{F}(s_i, t_j)| = \frac{\vec{F}(s_i, t_j) \cdot \vec{w}_{i,j}}{|\vec{w}_{i,j}|}$$

Combining these pieces, we find that the flux through $Q_{i,j}$ is approximated by

$$\begin{aligned}\text{Flux through } Q_{i,j} = |\vec{F}_{\perp Q_{i,j}}| \cdot S_{i,j} &= \left(\frac{\vec{F}_{i,j} \cdot \vec{w}_{i,j}}{|\vec{w}_{i,j}|}\right)(|\vec{w}_{i,j}|\Delta s \Delta t)\\ &= \left(\vec{F}_{i,j} \cdot (\vec{r}_s \times \vec{r}_t)\right)(\Delta s \Delta t)\,,\end{aligned}$$

where $\vec{F}_{i,j} = \vec{F}(s_i, t_j)$. Therefore we may approximate the total flux by

$$\text{Total Flux} = \sum_{i=1}^{n}\sum_{j=1}^{m} \left(\vec{F}_{i,j} \cdot \vec{w}_{i,j}\right)(\Delta s \Delta t)\,.$$

Taking the limit as $n, m \to \infty$ gives the following result.

Conceptually, the above limit shows how the argument in Subsection 12.6.2 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.

Theorem 13.9.7 *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*

$$\iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

In Figure 13.9.6, you can change the number of sections in the partition and see the geometric result of refining the partition. In the next example, we will look at how Theorem 13.9.7 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 13.9.8 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 13.9.7 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 13.9.9 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 13.9.7, 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. 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:

$$\begin{aligned}\vec{r}_s &= \langle \cos(t), \sin(t), 1\rangle\\ \vec{r}_t &= \langle -s\sin(t), s\cos(t), 0\rangle\end{aligned}$$

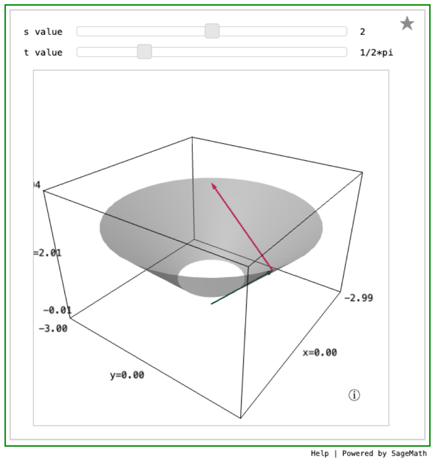
Taking the cross product of $\vec{r}_s$ and $\vec{r}_t$, we obtain the vector-valued function $\vec{w}$:

$$\vec{w} = \vec{r}_s \times \vec{r}_t = \langle -s\cos(t), -s\sin(t), s(\cos(t)^2 + \sin(t)^2)\rangle.$$

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
Embed

Figure 13.9.10 A plot of our surface with the parameterization vector in green and the normal vector $\vec{w}$ in magenta

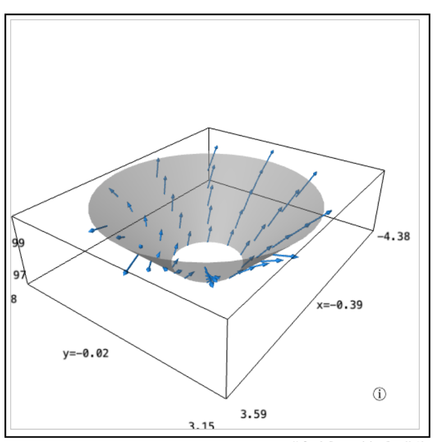
In Figure 13.9.10, the normal vector (as calculated from the parameterization) points inside the cone. This is the direction of positive flow when measuring the flux of the vector field. Our parameterization will also transform our vector field into a function of s and t. Specifically,

$$\vec{F}(s,t) = \langle (s\cos(t))(s) - 2, s\sin(t) + s\cos(t), 1\rangle$$

Applying Theorem 13.9.7 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
Embed

Figure 13.9.11 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 13.9.11 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 13.9.3 Checking the Visualization for Flux.

(a) Figure 13.9.12 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.

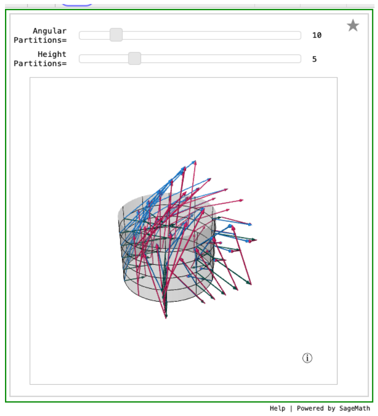
Help | Powered by SageMath

Standalone
Embed

Figure 13.9.12 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 13.9.13 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
Embed

Figure 13.9.13 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 explalin 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 13.9.14 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?

13.9.3 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

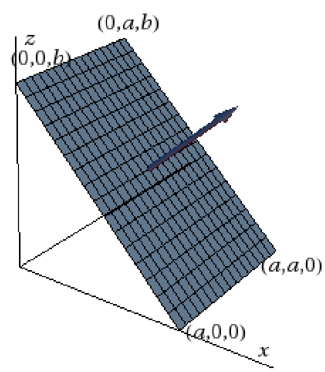
$$\iint_D \vec{F} \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

- 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

$$\iint_D \vec{F}(x, y, f(x, y)) \cdot \left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle dA.$$

13.9.4 Exercises

1. Compute the flux of the vector field $\vec{v} = -\tilde{i} + 5\tilde{j} - \tilde{k}$ through the rectangular region shown below, assuming it is oriented as shown and that $a = 1$ and $b = 4$.



flux = ____________________________________

2. Calculate the flux of the vector field $\vec{F}(x, y, z) = 6y\vec{j}$ through a square of side length 7 in the plane $y = 8$. The square is centered on the y-axis, has sides parallel to the axes, and is oriented in the positive y-direction.

Flux = ____________________

Hint. $y = 8$.

3. (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 $x^2 + y^2 + z^2 = 16$, 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. $z = \rho\cos(\phi)$.

4. Compute the flux of the vector field $\vec{F} = 8x^2y^2z\vec{k}$ 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 __________ $\leq r \leq$ __________

and __________ $\leq \theta \leq$ __________

(b) With this parameterization, what is $d\vec{A}$?

$d\vec{A} =$ ______________________________

(c) Find the flux of $\vec{F}$ through S.

flux = ______________________________

5. Compute the flux of the vector field $\vec{F} = z\vec{i} + 8x\vec{k}$ through the parameterized surface S oriented upward and given, for $0 \leq s \leq 4, \quad 2 \leq t \leq 4$, by

$$x = s^2, \quad y = 2s + t^2, \quad z = 5t.$$

flux = ______________________________

6. Are the following statements true or false?

(a) 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.

(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 the flux integral $\iint_S \vec{F} \cdot d\vec{A}$ is zero, then $\vec{F} = \vec{0}$.

(d) 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.

(e) 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.

(f) If S_1 is a rectangle with area 1 and S_2 is a rectangle with area 2, then $2\iint_{S_1} \vec{F} \cdot d\vec{A} = \iint_{S_2} \vec{F} \cdot d\vec{A}$

(g) The area vector of a flat, oriented surface is parallel to the surface.

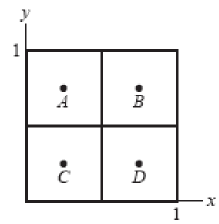
(h) If S is an open-ended circular cylinder centered about the z-axis, oriented away from the z-axis, and $\vec{F} = \langle x, y, 0\rangle$, then the flux of $\vec{F}$ through S is positive.

7. Let $\mathcal{S}$ be the square in the xy-plane shown in the figure below, oriented with the normal pointing in the positive z-direction. Estimate

$$\iint_{\mathcal{S}} \mathbf{F} \cdot d\mathbf{S}$$

where $\mathbf{F}$ is a vector field whose values at the labeled points are

$$\begin{aligned} \mathbf{F}(A) &= \langle -8, 8, 8\rangle, & \mathbf{F}(B) &= \langle -3, 5, 1\rangle \\ \mathbf{F}(C) &= \langle -5, -7, 6\rangle, & \mathbf{F}(D) &= \langle -7, -5, 9\rangle \end{aligned}$$



$\iint_S \mathbf{F}\cdot d\mathbf{S} \approx$ ____________________

8. $\mathrm{F} = 2x\boldsymbol{i} + 2y\boldsymbol{j} + z\boldsymbol{k}$ and σ is the portion of the cone $z = \sqrt{x^2+y^2}$ between the planes $z = 2$ and $z = 7$ oriented with normal vectors pointing upwards. Find the flux of the flow field F across σ: ____________________

9. **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 13.9.7 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 $0 \le x, y \le 2$

10. **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.

13.9.5 Notes to Instructors and Dependencies

This section relies on parameterized surfaces, which was first introduced in Section 12.6. While some of the activities in this section may be too much for a single student to do in a class setting, we suggest that different cases of the vector fields and surfaces can be split for small group work. Students can then present answers to the larger group.

13.10 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?

Preview Activity 13.10.1 In Preview Activity 13.5.1 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.

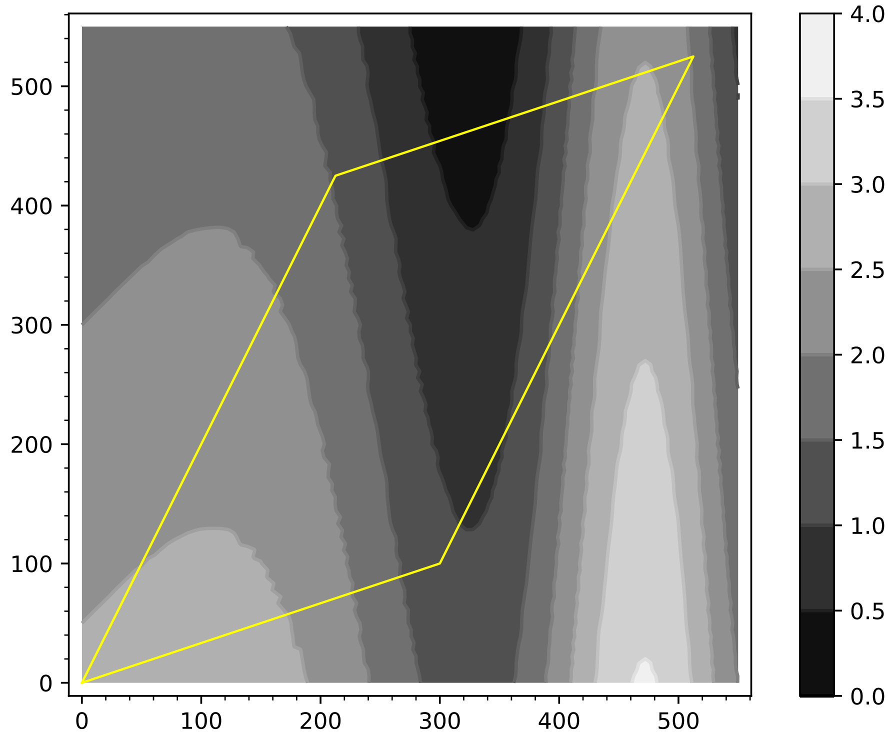


Figure 13.10.1 A plot of land with density of copium deposits and the edges of the plot drawn in yellow

The opening tasks of Preview Activity 13.5.1 had you estimating the amount of copium mined by driving along the edge of a plot of land (drawn in yellow on Figure 13.10.1). This iterpretation 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.

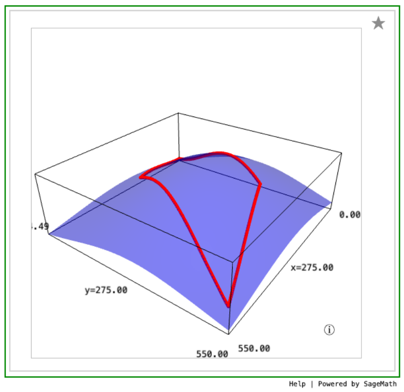
(a) In this task, we will interpret Figure 13.10.1 as a contour graph of the density of copium per unit area. This will allow us to compute the total amount of copium in our mining area by setting up a double integral. To approximate the total amount of copium available in our mine, do the following:

- **(i)** Break the mining plot (the area inside the yellow segments of Figure 13.10.1) 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 13.10.1, your mining plot was on a hill as represented in Figure 13.10.2. 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 13.10.2 is greater than on Figure 13.10.1
- the total amount of copium available on Figure 13.10.2 is lesser than on Figure 13.10.1
- the total amount of copium available on Figure 13.10.2 is the same as on Figure 13.10.1

Explain your reasoning for your choice.



Standalone
Embed

Figure 13.10.2 A 3D plot of the region including our mining area (shown with boundary in red)

In Section 13.9, 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 13.9.7 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 12.6.

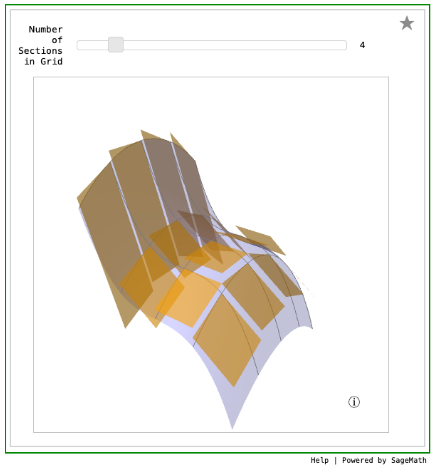
Just as we generalized line integrals of vector fields into scalar line integrals in Section 13.5, we will spend this section examining questions related to the accumulation of scalar-valued functions along surfaces in space as a means of generalizing flux integrals.

13.10.1 Defining surface integrals of scalar functions

In order to create a Riemann sum that will measure the acculumation 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 12.6.2 and Section 13.9, 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 $a \leq s \leq b$ and $c \leq t \leq d$. In our classic calculus style, 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 we slice $c \leq t \leq d$ 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 12.6 (specifically equation (12.6.1)) 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 13.10.3, where you can adjust the slider to control the number of rectangles. (For simplicity, the figure takes $m = n$.)



Standalone
Embed

Figure 13.10.3 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.

Theorem 13.10.4 *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*

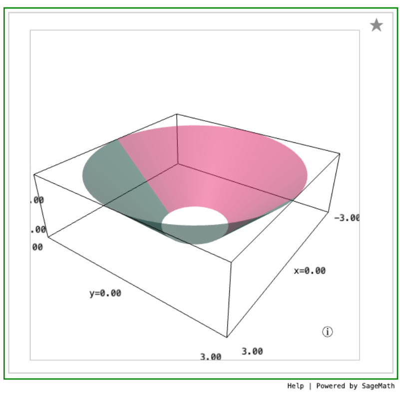
$$\iint_{S_1} f\, dS = \iint_D f(s, t)|(\vec{r}_s \times \vec{r}_t)(s, t)|\, dA.$$

In the following example, we both reason about the value of a scalar surface integral without working through computations and apply Theorem 13.10.4 to compute values of scalar surface integrals.

Example 13.10.5 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 13.9.9. 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 13.10.6, 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 13.10.6 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
Embed

Figure 13.10.6 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 13.9.9 in Theorem 13.10.4. 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 13.9.9 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 13.10.4 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 $1 \le s \le 3$ and $-\frac{\pi}{2} \le t < \frac{\pi}{2}$. As before, $|\vec{w}| = \sqrt{s^2(\cos(t)^2+\sin(t)^2)+s^2} = 2s$. Thus, Theorem 13.10.4 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. We

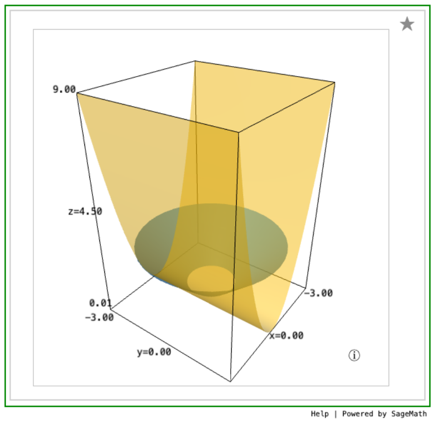
Example 13.10.7 For this example, we will consider the same surface S_1 as in Example 13.10.5, but we will change the scalar function we are integrating to something more complicated. Our goal is to calculate

$$\int_{S_1} z - x^2\,dS.$$

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 13.10.8, 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! Beacuse 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
Embed

Figure 13.10.8 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 arguement about the size of scalar surface integrals, but the algebraic calculation given by Theorem 13.10.4 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 13.10.2 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 $\iint_{S_1} x + z\, dS$.

13.10.2 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 13.10.3 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. 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 line integral.

(b) Explain in your own words what $\displaystyle\iint_{S_1} (2f)\, dS = 2\iint_{S_1} f\, dS$ 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 $\displaystyle\iint_{S_2} (f+g)\, dS = \iint_{S_2} f\, dS + \iint_{S_2} g\, dS$ 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 $\displaystyle\iint_{-S_2} f\, dS = \iint_{S_2} f\, dS$ means in the new analogy. It may be helpful to describe each side of the equation

[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.

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 13.10.9 Let us return to one of the problems from Activity 13.10.2. 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 helpul 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 13.10.4, 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 13.10.4 to compute the scalar surface integral as

$$\begin{aligned}
\iint_{S_1} x\,dS &= \int_0^{\pi/2}\int_0^{2\pi} [\cos(s)\sin(t)]\,(\sin(t))\,ds\,dt \\
&= [\sin(s)]_0^{2\pi}\int_0^{\pi}(\sin(t))^2\,dt = 0\cdot\int_0^{\pi}(\sin(t))^2\,dt = 0.
\end{aligned}$$

This result should match your answer for part a of Activity 13.10.2 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 13.10.4 Compute the value of the scalar surface integral in part c of Activity 13.10.2. 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.

13.10.3 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.

13.10.4 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 $\Phi(u,v) = (4u+3, u-v, 7u+v)$ 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 $\mathcal{D} = (u,v) : 0 \le u \le 8, 0 \le v \le 8$.

(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

$$y = 4 - z^2, \qquad 0 \le x, z \le 9; \qquad f(x,y,z) = z$$

$\iint_{\mathcal{S}} f(x,y,z)\, dS =$ ________________

3. Evaluate $\displaystyle\iint_S \sqrt{1+x^2+y^2}dS$ where S is the helicoid: $\mathbf{r}(u,v) = u\cos(v)\mathbf{i} + u\sin(v)\mathbf{j} + v\mathbf{k}$, with $0 \le u \le 3, 0 \le v \le 3\pi$

13.10.5 Notes to Instructors and Dependencies

This section relies parameterized surfaces, which was first introduced in Section 12.6. While some of the motivation for this section came from our treatment of flux integrals, it is not necessary to cover flux integrals first.

13.11 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 $\operatorname{curl}(\vec{F})$ through a surface with C as its boundary?

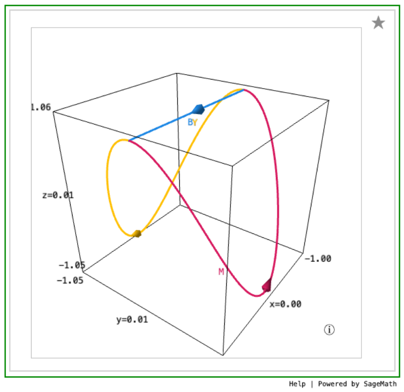
- Why does the flux integral of $\operatorname{curl}(\vec{F})$ through a surface with boundary only depend on the boundary of the surface and not the shape of the surface's interior?

When we studied Green's Theorem in Section 13.8, 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 13.11.1 In this activity, we will look at how we can apply the ideas about circulation along overlapping curves from the beginning of Subsection 13.8.1 to curves in space.

(a) For this part, consider the curves in Figure 13.11.1, where the yellow curve is Y, the blue curve is B, and the magenta curve is M.



Standalone
Embed

Figure 13.11.1 Three curves in space

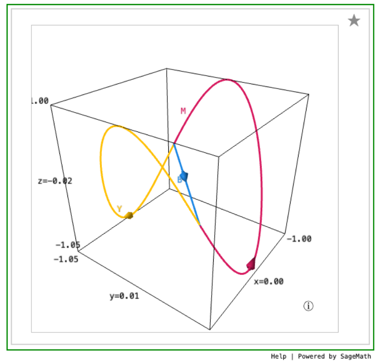
You should also go back and refamiliarize yourself with our notation for combining paths (as used in line integrals) from Convention 13.2.13. In our convention, $Y + M$ would be closed loop, but $Y - M$ would not make sense because the segment $-M$ does not begin where Y begins.

(i) Using the three segments in Figure 13.11.1, 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 $\int_{C_1+C_2} \vec{F} \cdot d\vec{r}$.

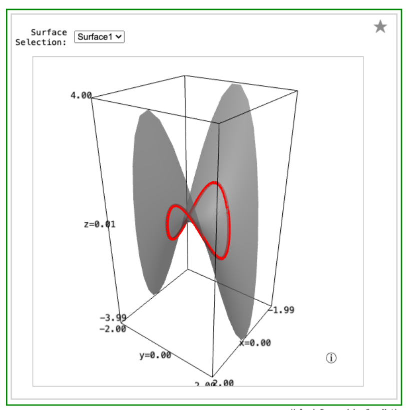
(iv) Explain how your arguments or equations from any of the parts above would or would not change if you considered the curves depicted in Figure 13.11.2.



Standalone
Embed

Figure 13.11.2 Three slightly different curves in space

(b) Let C be the simple closed curve consisting of the yellow and magenta curves in Figure 13.11.1. You can see C plotted in red in Figure 13.11.3. 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
Embed

Figure 13.11.3 Surfaces containing a common simple closed curve

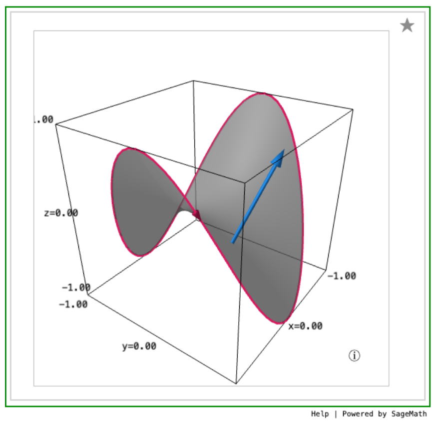
The simple closed curve consisting of the yellow and magenta curves in Figure 13.11.1 can be parameterized by $\langle \cos(t), \sin(t), \cos(2t) \rangle$ with $0 \le t \le 2\pi$. Let $C_3 = Y + M$. Use the given parameterization of C_3 to show that C_3 is on each of the following surfaces:

- $x^2 - y^2 = z$
- $z = x^4 - y^4$
- $z = 1 - 2y^2$
- $z = -\cos(\pi\sqrt{x^2 + y^2})(x^2 - y^2)$

13.11.1 Circulation in three dimensions and Stokes' Theorem

In part b of Preview Activity 13.11.1, 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 13.11.4, we show the curve C from Preview Activity 13.11.1 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
Embed

Figure 13.11.4 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 13.11.1, 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 13.7, the curl of a vector field in $\mathbb{R}^3$ measures the rotation of the vector field. Theorem 13.7.18 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 $\iint_D \operatorname{curl}(\vec{F}) \cdot (\vec{r}_s \times \vec{r}_t)\, dA$, 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 Preview Activity 13.11.1 and our discussion of Green's Theorem provide an intuitive description of why this theorem is true.

Theorem 13.11.5 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* $\operatorname{curl}(\vec{F})$ *through S. That is,*

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_D (\operatorname{curl}(\vec{F})) \cdot (\vec{r}_s \times \vec{r}_t)\, dA.$$

13.11.2 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 13.11.6 In this example, we will verify Stoke's Theorem for the curve used in Preview Activity 13.11.1 using the parameterization and surfaces from part b of Preview Activity 13.11.1. 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 13.3.6 to calculate the circulation, we have

$$\int_C \vec{F} \cdot d\vec{r} = \int_0^{2\pi} \langle \cos(2t) - \cos(t), \sin(t) + \cos(2t), \cos(t)\sin(t)\rangle \cdot \langle -\sin(t), \cos(t), -2\sin(2t)\rangle\, dt$$

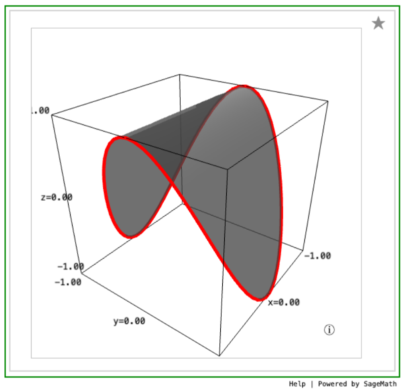
If you were to write out this dot product and combine like terms, then you would be left with four terms. Each of these can be evaluated using a few trig identities and substitutions. In fact, three of the four terms will correspond to functions that have as much area below the axis as above and thus will have integrals of zero.

$$\begin{aligned}\int_C \vec{F} \cdot d\vec{r} = &\int_0^{2\pi} 2\cos(t)\sin(t)\, dt \\ &- \int_0^{2\pi} \cos(2t)\sin(t)\, dt \\ &+ \int_0^{2\pi} \cos(2t)\cos(t)\, dt \\ &- \int_0^{2\pi} 2\cos(t)\sin(t)\sin(2t)\, dt\end{aligned}$$

$$= 0 - 0 + 0 - \pi = -\pi$$

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 Preview Activity 13.11.1, 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 13.11.7. 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
Embed

Figure 13.11.7 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 9 of Section 13.9, you may recognize our caclulation 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 $\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\rangle \, dA$.

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 13.9.7, 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 Preview Activity 13.11.1 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 simpmlifications 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

$$\int_0^1 \int_0^{2\pi} (-2s^2\cos(t))(s\cos(t) - 1) + (2s^2\sin(t))(1 - s\sin(t)) \, dt \, ds$$

$$= \int_0^1 \int_0^{2\pi} -2s^3 \cos(t)^2 + 2s^2 \cos(t) + 2s^2 \sin(t) - 2s^3 \sin(t)^2 \, dt \, ds$$
$$= \int_0^1 \int_0^{2\pi} -2s^3 + 2s^2 \cos(t) + 2s^2 \sin(t) \, dt \, ds.$$

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 13.11.2 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 $0 \le t \le 2\pi$.

- **(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 $y \le x$. Calculate the flux of $\mathrm{curl}(\vec{F})$ 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 $x, y, z \ge 0$. Find the flux of $\mathrm{curl}(\vec{G})$ through S_2.
- **(iii)** Write a sentence to explain why the sign of your answer to the previous two parts makes sense.

Activity 13.11.3

(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 Activity 13.11.3, 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.

13.11.3 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 6 and Exercise 8 of Section 13.8, 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 4 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 $\operatorname{curl}(\vec{F})$ from amongst possibly several reasonable surfaces with boundary C. The next two activities focus on the relationships between surfaces and their boundary.

Activity 13.11.4 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 13.11.5 In some sense, this activity considers the reverse problem of that considered in Activity 13.11.4. 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 13.11.1, 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 $0 \le t \le 2\pi$.

(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 $(2, 0, 0)$.

13.11.4 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 $\text{curl}(\vec{F})$ 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.

13.11.5 Exercises

1. Find $\int_C \vec{F} \cdot d\vec{r}$ where C is a circle of radius 2 in the plane $x + y + z = 3$, centered at $(1, 4, -2)$ and oriented clockwise when viewed from the origin, if $\vec{F} = 4y\vec{i} - 5x\vec{j} + (y - x)\vec{k}$

$\int_C \vec{F} \cdot d\vec{r} =$ ____________________

2. Use Stokes' Theorem to evaluate $\displaystyle\int_C \mathbf{F} \cdot d\mathbf{r}$ where $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + 8(x^2 + y^2)\mathbf{k}$ and C is the boundary of the part of the paraboloid where $z = 49 - x^2 - y^2$ 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 $\mathbf{F} = \langle 8z, 5x, 8y\rangle$.

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. 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 $\operatorname{curl}(\vec{G})$ 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 $\operatorname{curl}(\vec{F}) = \langle x - y^2, 4xy - y - 4, -4xz\rangle$.

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 $\displaystyle\int \cos(x)\, dx = \sin(x) + C$. 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 $\vec{F}$, $\operatorname{div}(\vec{H}) = 0$. Such a vector field is called a **solenoidal vector field** or **divergence-free vector field**.

5. For each of the following vector fields, determine whether a vector potential exists. If so, find one.

For this problem, enter your vectors with angle-bracket notation: < a, b, c >, not in ijk-notation.

(a) $\vec{F} = 8x\,\tilde{i} + \left(y - z^2\right)\,\tilde{j} + (3x - 9z)\,\tilde{k}$

$\vec{F}$ (☐ has a vector potential ☐ does not have a vector potential)

$\vec{H}$, $\vec{H} =$ ______________________________

*(If there is no potential function, enter **none** for the function.)*

(b) $\vec{F} = 8x\,\tilde{i} + \left(y - z^2\right)\,\tilde{j} + (3x + 9z)\,\tilde{k}$

$\vec{F}$ (☐ has a vector potential ☐ does not have a vector potential)

$\vec{H}$, $\vec{H} =$______________________________

*(If there is no potential function, enter **none** for the function.)*

6. Repeat the steps of part 13.11.6.b and part 13.11.6.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 13.11.6.

13.11.6 Notes to Instructors and Dependencies

This section relies heavily on understanding flux integrals Section 13.9 as well as the calculation of circulation around a closed curve (from Section 13.8 and

Section 13.2). Subsection 13.11.3 gives some reminders about the different ways to parameterize surfaces, which was first introduced in Section 12.6.

13.12 The Divergence Theorem

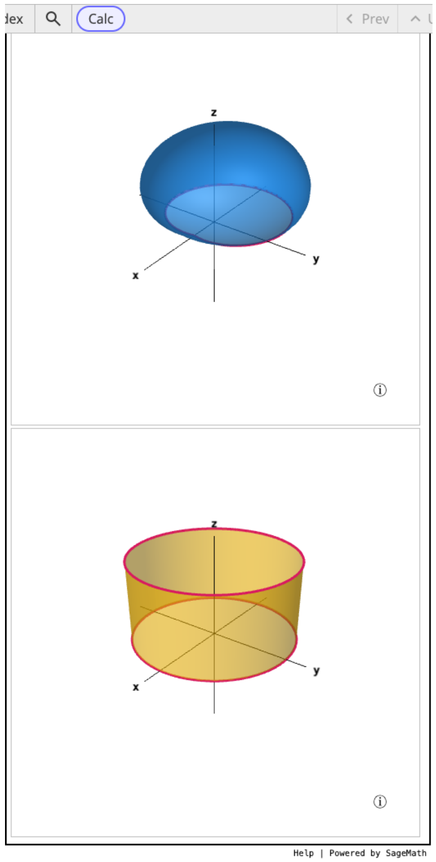
Motivating Questions

- 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?

In Section 13.6 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 13.9, 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 13.6. 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.

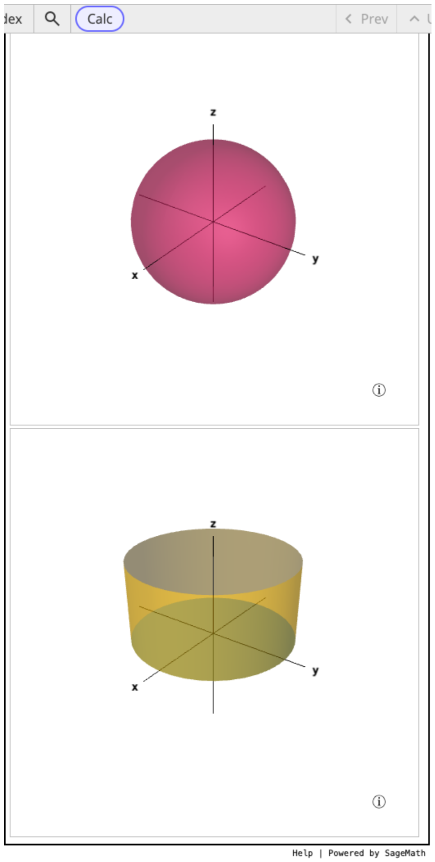
Recalling the surfaces we studied in Section 13.11, where we applied Stokes' Theorem, notice that all of these surfaces had the property that they had a **boundary** along which we calculated circulation. It turns out that giving a precise definition for boundary is challenging. For our purposes, however, 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 13.11, *none* of the surfaces were closed, because each had a boundary curve. On the other hand, in Figure 13.12.1, we show two more surfaces that are *not* closed, as demonstrated by the magenta curves marking the edge where the surface ends.



Standalone
Embed

Figure 13.12.1 Two surfaces that are *not* closed

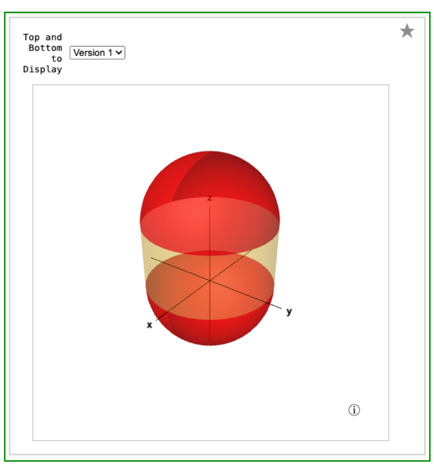
In fact, the yellow cylinder has two edges where the surface ends (does not meet itself). The surfaces in Figure 13.12.2 *are* closed because there is no edge to the surface. Note that the cylinder has been "filled in" with a cap at top and bottom (plotted in gray and green, respectively) to become a closed surface.



Standalone
Embed

Figure 13.12.2 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 13.12.3 illustrates three different ways to complete the cylindrical surface into a closed surface.



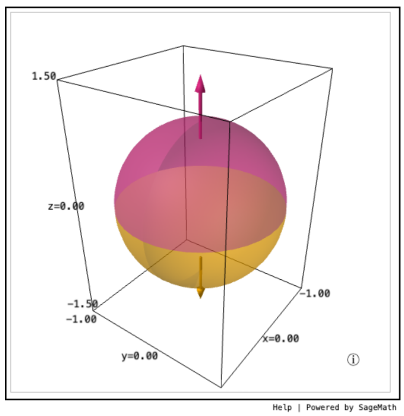
Standalone
Embed

Figure 13.12.3 Three ways to close a cylindrical surface by adding a top and bottom

Preview Activity 13.12.1 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.

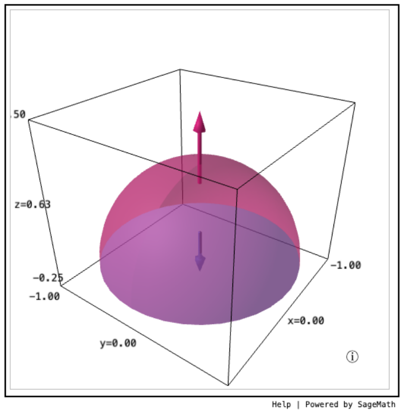
(a) In this part, we will consider the unit ball shown in Figure 13.12.4, with boundary and normal vectors as shown in the plot. If the flux integral of a poisonous gas through S_{top} is 15 and the flux integral of the poison gas through S_{bottom} is -3, is the interior of the sphere emitting or absorbing poisonous gas? Explain your reasoning.



Standalone
Embed

Figure 13.12.4 Unit ball with boundary given by the combination of S_{top} and S_{bottom}

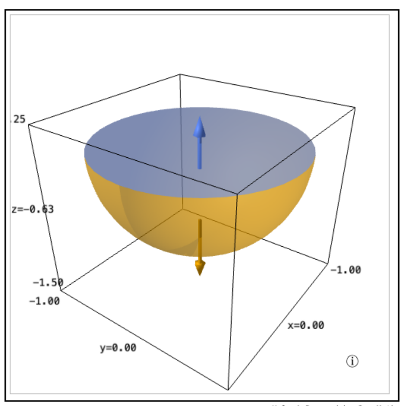
(b) In this part, we will consider the top half of the unit ball shown in Figure 13.12.5, with boundary and normal vectors as shown in the plot. If the flux integral of a poisonous gas through S_{top} is 15 and the flux integral of the poison gas through S_{mid} is -20, is the top half of the unit ball emitting or absorbing poisonous gas? Explain your reasoning.



Standalone
Embed

Figure 13.12.5 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 13.12.6, 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
Embed

Figure 13.12.6 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.

13.12.1 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 13.12.7 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 13.12.8 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 13.12.9. 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
Embed

Figure 13.12.9 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 13.9.7:

$$\begin{aligned}
S_1 &: \int_0^1 \int_0^{1-s} \langle (0)s - t, st - e^1, t(0-s)\rangle \cdot \langle 1, 0, 0\rangle \, dt\, ds \\
S_2 &: \int_0^1 \int_0^{1-s} \langle s(0) - t, 0(t) - e^s, t(s-0)\rangle \cdot \langle 0, -1, 0\rangle \, dt\, ds \\
S_3 &: \int_0^1 \int_0^{1-s} \langle st - 0, t(0) - e^s, 0(s-t)\rangle \cdot \langle 0, 0, 1\rangle \, dt\, ds \\
S_4 &: \int_0^1 \int_0^{1-s} \langle st - (1-s-t), t(1-s-t) + e^s, (1-s-t)(s-t)\rangle \\
&\qquad\qquad \cdot \langle 1, 1, 1\rangle \, dt\, ds
\end{aligned}$$

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:

$$\begin{aligned}
S_1 &: \int_0^1 \int_0^{1-s} -t \, dt\, ds \\
S_2 &: \int_0^1 \int_0^{1-s} -e^s \, dt\, ds \\
S_3 &: \int_0^1 \int_0^{1-s} 0 \, dt\, ds \\
S_4 &: \int_0^1 \int_0^{1-s} -1 + 2s + t + e^s - s^2 \, dt\, ds
\end{aligned}$$

Combining these we have:

$$\begin{aligned}
S_4 - S_1 + S_2 - S_3 : \quad & \int_0^1 \int_0^{1-s} -1 + 2s + 2t - s^2 \, dt\, ds \\
& \int_0^1 (s-1)^2 s \, ds
\end{aligned}$$

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 $0 \leq x \leq 1$, $0 \leq y \leq 1 - x$, and $0 \leq z \leq 1 - x - y$. Thus, our triple integral I will be set up and evaluated as

$$\begin{aligned}I &= \int_0^1 \int_0^{1-x} \int_0^{1-x-y} z + x\, dz\, dy\, dx\\ &= \int_0^1 \int_0^{1-x} 1/2(1 - x - y) + x(1 - x - y)\, dy\, dx\\ &= \int_0^1 \frac{1}{12}(\frac{x^3}{3})\, dx\end{aligned}$$

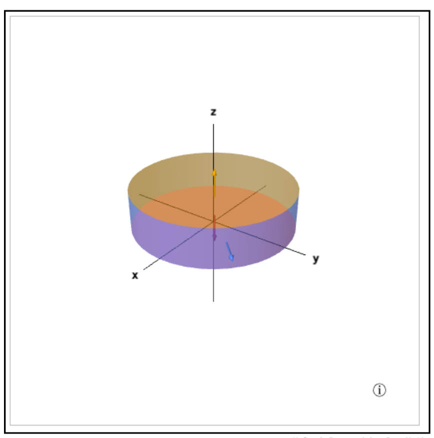
This evaluates to $\frac{1}{12}$, which matches with the flux calculations we did in the previous step to verify the Divergence Theorem for our example.

□

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 13.12.2 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 13.12.10. 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

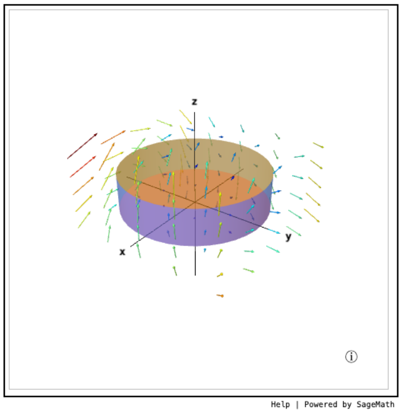
$$\vec{F} = \langle xy - 2z, y^2 - yz, 3x + z^2 \rangle.$$



Standalone
Embed

Figure 13.12.10 A closed cylindrical surface

(a) Figure 13.12.11 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
Embed

Figure 13.12.11 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 13.12.3

(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

13.12.2 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.

13.12.3 Exercises

1. Verify the Divergence Theorem for the vector field and region:
$\mathbf{F} = \langle 8x, 7z, 8y \rangle$ and the region $x^2 + y^2 \leq 1$, $0 \leq z \leq 2$
$\iint_S \mathbf{F} \cdot d\mathbf{S} =$ ____________________
$\iiint_{\mathcal{R}} \operatorname{div}(\mathbf{F})\, dV =$ ____________________

2. 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 $x^2 + y^2 \leq 4,\ 0 \leq z \leq 5$.
$\iint\limits_S \vec{F} \cdot d\vec{A} =$ ______________________________

3. 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. $x^2 + y^2 \leq z \leq 4$ ______________________________
B. $x^2 + y^2 \leq z \leq 4,\ x \geq 0$ ______________________________
C. $x^2 + y^2 \leq z \leq 4,\ x \leq 0$ ______________________________

4. 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 $x = 1, x = 4, y = 0, y = 3, z = 0, z = 1$, closed and oriented outward, and $\vec{F} = 4x^2\vec{i} + 3y^2\vec{j} + 4z^2\vec{k}$.

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 $x = 1$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _______________ $dz\, dy =$ _______________
where $a =$_____, $b =$_____, $c =$_____ and $d =$_____.
On $y = 3$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _______________ $dz\, dx =$ _______________
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 $z = 1$, $\int_S \vec{F} \cdot d\vec{A} = \int_a^b \int_c^d$ _______________ $dy\, dx =$ _______________
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 $\int_S \vec{F} \cdot d\vec{A} =$_________________________

5. 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.

6. 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.

7. 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.

8. 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.

13.12.4 Notes to Instructors and Dependencies

This section relies heavily on understanding flux integrals Section 13.9 as well as the meaning of the divergence of a vector field from Section 13.6 and triple integrals from Section 12.7.

Index