Multivariable Calculus

Chris Johnson

January 7, 2022

Contents

Contents ii

Introduction to the Course iv

1 Preliminaries 1
- 1.1 Three-dimensional space . . . 1
- 1.2 Higher dimensional space . . . 24
- 1.3 Vectors . . . 27
- 1.4 Linear transformations and matrices . . . 46
- 1.5 The dot product . . . 69
- 1.6 The cross product . . . 80
- 1.7 Lines and planes . . . 87

2 Curves 99
- 2.1 Vector-valued functions and parametric curves . . . 99
- 2.2 Limits, continuity, and tangent vectors . . . 103
- 2.3 Arclength . . . 114
- 2.4 Curvature . . . 118
- 2.5 Motion in space and integrals of vector-valued functions . 130

3 Functions of Multiple Variables 139
- 3.1 Basic Ideas . . . 139
- 3.2 Limits, continuity, and the sandwich theorem . . . 149
- 3.3 Partial derivatives . . . 156
- 3.4 Directional derivatives . . . 164
- 3.5 The total derivative and the chain rule . . . 172
- 3.6 Tangent planes . . . 184
- 3.7 Linearization . . . 199
- 3.8 Optimization . . . 206
- 3.9 Lagrange multipliers . . . 213

4 Integration **223**
4.1 Integration in two variables 223
4.2 Iterated integrals 234
4.3 Double integrals over general regions 243
4.4 Double integrals in polar coordinates 251
4.5 Triple integrals 263
4.6 Cylindrical and spherical coordinates 274
4.7 Change of variables 284

5 Vector Calculus **304**
5.1 Vector fields 304
5.2 Line integrals 310
5.3 The fundamental theorem of line integrals 323
5.4 Green's theorem 333
5.5 Curl and divergence 344

Appendix A Proofs of results from the text **357**
A.1 Proofs from Chapter 1 357

Appendix B Sets **369**
B.1 Introduction 370
B.2 Basic Ideas 370
B.3 Unions and Intersections 373
B.4 Cartesian Products 378
B.5 Standard Notations 380
B.6 Maps Between Sets 384
B.7 Summary 387

Introduction to the Course

Difficulties strengthen the mind as labor does the body.

SENECA THE YOUNGER

Welcome to Math 256, which is the third and final course in the calculus sequence at Western Carolina University. Our goal in this course is to take the calculus from the first two courses (Math 153 and 255 at WCU) and make it "useful." That is, the tools you learned in those classes (such as differentiation and integration) and their applications (such as solving optimization problems) can not be directly applied in many situations because they have a fundamental limitation: they assume you are only working with one variable. For example, the calculus you know thus far relating a particle's position to its velocity via differentation currently only applies when the particle is moving in a straight line. Particles can, of course, move in much more complicated ways. One of our primary motivations in this course is to extend the calculus you currently know to these more general situations. To do this we will have to come to terms with higher-dimensional space and functions of several variables, and the first part of the course is essentially devoted to making sense of these ideas and learning about the basic tools we need to do mathematics in these more general settings.

After spending some time setting up the basic "environment" for our mathematics, we begin discussing how to do calculus with multivariable functions. That is, we will learn how to differentiate and integrate these types of functions. As we will see, much of this material is a straightforward generalization after we've developed the appropriate background. (There are, however, some subtleties and quirks that can trip us up if we aren't careful.) We will also see applications of calculus of multivariable functions, for example in solving optimization problems including problems which have constraints.

Finally, towards the end of the course we will introduce the main ideas of "vector calculus." Whereas much of the material up to this point will be a straight-forward generalization of material you've seen before, here we start to develop some fundamentally new ideas. These topics (namely vector fields and line integrals) are of fundamental importance in modern mathematics and physics. (In fact, understanding the various conditions in which we can solve certain problems relating to these topics influenced a large part of research mathematics in the late 19th and early 20th centuries.) Because of time considerations, our discussion of these topics will necessarily be incomplete, but we will at least be able to describe some of the main ideas and important results.

How to study for this course

This course covers a lot of material, so we will necessarily have to move through some of that material quickly. Additionally, each topic in this course builds on the previous topics. As a consequence, it can be very difficult to get up to speed if you start to fall behind, **and so you need to take studying for this course seriously at the very beginning of the semester**! You will be expected to watch video lectures outside of class, read the lecture notes and the textbook, and complete assignments on a regular basis. The best way to do all of this is to make a plan about when you're going to study at the very start of the semester.

Everyone is different and everyone learns differently, but each student needs to make an effort to study this material *outside of class on a regular basis* if they hope to be successful. Some general guidelines for studying are the following:

- Make a point to study every day.
- Read the lecture notes and textbook regularly, ideally *before* coming to class.
- Read and study incredulously: don't simply take what you read in the notes or hear in a lecture as a factoid to memorize, but think about why it's true.
- Start on assignments early.
- Discuss the material with other students in the class; work together on assignments.
- Look for extra problems to practice with online or in a textbook.
- Ask lots of questions.

It's important to emphasize that you need to be working outside of class on a regular basis, and that just coming to class and hoping you will understand things is not a strategy for success.

Serious study, especially of a technical topic such as multivariable calculus, can be very daunting, especially at first. If you can get into a habit of studying, though, it does become easier with time. For example, if you have an hour inbetween your classes at some point in the day where you don't have anything else scheduled, it would be a good idea to make a habit of going to the library and just reading the notes and working on problems during that time. Even if it's hard to find an entire hour free in your schedule, studying for just fifteen or twenty minutes each day is much, much better than not studying at all.

Recognize this class will be difficult

There is no denying that this class will be challenging. The material can be technical, new ideas we introduce can be hard to wrap your head around at first, and there is a lot of work involved. It's easy to feel overwhelmed, especially if you allow yourself to slip into bad habits such as putting assignments off until the last minute and limiting your studying to cramming before an exam. The best way to avoid these bad habits is to recognize up front that this class will be hard and you will have to work hard to understand the material. Again, get in the habit of studying every day and starting assignments early. Even when you feel frustrated, it's important to make an effort to keep trying to work with the material and understand it by studying regularly. **The worst thing you can do is give up.**

It's okay to be confused and frustrated while studying. Learning is hard, especially when it comes to difficult material. Getting frustrated or confused doesn't mean you're a bad student and it doesn't mean you can't learn the material. The most important thing is to persist, and even though it can be hard to believe it at the time, you can learn and understand the material.

Get help when you need help

You are always encouraged to discuss the material in this course with other students and the instructor. When you are struggling with a topic or can't seem to get started on a particular problem, sometime just bouncing ideas off another person can help you reach that *"ah-ha!"* moment when the material starts to click. These moments don't happen out of the blue, however: you have to make an effort to think about the material and talk through it with someone to have those epiphanies.

You should always feel welcome to contact the instructor when you are having trouble. This can be done during office hours, or over email, or just by asking a question at the start or end of class. *Having questions* *is a good thing!* Questions show that you're thinking about the material, and that's what you have to do to understand the material, which in turn is what you have to do to do well in this class.

In addition to the instructor and your classmates, you should consider visiting the Mathematics Tutoring Center located in Killian Annex when you are having trouble. While not every tutor will necessarily be able to help you with every problem, there are very good students that understand this material very well working in the MTC, and they can be a very valuable resource as you study and work on assignments.

Take breaks

Even though studying regularly is important, it's also important that you take breaks while studying. You may find that it's difficult to focus on reading and thinking deeply for long periods of time if you're not accustomed to intense studying. For example, you may find that after only ten minutes it's hard to pay attention and that your eyes just glance over words on a page without you consciously thinking about what you're reading. This is normal and it does get easier to study for longer periods over time *as long as you make an effort to study regularly*. When you notice that you're not "really" studying, take a break to do something you enjoy for a few minutes. Just five or ten minutes might be all you need to reset and be ready to focus again.

Remove distractions

Students sometimes believe they are capable of multitasking and try to do several things while they are studying. For example, you might "study" while simultaneously responding to texts from your friends, and periodically checking your email (or Snapchat or Twitter or Tinder, or whatever), while also listening to music, while also chatting with a friend that's physically nearby. This is not truly studying. To actually study you need to remove yourself from all these distractions and focus on one thing at a time. This might mean that you need to physically isolate yourself in the library while you're studying and you might need to turn your phone off (*off* is better than silenced) while you're trying to study.

Don't just find time to study, make time to study

Everyone has a lot of commitments and responsibilities in their life outside of their classes, and that's a good and normal thing. Sometimes, however, these other commitments can make you think you have less time than you really do and make it seem like there aren't enough hours in the day for you to study. Learning to effectively manage your time and prioritize your studying is really the big secret to success in your classes. This might mean that you have to concessions with some of your free time in order to make time to study. For example, perhaps there's a particular video game that you enjoy playing and you play it for a few hours each day. If you're playing video games, or watching movies, or aimlessly scrolling through Twitter/Facebook/Instagram/Snapchat/Whatever, etc. for a few hours each day, then you can't really claim that you don't have time to study. You instead have to put your studying ahead of these other activities. (Notice this doesn't mean you have to give up those activities! It's important to have other things in your life you enjoy doing, you just have to prioritize studying if you want to do well in your classes.) You might have to make a deal with yourself that you will not play that game you love until after you've spent at least thirty minutes reading and started an upcoming homework assignment.

Math is not a spectator sport

You likely have been in the situation before where an instructor works a problem on the board in class and everything makes complete sense, but when you try to do a similar problem on your own (either while studying independently, or in doing a homework assignment, or a problem on an exam) you don't even know where to start. This is extremely common, but the only way to rectify this is for *you* to solve problems by yourself.

Unfortunately a lot of students fall into the trap of seeing someone else work through a problem, or reading through someone else's solution in a book or online, and thinking that they suddenly understand how to solve the problem on their own. However, as you have probably experienced, this is not a substitute for actually doing the hard work of thinking through a problem on your own. People sometimes say *math is not a spectator sport: the only way to learn math is to do math*.

I mention this here mostly go caution you that while discussing the material with other people, reading in the book or lecture notes, and looking material up online is all well and good, until you sit down and work through a problem on your own (i.e., without having to look at notes, ask other people for help, etc.), you don't truly know the material. The first time you work a particular type of problem you probably will need to look things up, ask for help, and so on, *and that's completely fine*, but your goal should be to repeat these problems until they become second-nature and you can complete them independently.

Other resources

There are lots of resources online that you may want to consider looking at while studying. Simply Googling for *"multivariable calculus"* will bring up many sites with notes, exercises, sample exams, etc., and you are welcome (and encouraged!) to use any resource that you find helpful. Here are a few particular things you might find helpful.

OpenStax Textbook

https://openstax.org/details/books/calculus-volume-3
This is the official textbook for our course. We will cover some of the material out of order compared to the textbook; we will also omit some things from the book; and cover some things not mentioned in the book. I think of the textbook as the secondary resource for this class and these notes as the primary resource. One nice feature of the textbook, however, is that it has a lot of extra exercises for you to practice with, so I encourage you to look at these exercises in the textbook that correspond to what we are discussing in class each week.

Paul's Online Calculus Notes

Notes: https://tutorial.math.lamar.edu/Classes/CalcIII/CalcIII.aspx
Pratice Problems: https://tutorial.math.lamar.edu/Problems/CalcIII/CalcIII.aspx
Paul's Online Calculus Notes are a nice and easy-to-read reference for most of the material we will discuss in this class. These notes also have a nice set of practice problems which have well-written solutions available.

MathInsight

Multivariable Calculus: https://mathinsight.org/thread/multivar
Vector algebra: https://mathinsight.org/thread/vector_algebra
MathInsight is a nice website that tries to provide intuitive descriptions of a variety of topics related to this course. Some of the topics are a bit "sporadic" and so it's not a complete, comprehensive resource for everything we will discuss, but it does have nice descriptions of many topics.

A note on the lecture notes

These lecture notes have evolved from my original handwritten notes containing examples I wanted to cover when I first taught multivariable calculus in graduate school. Over time I've tried to flesh out details and turn the notes into something that students can read to supplement the textbook. At this point I think of these notes as the primary resource for the course, and so *if* a student wasn't able to read both these lecture notes and the textbook, I think they should prioritize the lecture notes. If there is one major problem with the notes, it is that they don't contain many exercises for students. The textbook, however, has *lots* of exercises and so I would encourage students to look there for extra problems to practice with.

These lecture notes are a perpetual work in progress and are likely to contain typos and errors in places, but hopefully these mistakes are minor issues and not fundamental problems. Please feel free to email me (`cjohnson@wcu.edu`) if you notice any mistakes.

Proofs

Some, but not all, of the theorems that appear here have an accompanying proof. You should not feel obligated to read through these proofs if you do not want to. The proofs are mostly included for the sake of completeness and to satisfy the curiosity of any interested student. Students that are mathematics majors may want to make an effort at understanding the proofs just as a personal exercise to familiarize themselves with proofs, but the proofs are not necessary to understand the main ideas, computations, or applications of the material here.

At some point in the future I hope to add proofs of each theorem that appears in the notes, but place the proofs in an appendix so they don't distract from the main text.

Preliminaries

1

The White Rabbit put on his spectacles. "Where shall I begin, please your Majesty?" he asked. "Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

LEWIS CARROLL
Alice's Adventures in Wonderland

Multivariable calculus is concerned with extending the calculus you have learned thus far to functions of several variables. As a consequence, we will be doing calculus in higher dimensional space. That is, when considering a function of one variable, $f(x)$, we are often interested in the graph $y = f(x)$ which is a curve in the plane. If we have a functions of two variables, however, say $f(x, y)$, then instead of defining a curve in the two-dimensional plane, the graph of such a function is a surface in three-dimensional space. We can go further and consider the graph of a function of three variables, $f(x, y, z)$, which would be three-dimensional object living in four-dimensional space. Though we may not be able to easily visualize such an object, all of the calculus we will develop will apply to these and even higher-dimensional objects.

In this chapter we introduce the basic ideas of three-dimensional (and higher dimensional) space, as this material may not be familiar to many students. That is, the goal of this first chapter is simply to set the stage and prepare us with the basic tools we need do calculus in several dimensions.

1.1 Three-dimensional space

In this section we will introduce three-dimensional space and briefly talk about some of the geometric objects that live in three dimensions. In particular, we'll describe the distance formula for telling how far two points in 3-space are from one another. We'll then use this formula to derive the equation of a sphere. Finally, we'll discuss the coordinate planes and some very simple types of surfaces.

Introduction

What does it mean to say something is one-dimensional, two-dimensional, three-dimensional, or eleven-dimensional? We hear a lot about how we all live in a three-dimensional world, and maybe you can believe that we really live in a four-dimensional world where time is the fourth dimension. But what does all of this really mean?

For us the notion of "dimension" will be a little bit informal and hand-wavy. We could spend some time trying to formally define dimension, but this would take up too much time. For what we want to do in this class, we'll just say that if an object is n-dimensional, this means you need *at least* n numbers to specify a point in that object.

Consider the real number line. If you want to specify a point on the line, you just need a single number: this number tells you how far away from zero you are on the line (this is the absolute value of the number) and whether you're to the left or right of zero (depending on whether that number is positive or negative). Thus the real line is one-dimensional.

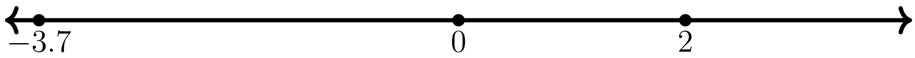


Figure 1.1: The real number line is one-dimensional.

What about the Euclidean plane (the xy-plane)? Here you need two numbers to specify a point: an x-coordinate and a y-coordinate. Once you have those two pieces of information, you can uniquely identify a point in the plane. As we need two pieces of information to determine a point in the plane, the plane is two-dimensional.

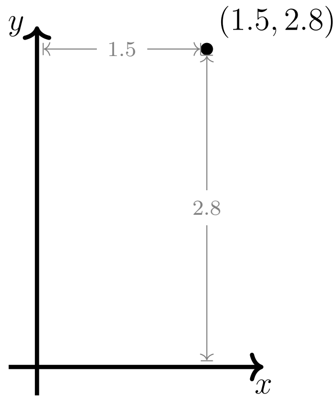


Figure 1.2: The plane is two-dimensional.

This notion of dimension may seem straight-forward, but it can be a little bit misleading. For example, what's the dimension of a circle (e.g., what's the dimension of the collection of points satisfying $x^2 + y^2 = 1$)? Your first guess might be that the circle is two-dimensional since it lives inside the two-dimensional plane, but the circle is only one-dimensional! If you know the angle θ of the point, measured from the positive x-axis, then you know the point; see Figure 1.3. Since you only need one number to specify a point on the circle, the circle is one-dimensional.

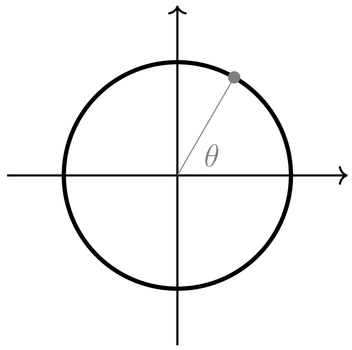


Figure 1.3: The circle is one-dimensional.

Notice that in the case of the circle, you could also use two numbers (x- and y-coordinates) to specify a point. We are defining dimension to use *as few numbers as possible*, however, and this is why the circle is one-dimensional.

As another simple example, consider an *annulus* (aka a ring). This is the collection of all points satisfying an equation of the form $a \leq x^2 + y^2 \leq b$. The case when $a = 1$ and $b = 4$ is illustrated in Figure 1.4 on the following page. What's the dimension of this annulus? Notice here we can't get away with just knowing an angle, because there will be lots of points that have the same angle. Once we specify an angle, we then need to specify which one of these points we mean and this makes us use up another number. So the annulus is two-dimensional.

In this course we'll mostly care about two-dimensional objects that live inside of three-dimensional space. It will take some time for us to get through all of the details and background of what exactly this means, so we'll begin by simply defining three-dimensional space. Before we do that, though, we need to make one small technical detour.

Sets

In this course it will be convenient sometimes to use the language of "set theory" to describe some of the objects we will consider. We won't spend any time in the "theory" of set theory; we will just adopt the language and notation, and even this we will do in a rather informal way.

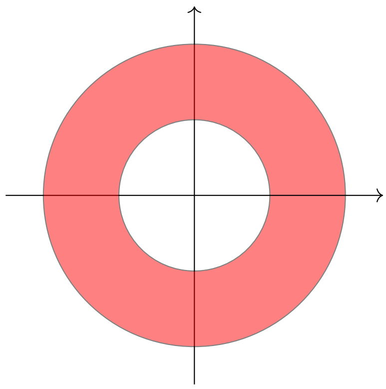


Figure 1.4: The annulus, $1 \leq x^2 + y^2 \leq 4$, is two-dimensional.

Remark.
Most modern mathematics is described using set theory, and if you continue to take mathematics courses you will eventually take a course on logic and proof where you will learn more about the nitty gritty details of set theory that we're skipping over here.

A ***set*** is simply an unordered collection of objects, and those objects are called the ***elements*** of the set. The sets we will consider in this course will consist of either numbers or points in space. It's a pretty good bet that every single mathematical "object" you've dealt with before is actually set, even if you've never heard the terminology; circles, lines, and solutions to equations are all examples of sets.

The collection of all real numbers (all numbers you can represent as a decimal expansion, including infinite expansions) is denoted $\mathbb{R}$. Geometrically this corresponds to the number line: every point on the line corresponds to a unique real number, and every real number determines a unique point. The two-dimensional plane – the set of all ordered pairs (x, y) – is denoted $\mathbb{R}^2$ because it takes two real numbers to specify any point.

Geometric objects living inside the plane can also be described as sets. The unit circle centered at the origin, for example, can be thought of as the following set which we'll denote C:

$$C = \left\{(x, y) \,|\, x^2 + y^2 = 1\right\}.$$

What this notation means is that C refers to the collection of all ordered pairs, (x, y), which satisfy the equation $x^2 + y^2 = 1$.

Now say that P is some point in the plane. If P is an element of the circle – if the coordinates of P satisfy the equation $x^2 + y^2 = 1$ – then we write $P \in C$. The symbol $\in$ establishes a relationship between the point P and the circle C: it says P is a point that lives on the circle C. If Q was some other point in the plane which did not live on the circle C, we would write $Q \notin C$. For example, $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) \in C$, but $(3, -2) \notin C$.

We will use these symbols $\in$ and $\notin$ repeatedly throughout the semester, so it's good to go ahead and get used to them.

Remark.
In this course, "sets" are just a language that we will use from time to time. We don't need to know anything very deep about sets, and the things that we do need to know we will see through examples over and over. If you're particularly curious, though, you may read Appendix B to see a little bit more about sets. This, however, is completely optional and you shouldn't feel obligated to read that appendix if you don't want to.

Cartesian coordinates

Recall that the two-dimensional Euclidean plane is represented by $\mathbb{R}^2$: that is, every point in the plane can be represented by a pair of two real numbers called the *x-coordinates* and *y-coordinates* of that point. Three-dimensional space is basically the same idea, but with one extra piece of information which we'll denote with the letter z.

That is, three-dimensional space is represented by triples (x, y, z) of real numbers. The collection of all such triples is denoted $\mathbb{R}^3$:

$$\mathbb{R}^3 = \left\{(x, y, z) \middle| x, y, z \in \mathbb{R}\right\}.$$

Verbally, we sometimes say "*3-space*" or "*R-3*" to denote this space; it's just a little bit quicker than saying "*three-dimensional space*" every single time.

Just as the two-dimensional plane is normally represented graphically as a pair of orthogonal lines (i.e., two lines that meet at a $90°$-angle), called the *x-axis* and *y-axis*, in 3-space we have three lines that meet at $90°$-angles: the usual x- and y-axes, plus a new *z-axis*. See Figure 1.5 on the next page.

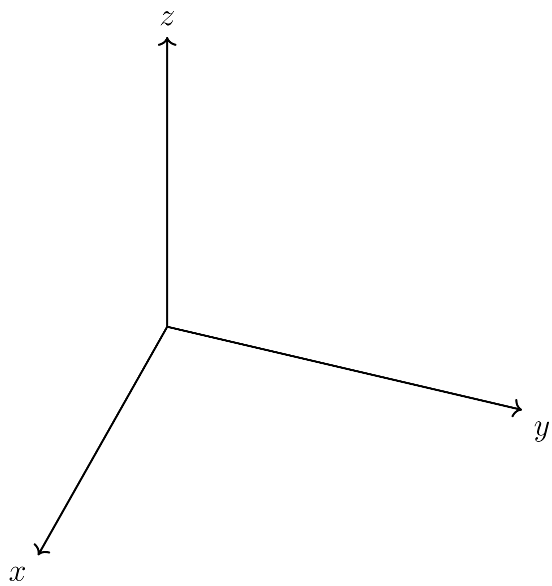


Figure 1.5: The x-, y-, and z-axes in $\mathbb{R}^3$.

By convention, we will use the ***right-hand rule*** to tell us how the x-, y-, and z-axes are oriented. This means that if you take your right hand, and your fingers point in the direction of the positive x-axis with your palm pointing towards the positive y-axis (so, in closing your hand to make a fist, your fingers would curl from the positive x-axis towards the positive y-axis), then your thumb will point in the direction of the positive z-axis. This is just a convention that people have adopted over the years. This is similar to the way that in the two-dimensional plane, by convention, the positive x-axis points to the right, while the positive y-axis points up. Of course, you could just arbitrarily decide, in two dimensions, that you're going to draw the positive y-axis pointing to the left and the positive x-axis will point up, but then the pictures you draw in the plane will be flipped around from what other people draw. By having a convention like this, you're sure that any picture you draw matches up with anyone else's picture. The right-hand rule is the same sort of idea in three dimensions: it's just there so that the images people draw in three dimensions are consistent.

Remark.
Since the page or screen you are reading this document on is two-

dimensional, any three-dimensional picture we try to draw, such as the axes in Figure 1.5 can be a little bit ambiguous. It is helpful, therefore, to sometimes use software to graph objects in three-dimensional space so that you can rotate the objects on the screen and get a better feeling for what's happening.

If you imagine that the usual Euclidean, xy-plane is the floor, then the z-axis sticks up out of the floor pointing straight up. This tells us how the (x, y, z)-coordinates of a point specify the point's position in 3-space. Say we have some point, call it P, and the coordinates of P are $P = (a, b, c)$.

Remark.
Some people use the notation $P(a, b, c)$ whereas we use $P = (a, b, c)$ here. Both of these are common, so you should be aware of these notations if you look in different textbooks or for material online.

In order to find the point the point $P = (a, b, c)$ in 3-space, we first locate the point (a, b) in the xy-plane, and then c tells us how far above or below the point (a, b) in the xy-plane the point P is. The standard convention is that if $c > 0$, then the point is above the xy-plane; and if $c < 0$, then the point is below the xy-plane.

Consider the point $P = (1, 2, 3)$. To locate this point, we find the point $(1, 2)$ in the xy-plane, and then go straight up three units from that point. See Figure 1.6 on the following page.

Another way to think about this is to move from the origin to the point, one step at a time, using the coordinates to tell you how far to step each time. That is, to find $P = (a, b, c)$, we start at the origin and then walk a units along the x-axis. (The direction we walk on the axis – left or right – depends, of course, on whether the value a is negative or positive.) Then we turn so that we're pointing in the same direction as the y-axis (or in the opposite direction, if $b < 0$), and walk b units. Finally, we look straight up (or down) and walk (fly?) c units to our point. In the case of the points $P = (1, 2, 3)$ and $Q = (3, -2, -1)$, this is illustrated in Figure 1.7 on the next page.

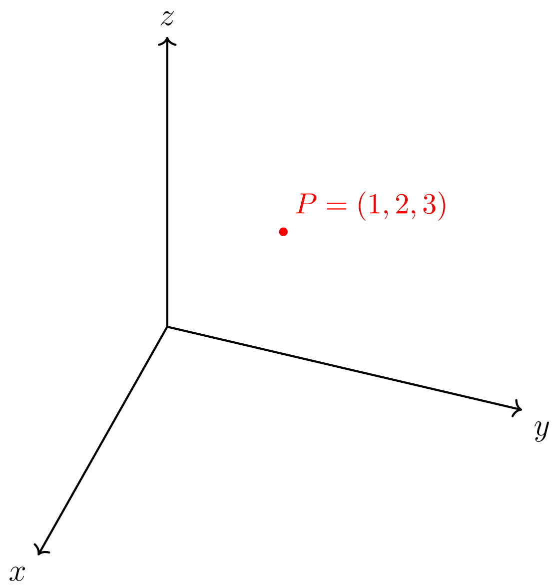


Figure 1.6: The point $P = (1, 2, 3)$.

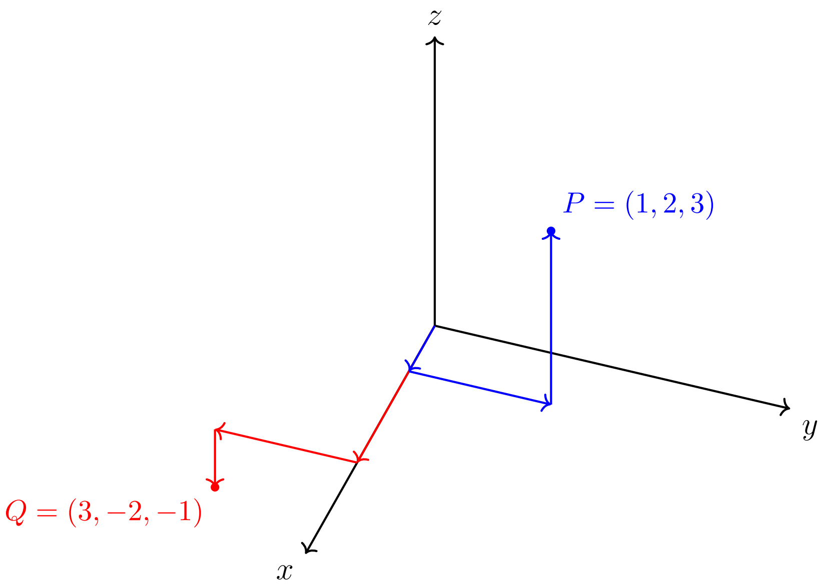


Figure 1.7: The points $P = (1, 2, 3)$ and $Q = (3, -2, -1)$, with lines indicating how to get to the points from the origin.

The distance formula

Given two points $P, Q \in \mathbb{R}^3$, we'd like to know how far away P and Q are from one another. That is, if we were to fly a spaceship around from point P to point Q, how far would that ship have to go? We will suppose that the shortest path between the two points, P and Q, is given by a line segment.[1] So, really the question we need to answer is "how long is the line segment that connects P to Q?" Before we answer that question in three-dimensions, let's briefly recall the answer in two dimensions.

Distances in two dimensions

In order to compute distances in two dimensions, we need to recall the Pythagorean theorem:

Theorem 1.1 (The Pythagorean Theorem)**.**
If T is a right triangle whose hypotenuse has length c, and whose other legs have lengths a and b, then $a^2 + b^2 = c^2$.

Remark.
In order to streamline these notes, proofs of theorems are not given in the main part of the text and are instead placed in Appendix A on page 357. The proofs are included mostly for completeness and to satisfy any particularly interested students, but are not required reading for any student.

We can use the Pythagorean theorem to determine the distance between two points. Let's say $P = (x_0, y_0)$ and $Q = (x_1, y_1)$ are two points in the plane and we draw a line segment between them, as in Figure 1.8 on the next page. Then we add two more lines to get a right triangle whose hypotenuse is the line between P and Q, as in Figure 1.9. Notice that the lengths of the legs of this triangle are $\Delta x = |x_1 - x_0|$ and $\Delta y = |y_1 - y_0|$.

---

[1] You might think it sounds silly to *suppose* the shortest path is a line segment – isn't the shortest path *always* a line? It depends on the geometry you're working with. We'll almost exclusively consider Euclidean geometry in this class, and so the distance is given by the length of a line segment. However when we discuss arclength of curves later, we'll very briefly talk about hyperbolic geometry, and there you'll see an example of a geometry where the shortest path between two points *isn't* a straight line.

This means that the length of the hypotenuse (and so the distance between P and Q) is given by

$$\sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}.$$

This distance between P and Q is sometimes denoted $|PQ|$ or $d(P, Q)$.

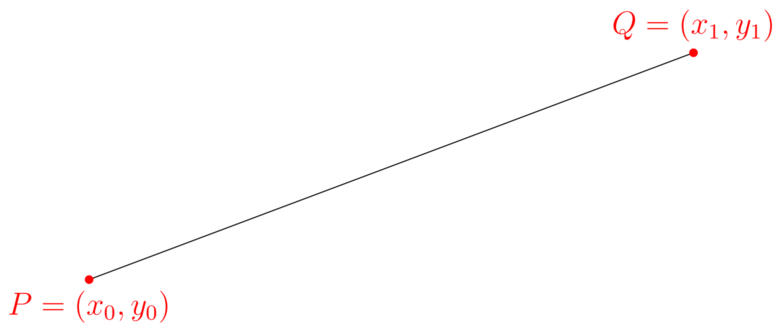


Figure 1.8: Two points in the plane.

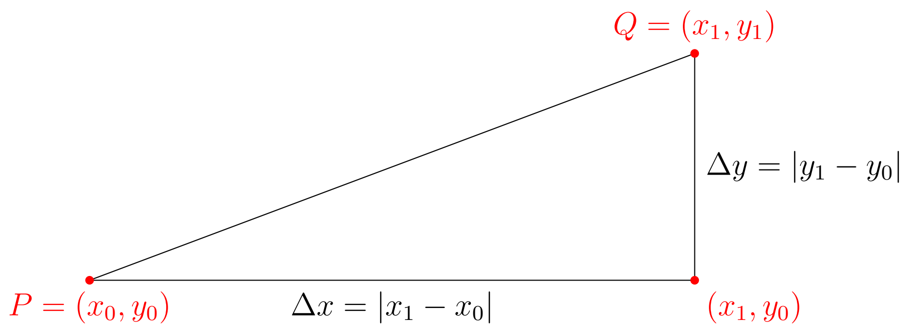


Figure 1.9: Given any two points, we can construct a triangle whose hypotenuse has the same length as the distance between those two points.

Distances in three dimensions

To compute the distance between two points P and Q in 3-space, we want to do the mimic what we did in two dimensions: represent the distance between two points as the hypotenuse of a triangle, and then use the Pythagorean theorem to determine the hypotenuse's length. Let's say our points have coordinates $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$. We draw a line segment between these two points, and then construct a triangle whose hypotenuse is this line segment, as in Figure 1.10 on the following page.

Remark.
The triangle in Figure 1.10 really is a right triangle, but because of the perspective it may not appear to be a right triangle on a screen or piece of paper.

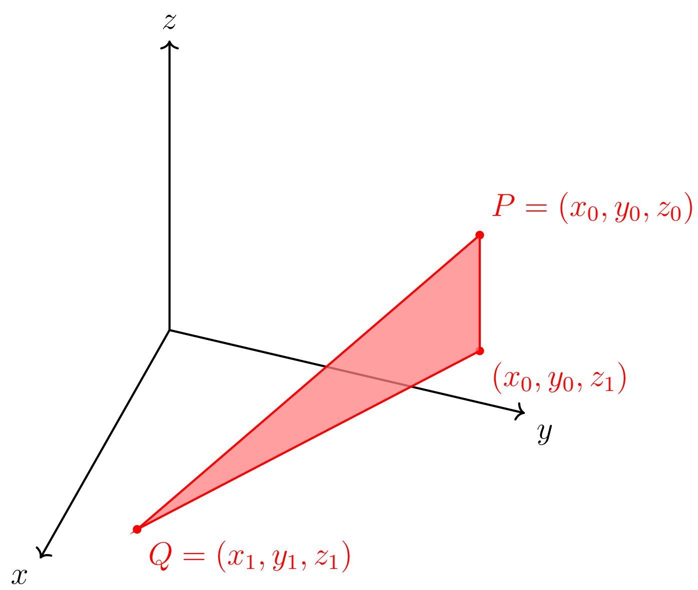


Figure 1.10: The triangle whose hypotenuse we care about.

Now we have a little bit of a problem: to figure out the length of this hypotenuse, we need to know the lengths of the other sides of a triangle. In two dimensions this was relatively straight forward, but in three dimensions it seems to be a little less obvious. In order to figure out what these are, we imagine our triangle as sitting inside of a box where the points P and Q are at opposite ends of the box, as pictured in Figure 1.11 on the next page.

Because of the way we've constructed this box, having P and Q in opposite corners, the dimensions of the box are $\Delta x \times \Delta y \times \Delta z$. (Convince yourself of this fact by finding the coordinates of all the corners of the box.) This means that one of the legs of our triangle (the right-most leg

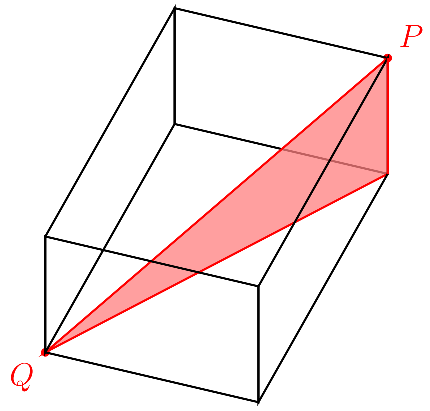


Figure 1.11: Our triangle sitting inside of a box.

in the pictures above) has length Δz. We just need to figure out what the length of the other leg is. To do this we notice that this leg is the hypotenuse of another triangle: the one in blue in Figure 1.12 on the following page.

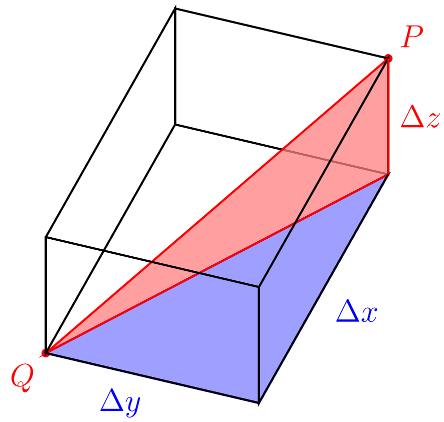


Figure 1.12: One leg of the red triangle is also the hypotenuse of the blue triangle.

As the other two sides of the blue triangle are on the edges of the box, we know their lengths are Δx and Δy. Thus the hypotenuse of the blue triangle is given by $\sqrt{\Delta x^2 + \Delta y^2}$, and so the legs of the red triangle have lengths Δz and $\sqrt{\Delta x^2 + \Delta y^2}$. We are thus able to say that the length of the hypotenuse of the red triangle (aka the distance from P to Q) is given by

$$\sqrt{\left(\sqrt{\Delta x^2 + \Delta y^2}\right)^2 + \Delta z^2}$$
$$= \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$
$$= \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}$$

This is the ***distance formula*** for calculating the distance between two points, $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$, in 3-space. As in two-dimensions, this distance is often denoted $|PQ|$ or $d(P, Q)$.

Spheres

We can use our three-dimensional distance formula above to determine the equation of a sphere in 3-space. As a warm up, though, let's first recall how to derive the equation of a circle in the plane.

Equation of a circle in the plane

In two dimensions, a ***circle*** centered at P with radius $r > 0$ is the collection of all the points Q in $\mathbb{R}^2$ whose distance from P is equal to r. Thus, as a set of points, such a circle is given by

$$\left\{Q \in \mathbb{R}^2 \,\middle|\, |PQ| = r\right\}.$$

Written in terms of xy-coordinates, if we suppose $P = (x_0, y_0)$ and $Q = (x, y)$, then this set may be written as

$$\left\{(x, y) \in \mathbb{R}^2 \,\middle|\, \sqrt{(x - x_0)^2 + (y - y_0)^2} = r\right\}.$$

Normally when we want to consider the set of all the points satisfying some equation, we just write the equation and don't bother with the set notation. Our equation is thus

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = r.$$

Squaring both sides of the equation this becomes

$$(x - x_0)^2 + (y - y_0)^2 = r^2,$$

which is the familiar equation of a circle of radius r centered at (x_0, y_0).

Equation of a sphere in 3-space

To define a sphere in three dimensional space, we will mimic the derivation of a circle in two dimensions. By definition, a ***sphere*** centered at a point $P \in \mathbb{R}^3$ with radius $r > 0$ is the set of all points $Q \in \mathbb{R}^3$ which are distance r from P. That is,

$$\left\{Q \in \mathbb{R}^3 \,\middle|\, |PQ| = r\right\}.$$

If we suppose $P = (x_0, y_0, z_0)$, then using our distance formula from above, this becomes

$$\left\{(x, y, z) \in \mathbb{R}^3 \middle| \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = r\right\}.$$

Dropping the set notation and just writing the equation we have

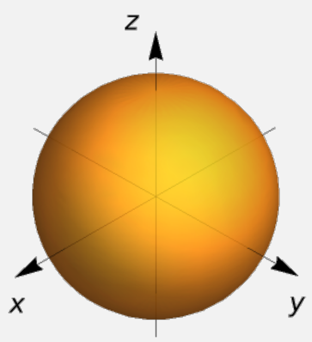
$$\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} = r,$$

and if we square both sides this turns into

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

Example 1.1.
The sphere with radius 1 centered at the origin (sometimes called the *unit sphere*) is given by the equation

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



Example 1.2.
The sphere of radius 2 centered at the point $(1, 0, -2)$ is given by the equation
$$(x-1)^2 + y^2 + (z+2)^2 = 4.$$

Standard Form

The nice thing about this formula is that if someone just hands us an equation like
$$(x+2)^2 + (y-1)^2 + (z+1)^2 = 25,$$
we immediately know that this equation represents a sphere centered at $(-2, 1, -1)$ with radius 5.

This way of representing a sphere is called *standard form*.

Notice in Example 1.2, we could FOIL-out each of the terms $(x-1)^2$ and $(z+2)^2$:

$$(x-1)^2 = x^2 - 2x + 1$$
$$(z+2)^2 = z^2 + 4z + 4$$

If we plug this into our original equation,

$$(x-1)^2 + y^2 + (z+2)^2 = 4.$$

and combine like-terms, we suddenly have the uglier equation

$$x^2 + y^2 + z^2 - 2x + 4z = -1.$$

If we were just given this equation, it's not immediately clear that the equation represents a sphere. Even if we somehow knew this was the equation of some sphere, now the equation becomes which sphere is it? What's the center? What's the radius?

It'd be nice if we had some way of taking an equation like

$$x^2 + y^2 + z^2 - 2x + 4z = -1.$$

and turning it back into standard form to find the center and radius. The trick to doing this is to complete the square. Let's briefly recall how completing the square works.

Theorem 1.2 (Completing the Square).
The quadratic polynomial

$$x^2 + bx + c$$

can be rewritten as

$$\left(x + \frac{b}{2}\right)^2 - \frac{b^2}{4} + c.$$

Example 1.3.
Complete the square to put the following equation in standard form:

$$x^2 + y^2 + z^2 - 6x + 2y - 4z = 11.$$

The idea is that we'll complete the square three times, one for each variable, and then move all the constants to one side. As a preliminary step, lets group all the x's together, all the y's together, and all the z's together.

$$x^2 - 6x + y^2 + 2y + z^2 - 4z = 11$$

Now we want to complete the square on $x^2 - 6x$, on $y^2 + 2y$ and on $z^2 - 4z$. In the case of $x^2 - 6x$, we take half of the coefficient of the middle term, and square it. Here, $x^2 - 6x = x^2 - 6x + 0$, so $6x$ is the "middle" term. Half of this is 3, and squaring gives us 9. We *want* to add 9 to our expression because then we'd have

$$x^2 - 6x + 9 = (x - 3)^2$$

However our $x^2 - 6x$ stuff is part of an equation: we can't just magically add a 9 to one side without adding it to the other side. So, completing the square just on the "x-part" of the equation, we have

$$\begin{aligned} & x^2 - 6x + y^2 + 2y + z^2 - 4z = 11 \\ \Longrightarrow\ & x^2 - 6x + 9 + y^2 + 2y + z^2 - 4z = 11 + 9 \\ \Longrightarrow\ & (x - 3)^2 + y^2 + 2y + z^2 - 4z = 20 \end{aligned}$$

So we're one third of the way there. We just need to complete the square on the y-part and z-part of our equation.

For the y-part we have $y^2 + 2y$. We want to add half the middle term squared, so we want to add 1, because then we know $y^2 + 2y + 1 = (y+1)^2$. Again, since this is an equation, so we have to add 1 to each side of the equation:

$$\begin{aligned} &(x-3)^2 + y^2 + 2y + z^2 - 4z = 20 \\ \Longrightarrow\ &(x-3)^2 + y^2 + 2y + 1 + z^2 - 4z = 20 + 1 \\ \Longrightarrow\ &(x-3)^2 + (y+1)^2 + z^2 - 4z = 21. \end{aligned}$$

Finally we need to complete the square on the z-part of the equation. For $z^2 - 4z$ we want to add a 4 because $z^2 - 4z + 4 = (z-2)^2$.

$$\begin{aligned} &(x-3)^2 + (y+1)^2 + z^2 - 4z = 21 \\ \Longrightarrow\ &(x-3)^2 + (y+1)^2 + z^2 - 4z + 4 = 21 + 4 \\ \Longrightarrow\ &(x-3)^2 + (y+1)^2 + (z-2)^2 = 25 \end{aligned}$$

Now that we have our equation in standard form, we see that this is the equation of a sphere with radius 5 centered at the point $(3, -1, 2)$.

Example 1.4.
Find the radius and center of the sphere given by the equation

$$x^2 + z^2 + 2x = -49 - y^2 - 14z$$

Of course, the very first step is to get all of the variables on one side, and all the constants on the other side, then group the x's, y's, and z's together:

$$\begin{aligned} &x^2 + z^2 + 2x = -49 - y^2 - 14z \\ \Longrightarrow\ &x^2 + 2x + y^2 + z^2 + 14z = -49. \end{aligned}$$

Completing the square for the x's gives

$$\begin{aligned} & x^2 + 2x + y^2 + z^2 + 14z = -49 \\ \Longrightarrow\ & (x+1)^2 + y^2 + z^2 + 14z = -48. \end{aligned}$$

Notice we don't actually need to do completing the squares to the y's: the stuff with y's is already squared!

Completing the square for the z's gives

$$\begin{aligned} & (x+1)^2 + y^2 + z^2 + 14z = -48. \\ \Longrightarrow\ & (x+1)^2 + y^2 + (z+7)^2 = 1. \end{aligned}$$

So we see that this is the equation of a sphere with radius 1 centered at the point $(-1, 0, -7)$.

The Coordinate Planes and their Translates

Recall that in the case of the two-dimensional plane, the x-axis is given by the equation $y = 0$; and the y-axis is given by the equation $x = 0$. Recall that writing an equation like $y = 0$, in two dimensions, is really just short-hand for the set

$$\left\{(x, y) \in \mathbb{R}^2 \middle| y = 0\right\}.$$

What if we're in three dimensions? Does the equation $y = 0$ still represent the x-axis? Let's consider what the equation $y = 0$ actually represents in three dimensions:

$$\left\{(x, y, z) \in \mathbb{R}^3 \middle| y = 0\right\}.$$

So, we're asking for all of the points (x, y, z) where $y = 0$: the x's *and* z's can both be anything else: points like $(-14, 0, \pi^2)$ live inside this set.

A portion of the collection of points, in $\mathbb{R}^3$, given by the equation $y = 0$ is shaded in blue in Figure 1.13. To (hopefully) make the picture a little bit easier to understand, the negative parts of the x-, y-, and z-axes are also drawn.

This collection of points is called the xz-*plane*, and consists of all the (x, y, z)-points you can get by forcing the y-coordinate to be zero, but letting the x- and z-coordinates be anything you'd like. If you imagine the

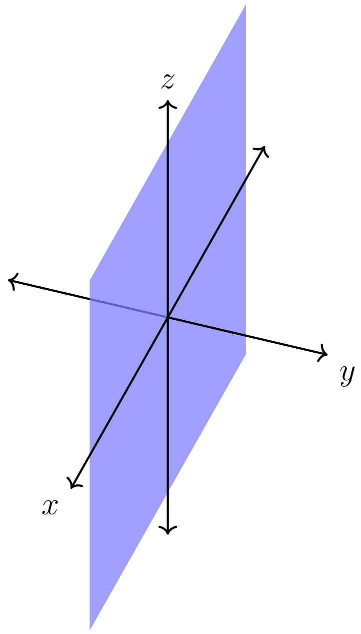


Figure 1.13: $y = 0$

three-dimensional coordinates in a room where the floor is the xy-plane and the origin is in one of the corners of the room, then the xz-plane is one of the walls.

The xy-plane, in three dimensions, is of course given by the equation $z = 0$: The xy-plane, as a set, is $\left\{(x, y, z) \in \mathbb{R}^3 \middle| z = 0\right\}$. This is given in Figure 1.14.

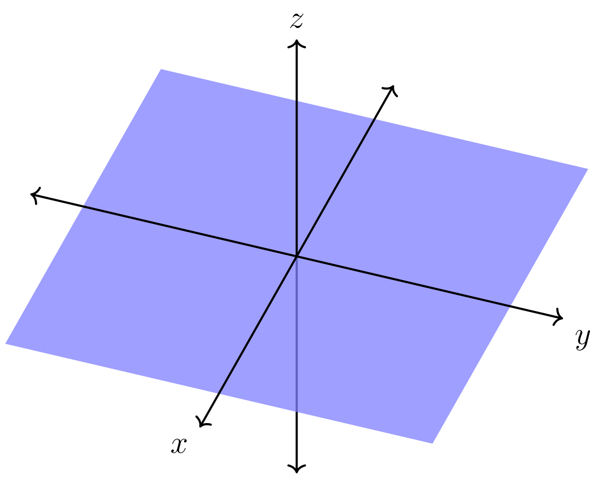


Figure 1.14: The xy-plane is given by $z = 0$.

Finally, the set of points satisfying $x = 0$ is called the *yz-plane* and is shown in Figure 1.15.

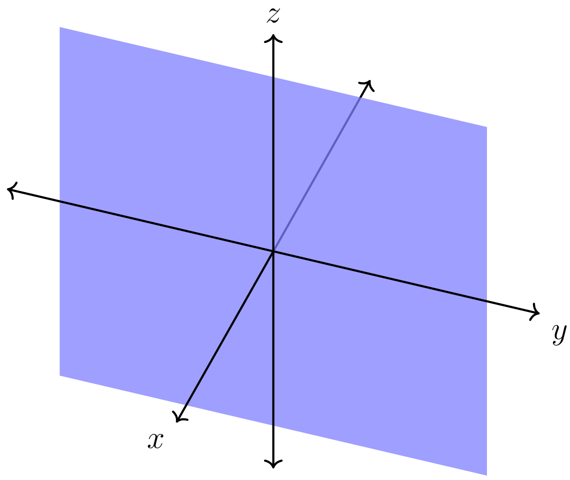


Figure 1.15: The yz-plane is given by $x = 0$

These three planes, the xy-plane, the yz-plane, and the xz-plane, are called the *coordinate planes*: just like how the lines $x = 0$ and $y = 0$, in two dimensions, are called the *coordinate axes*.

For the time being, let's denote these planes by P_{xy} for the xy-plane; P_{xz} for the xz-plane; and P_{yz} for the yz-plane.

$$P_{xy} = \left\{(x, y, z) \in \mathbb{R}^3 \middle| z = 0\right\}$$
$$P_{xz} = \left\{(x, y, z) \in \mathbb{R}^3 \middle| y = 0\right\}$$
$$P_{yz} = \left\{(x, y, z) \in \mathbb{R}^3 \middle| x = 0\right\}$$

There are three useful maps, called *projections*, which take a point in $\mathbb{R}^3$ and map it to the closest point in each of the projection planes. Suppose that ρ_{xy} is the projection to the xy-plane. Then ρ_{xy} is a map $\rho_{xy} : \mathbb{R}^3 \to P_{xy}$. The way this map works is that it basically just takes the z-coordinate of a point, and replaces it with a zero:

$$\rho_{xy}(a, b, c) = (a, b, 0).$$

This takes in any point in $\mathbb{R}^3$, and spits out a point that lives on the xy-plane, P_{xy}. In Figure 1.16 several examples are shown. This seems like a very simple operation to perform (and it is), but the important geometric property of this map is that the point $\rho_{xy}(P)$ is point on P_{xy} that is as close

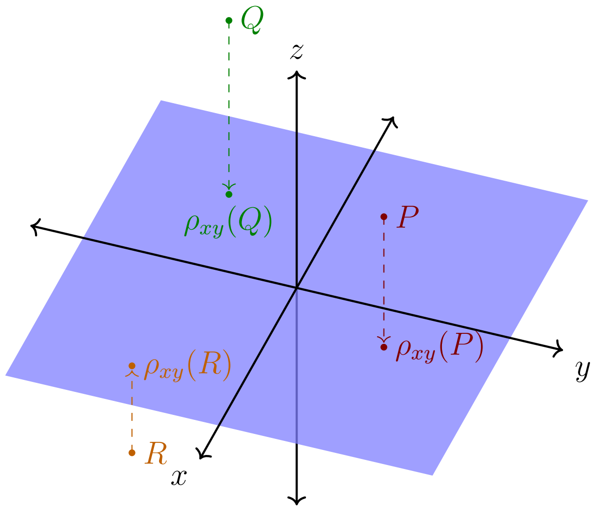


Figure 1.16: Each coordinate plane comes with a projection map.

as possible to P. (To make sure you're understanding what's going on, you should think a little bit about why this is: how do you know there isn't another point on P_{xy} that's just as close, or closer?)

The projections ρ_{xz} and ρ_{yz} are defined in the same way: they replace one of the coordinates with zero.

$$\rho_{xz}(a,b,c) = (a,0,c)$$
$$\rho_{yz}(a,b,c) = (0,b,c)$$

The important geometric property of ρ_{xz} is that for any point $P \in \mathbb{R}^3$, $\rho_{xz}(P)$ is the point on the xz-plane which is as close as possible to P. Similarly, $\rho_{yz}(P)$ is the point on the yz-plane which is as close as possible to P.

Notice that if we were to take the equation of a coordinate plane, say the yz-plane $(x = 0)$, and modify it just a little bit by changing the number 0 to something else (like maybe $x = 3$) then we get another plane. This plane is parallel to the original plane, but slid three units in the x-direction. See Figure 1.17.

Some Simple Surfaces

Finally, we consider some very simple examples of surfaces. Let's start off with a very simple question: what does the equation $y = x$ represent in three dimensions?

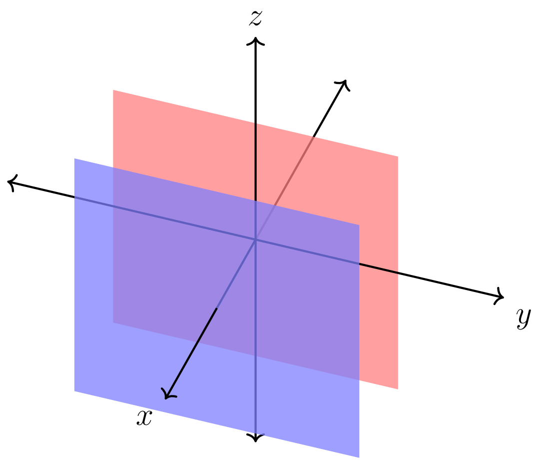


Figure 1.17: The plane $x = 3$, in blue, is parallel to the yz-plane, in red.

We know that, in two dimensions, $y = x$ is just a line. In three-dimensions we have this extra z-coordinate we can play with, but there's no restriction on what z has to be: it can be anything we want. So, we take the line $y = x$ in the xy-plane, and can move it up or down as much as we like. This gives us yet another plane. See Figure 1.18.

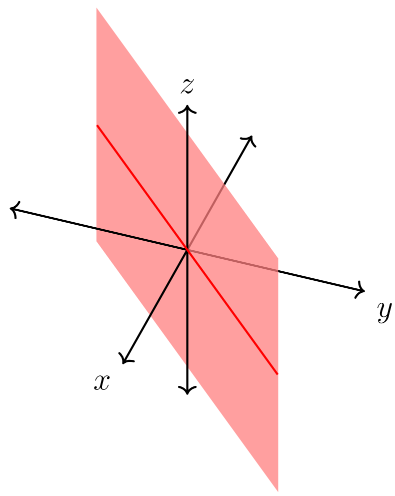
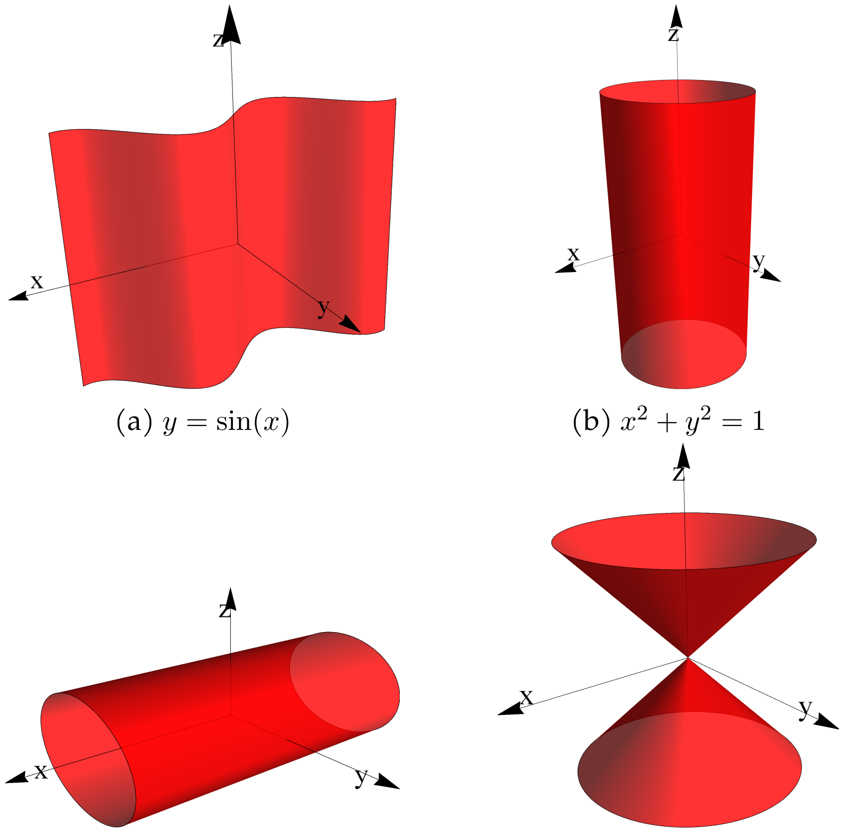


Figure 1.18: The equation $y = x$ gives a plane in $\mathbb{R}^3$. (The line $y = x$ in the xy-plane is darkened.)

In general, the graph of any function $y = f(x)$ in the xy-plane turns into a surface in $\mathbb{R}^3$. There are no restrictions on z, so pick up the curve, then move it straight up or down as much as you'd like. The graph $y =$

$\sin(x)$, for example, becomes an infinitely long wavy wall, like a giant piece of corrugated cardboard. See Figure 1.19a.



(a) $y = \sin(x)$

(b) $x^2 + y^2 = 1$

(c) $4y^2 + 9z^2 = 36$

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

Figure 1.19: Some examples of surfaces.

Of course, we don't have to limit ourselves to graphs of functions. We can consider other, more interesting, things like equations. In two dimensions, the equation $x^2 + y^2 = 1$ gives us a circle. In three dimensions we get an infinite cylinder: we take the circle in the xy-plane and move it up or down as much as we'd like. See Figure 1.19b.

Finally, we don't have to limit ourselves to equations that involve only x's and y's: we can throw some z's in there too. For example, consider an equation like $4y^2 + 9z^2 = 36$. This is an ellipse in the yz-plane. In three-dimensional space, though, we have to ask what the x-coordinate can be. Since this equation places no restriction on what x could or could not be, the x-value can be anything. Thus we can take that ellipse in the yz-plane, then push and pull it back and forth along the x-axis as much as we want. We then have an infinitely long, elliptical pipe as in Figure 1.19c.

And of course, to finish things off, we could have an equation that involves all three of our variables. Consider the equation

$$z^2 = x^2 + y^2$$

What does the set of points satisfying this equation look like? To figure this out, let's suppose that we only look at the points satisfying this equation which also satisfy the equation $z = 1$. That is, we have (x, y, z)-triples which look like $(x, y, 1)$ where $x^2 + y^2 = 1$. Of course, the equation $z = 1$ is just a plane (parallel to the xy-plane, just moved up one unit). Now, on this plane we consider the points satisfying $x^2 + y^2 = 1$, which is just a circle of radius one centered at the origin. This circle is one "slice" of our surface.

Let's try the same thing again, but suppose that $z = 2$. Then we're looking at the plane $z = 2$, and inside of that plane consider the points that satisfy the equation $x^2 + y^2 = 4$. Why "= 4" this time? Because $x^2 + y^2 = z^2$, but we are assuming $z = 2$. This, in the $z = 2$ plane, gives us a circle of radius 2 centered at the origin. We can repeat this idea for each value of z, and find that we get lots and lots of circles of different radii. (In particular, if z is negative, z^2 is positive, and so this still works for values of $z < 0$.)

We get our surface by gluing all of these circles together, and this gives us two infinite cones whose tips meet at the origin, as in Figure 1.19d.

1.2 Higher dimensional space

While we are primarily going to be concerned with two- and three-dimensional space in this course and won't spend much time discussing higher dimensional space. The reasons for this have more to do with psychology than mathematics, though, and so we will go ahead and briefly mention higher dimensional space and why you might be interested in such spaces.

The two-dimensional Euclidean plane can be represented as the set of all ordered pairs (x, y) of real numbers, and this set is denoted $\mathbb{R}^2$. Three-dimensional space is represented as all ordered triples (x, y, z) of real numbers, and this set is denoted $\mathbb{R}^3$. Given these two examples, it's probably not hard to believe that we could think of four-dimensional space as the set of ordered "quadruples" of real numbers, (w, x, y, z), and this space would be denoted $\mathbb{R}^4$. This description doesn't really tell us anything about how to visualize four-dimensional space, but it seems like an obvious generalization of how we represent two- and three-dimensional space. There's also no great reason to stop this construction at four dimensions. We may not be able to easily visualize it, but we *can* easily define five-dimensional space, six-dimensional space, seven-dimensional space, and so on.

In general, we will define ***n-dimensional space*** as the set of all ordered ***n-tuples*** of real numbers. That is, for us n-dimensional space will consist of all possible ways of writing down a list of n real numbers. When n is small (two, three, or maybe four), we'll use the letters at the end of the alphabet for the individual ***coordinates*** of our points. When is large, though, we will denote the coordinates by x_1, x_2, x_3, and so on, up through x_{n-1} and x_n. A point in this n-dimensional space then can be written as

$$(x_1, x_2, x_3, ..., x_{n-1}, x_n)$$

and we write $\mathbb{R}^n$ as a short-hand for this set of all possible n-dimensional points.

It might be hard to believe right now, but all of the mathematics we are about to develop will easily carry over to these higher-dimensional spaces. We will usually do problems and examples and discuss applications in two and three dimensions, but this is really just so that we can draw pictures and more easily visualize what's happening.

Remark.
It's even possible to consider infinite-dimensional space, but here things do actually start to get more involved. That is, while the jump from one dimension to two dimensions or three dimensions is basically just like the jump from two dimensions to five-thousand dimensions, the jump to infinitely many dimensions has some surprising technical subtleties that we aren't going to tackle. As such we will never discuss infinite dimensional spaces in this class.

Even though we can't really visualize what's happening in five, ten, or seventy-eight dimensions, there are real world, practical problems that really do take place in these spaces. In general, many real-world problems in science, engineering, economics, business, medicine, and many other fields often rely on several different pieces of information. For example, a business interested in modeling the growth of their expenses over the coming years may need to know about the cost of raw materials they require, the number of employees they will have, the price of maintenance on any machinery they own, the cost of keeping attorneys on retainer, the cost of their taxes, and so on. Each one of these different pieces of information gives us a variable; that is, each piece of information gives us one coordinate, and so one dimension. We see, then, that the number of pieces of information determine the number of dimensions of the space in which our problem lives. In the case of the business mentioned above, there are five particular pieces of information that are mentioned, and so the problem of modelling the business' expenses takes place in five-dimensional space.

As another example, an engineer designing a certain type of robot may be interested in writing a computer program that controls the robot, and so needs to represent all possible configurations of the robot's various components. If the robot contains an arm attached to a motor at one end, we need one piece of information to determine the state of the robot's arm (e.g., the angle at which the arm is held at). The robot may have a second arm attached at the end of the first arm, and so a second piece of information is required to tell us the angle between the arms. A complicated robot may have very many arms attached end-to-end in this way, as in Figure 1.20 and each arm gives us a piece of information. The set of all possible configurations of the arm, then, is n-dimensional where n is the number of segments in the arm. For example, if the arm contains six segments, then the space of all configurations is six-dimensional.

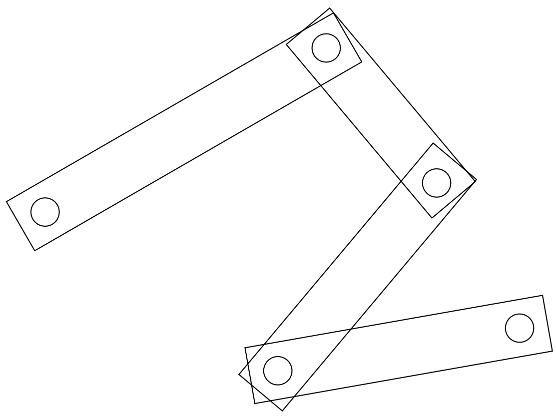


Figure 1.20: One configuration of a robot's arm. This arm consists of four segments, so the space of configurations is four-dimensional.

1.3 Vectors

Physical motivation

Newton's first two laws of motion basically say that moving objects don't just magically change speed or direction. Instead, for such a change to occur, a *force* must act on an object. For example, suppose you're an astronaut in outer space, floating around in zero gravity.[2] Now suppose you notice that some little bit of a meteor is floating towards you. Newton's first two laws say that this little meteor chunk is going to continue moving towards you in a straight line, and at a constant speed, until something happens to change that. Thus if you want that little chunk of meteor to change speed or direction, you have to do something to make that happen. So as the meteor gets close to you, you give it a little push to the side so that it won't ram into your spaceship. What you've done is apply a force (the push) to the object, and that force made the object change direction.

There are basically two important properties of the push you gave the meteor: how hard you pushed it, and the direction pushed it. Any force (anything that makes an object change speed or direction) has these two properties: it has a direction (where are you trying to make the object go), and it has a magnitude (how hard the force pushes/pulls).

It turns out that many other physical quantities we care about have this property of being defined by a direction and a magnitude. For example, velocity, momentum, and torque all are described by a direction and a magnitude. Since this idea of pairing direction and magnitude comes up so frequently, we give these quantities that have both a magnitude and a direction a special name: we call them ***vectors***.

Since a "vector" is a new, special quantity that we're introducing, we'll need some way of distinguishing vector quantities from the usual numerical quantities we typically deal with. For this reason we sometimes call real numbers ***scalars***. That is, a scalar is just a number, but a vector is this quantity that has both direction and magnitude. (We will see the reason for the word *scalar* instead of just saying *number* soon.)

[2]You may know there's no such thing as zero gravity, but just for the sake of this example, suppose zero gravity actually existed.

Geometric point of view

In two dimensions

Geometrically, a vector is just an arrow. For the moment let's just focus on two-dimensional vectors, and so we mean an arrow in the plane. The "direction" of the vector is then (not surprisingly) the direction of this arrow; it just says where does the arrow point. The "magnitude" of the vector is the length of the arrow. Some examples are given in Figure 1.21. The point at the tip of the arrow is called the *tip* of the vector, and the point at the opposite end of the arrow is called the *tail*.

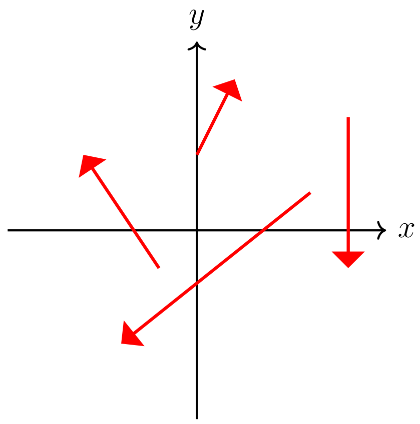


Figure 1.21: Some vectors in two dimensions.

Just like any other mathematical object, it's convenient to give vectors a name when referring to them, so we don't have to say something like "the arrow that's 3 units long and points $45°$ above the x-axis" each time we refer to the vector. Textbooks usually denote vectors in one of two ways, either by making it bold, or putting an arrow over it. That is, if we want to signify that "u" is the name of a vector, we might write **u** or $\vec{u}$. In hand-written notes, people usually prefer the $\vec{u}$ notation because it's a lot easier to draw an arrow than it is to make something bold. In these notes we will denote vectors by writing an arrow over them.

One of the important properties of vectors, which may seem strange at first, is that you can slide them around to your heart's content and you don't change the vector. For example, all of the arrows in Figure 1.22 represent the same vector. The reason we consider all of the arrows in Figure 1.22 as representing the same quantity is that we are explicitly defining a vector as something defined by its direction and magnitude; notice we *are not* saying the location of the vector is important. Thus if two arrows have the same length and b

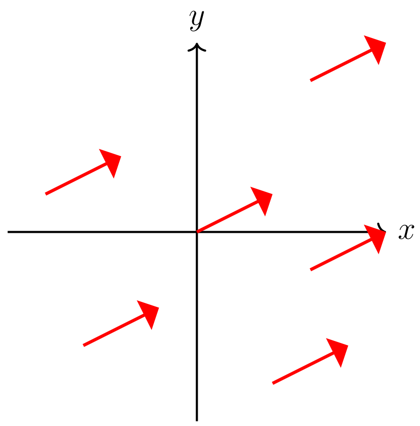


Figure 1.22: Translating (sliding) a vector around doesn't change it. All of the vectors in this picture are the same.

In three dimensions

Vectors in three dimensions are the same idea as vector in two dimensions, we just have another dimension for our arrows to move around in. See Figure 1.23.

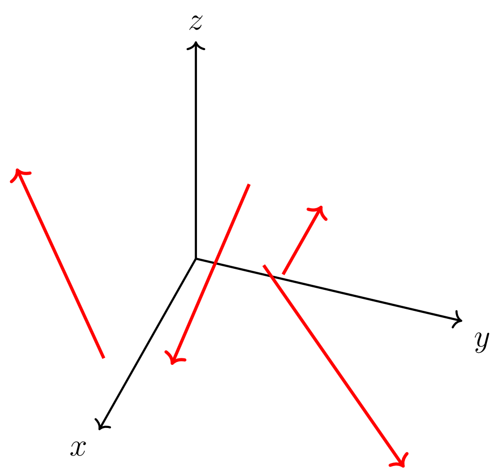
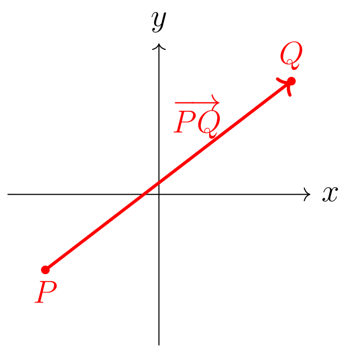


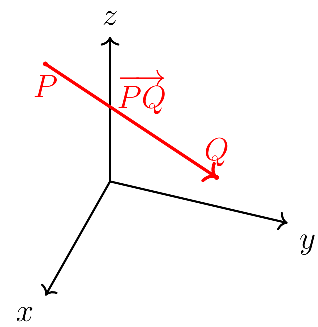
Figure 1.23: Some vectors in $\mathbb{R}^3$.

Displacement

Let's say that P and Q two points, either both in the plane, or both in 3-space. The *displacement vector* from P to Q, is just the arrow (vector) that starts at P and ends at Q. This is denoted $\overrightarrow{PQ}$. See Figures 1.24a and 1.24b for two- and three-dimensional examples.



(a) 2-dim. displacement.



(b) 3-dim. displacement.

Figure 1.24: Displacement vectors.

Vector addition

One of the most important properties of vectors is that two vectors can be combined together to produce a new vector. This is accomplished by an operation called *vector addition*. There are two different ways we can describe vector addition graphically. The first way is called the *triangle law*.

To add $\vec{v}$ and $\vec{u}$ together, we slide the vectors around so that the tail of $\vec{v}$ is at the tip of $\vec{u}$. Then $\vec{u} + \vec{v}$ is the vector that connects the tail of $\vec{u}$ to the tip of $\vec{v}$. See Figure 1.25.

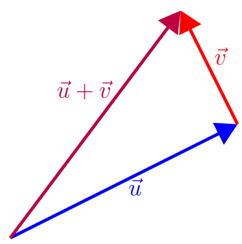


Figure 1.25: The triangle law.

The physical interpretation of the triangle law is that if we apply two forces, $\vec{F}_1$ and $\vec{F}_2$, to some object, this is the same thing as applying the force $\vec{F}_1 + \vec{F}_2$.

The other way we can add the vectors together is to use the *parallelogram law*. We take the vectors $\vec{u}$ and $\vec{v}$ and use them as the sides of a parallelogram. That is, we take two copies of each vector and slide them around until we get a parallelogram. We then add in a vector from the corner of the parallelogram with two tails, to the corner with two tips. See Figure 1.26.

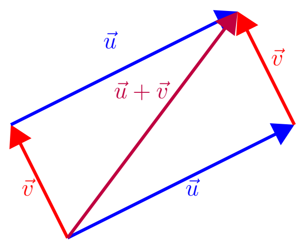


Figure 1.26: The parallelogram law.

Let's suppose that $\vec{u}$ and $\vec{v}$ point in opposite directions and have the same length. What is $\vec{u} + \vec{v}$ supposed to be? Thinking about these vectors as forces, we give an object a push in some direction, then give a push of the same magnitude in the opposite direction. These two things cancel each other out. The vector we get by adding two such opposites is called the *zero vector* and denoted $\vec{0}$. This is the unique vector that has no direction and no length: it's the "nothing" vector.

Scalar multiplication

The second important operation we can perform with vectors is called *scalar multiplication*. Scalar multiplication takes a scalar (real number) and a vector, and combines them together to get another vector. If $\lambda \in \mathbb{R}$ is our scalar and $\vec{v}$ is the vector, the product of λ and $\vec{v}$ is denoted $\lambda \cdot \vec{v}$, or simply $\lambda\vec{v}$.

To describe scalar multiplication we consider three separate cases. First suppose that $\lambda > 0$, then $\lambda\vec{v}$ is the vector that has the same direction as $\vec{v}$, but with length multiplied by λ. For example, $2\vec{v}$ points in the same direction as $\vec{v}$, but has twice the length. $\frac{1}{3}\vec{v}$ points in the same direction as $\vec{v}$, but has one third the length. See Figure 1.27.

In the case of $\lambda < 0$, we do the same thing, except we also flip the vector around to point in the opposite direction. That is, if $\lambda < 0$, then $\lambda\vec{v}$

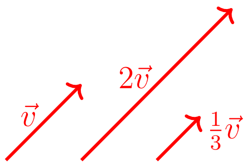


Figure 1.27: Scalar multiplication with $\lambda > 0$.

points in the opposite direction as $\vec{v}$, and the length is multiplied by $|\lambda|$. See Figure 1.28.

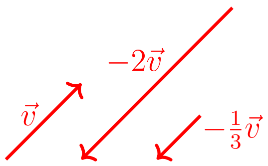


Figure 1.28: Scalar multiplication with $\lambda < 0$.

If $\lambda = 0$, then $\lambda\vec{v}$ is defined to be the zero vector: $0 \cdot \vec{v} = \vec{0}$.

The reason real numbers are called scalars when we're talking about vectors is because they *scale* vectors: they make them longer or shorter (and possibly make them point in the opposite direction).

By convention, we write $-\vec{v}$ to mean $-1 \cdot \vec{v}$ and $\vec{u} - \vec{v}$ really means $\vec{u} + -\vec{v}$.

We say that vectors $\vec{u}$ and $\vec{v}$ are *parallel* if they point in the same (or opposite) directions. This is the same thing as saying $\vec{u} = \lambda\vec{v}$ for some non-zero $\lambda \in \mathbb{R}$.

Algebraic point of view

In the last section we defined vectors as arrows. These arrows are a very convenient way of thinking about vectors and make it easy to draw pictures and reason geometrically, and throughout the semester we'll be drawing lots of these arrows. However, these arrows have a problem: they're pretty clumsy to work with for longer, more involved calculations. We need some way of taking the arrows and turning them into things we can manipulate more directly: we want to somehow replace the arrow with a collection of numbers. To do this we take advantage of the fact that you can slide vectors around as much as you'd like and you don't change the vector.

What we'll do is slide the vector so that it's tail is at the origin. The tip of the arrow will then be at some point, and we'll use that point's coordinates as *components* of the vector, which we list between two angle brackets, $\langle$ and $\rangle$. That is, if $\vec{u}$ is some vector and we slide it so that its tail is at the origin and its tip is then at the point (a, b, c), we write

$$\vec{u} = \langle a, b, c \rangle .$$

It should be noted that other people use different notations for vectors. Some common ones are the following.

$$\begin{aligned} \vec{u} &= [a, b, c] \\ \vec{u} &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ \vec{u} &= a\vec{\imath} + b\vec{\jmath} + c\vec{k} \\ \vec{u} &= a\widehat{\imath} + b\widehat{\jmath} + c\widehat{k} \end{aligned}$$

You should be aware of each of these notations: each one is very common, and useful in certain situations. We will go back and forth between the various notations.

The $\vec{\imath}$, $\vec{\jmath}$, and $\vec{k}$ listed above are sometimes called the *standard basis vectors*. These are unit vectors that point in the positive x, y, and z directions, respectively. In components,

$$\begin{aligned} \vec{\imath} &= \langle 1, 0, 0 \rangle \\ \vec{\jmath} &= \langle 0, 1, 0 \rangle \\ \vec{k} &= \langle 0, 0, 1 \rangle . \end{aligned}$$

Vector addition and scalar multiplication

Representing vectors algebraically, our two basic vector operations are about as easy to describe as could be. If $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$, then their sum is given by

$$\begin{aligned} \vec{u} + \vec{v} &= \langle u_1 + v_1, u_2 + v_2, u_3 + v_3 \rangle \\ &= \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix} \\ &= (u_1 + v_1)\vec{\imath} + (u_2 + v_2)\vec{\jmath} + (u_3 + v_3)\vec{k}. \end{aligned}$$

The scalar product $\lambda\vec{u}$ is given by

$$\begin{aligned}\lambda\vec{u} &= \langle \lambda u_1, \lambda u_2, \lambda u_3 \rangle \\ &= \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \\ \lambda u_3 \end{bmatrix} \\ &= \lambda u_1 \vec{\imath} + \lambda u_2 \vec{\jmath} + \lambda u_3 \vec{k}\end{aligned}$$

Displacement

Recall that if P and Q are two points (either both in $\mathbb{R}^2$ or both in $\mathbb{R}^3$), then $\overrightarrow{PQ}$ denotes the displacement vector from P to Q. We want to figure out what the components of this vector are. Just to keep things simple, let's start off in two dimensions. Say $P = (x_0, y_0)$ and $Q = (x_1, y_1)$. Let $\vec{P}$ and $\vec{Q}$ be the vectors from the origin to P and Q, respectively. Our goal is to find the components of the red vector in Figure 1.29.

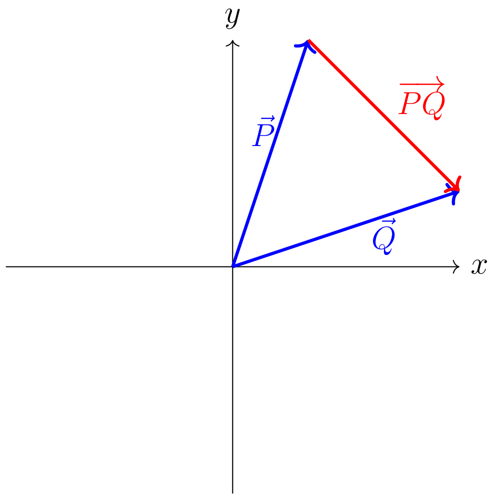


Figure 1.29: We want to find the displacement vector from P to Q.

One thing to notice about this displacement vector is that $\vec{Q} - \overrightarrow{PQ} = \vec{P}$. (If we make $\overrightarrow{PQ}$ point in the opposite direction, then $\vec{P}$ is the sum of $\vec{Q}$ and $-\overrightarrow{PQ}$ by the triangle law.) Thus, solving for $\overrightarrow{PQ}$ gives us

$$\overrightarrow{PQ} = \vec{Q} - \vec{P}.$$

Thus if $P = (x_0, y_0)$ and $Q = (x_1, y_1)$, then

$$\overrightarrow{PQ} = \langle x_1 - x_0, y_1 - y_0 \rangle .$$

The exact same idea works in three dimensions: if $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$, then

$$\overrightarrow{PQ} = \langle x_1 - x_0, y_1 - y_0, z_1 - z_0 \rangle .$$

Magnitude

Our algebraic notation makes it very easy to figure out the magnitude of the vector. We slide the vector around so that its tail is at the origin and its tip is at the point (u_1, u_2). Then the length of the vector (the arrow) is just the length of the line segment from $(0,0)$ to (u_1, u_2), which is of course just $\sqrt{u_1^2 + u_2^2}$. In three dimensions the exact same idea works, and so the length of $\langle u_1, u_2, u_3 \rangle$ is $\sqrt{u_1^2 + u_2^2 + u_3^3}$. This quantity is sometimes called *magnitude*, or *length*, or the *norm* of the vector. Two common notations for the length of the vector are

$$|\vec{u}| \text{ and } \|\vec{u}\|.$$

The magnitude has various convenient properties, each of which is easy to justify using our algebraic notations.

Theorem 1.3 (Properties of magnitudes)**.**
Let $\lambda \in \mathbb{R}$ be a scalar, and let $\vec{u}$ and $\vec{v}$ be vectors (both of which are either two-dimensional or both of which three-dimensional). The following properties.

(*i*) $\|\lambda \vec{v}\| = |\lambda| \|\vec{v}\|$

(*ii*) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

(*iii*) $\|\vec{u} + \vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$ *if and only if $\vec{u}$ and $\vec{v}$ are parallel.*

(Properties (ii) and (iii) above are sometimes called the *triangle identity*.)

Equality of vectors

We say that two vectors, $\vec{u}$ and $\vec{v}$, are *equal* if they have the same dimension (they're both two-dimensional, or they're both three-dimensional) and their components are equal. That is, in the two-dimensional case, if $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$, then $\vec{u} = \vec{v}$ if

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2. \end{aligned}$$

The three-dimensional case is the same, except that we also require that $u_3 = v_3$. Notice that this is the same thing as saying $\vec{u} - \vec{v} = \vec{0}$.

Algebraic properties of vectors

The operations of vector addition and scalar multiplication satisfy several useful properties, listed below.

Theorem 1.4.
Let $\lambda, \mu \in \mathbb{R}$, and let $\vec{u}$, $\vec{v}$, and $\vec{w}$ be three vectors (both of which are either two-dimensional, or three-dimensional). Then the following properties hold.

(*i*) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

(*ii*) $\vec{u} + \vec{0} = \vec{v}$

(*iii*) $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$

(*iv*) $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$

(*v*) $(\lambda\mu)\vec{v} = \lambda(\mu\vec{v})$

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

(*vii*) $\vec{u} - \vec{v} = \vec{u} + (-1 \cdot \vec{v})$

(*viii*) $\vec{v} - \vec{v} = \vec{0}$

Examples

Now we'll do some semi-complicated examples to hopefully make everything as concrete as possible.

Example 1.5.
Suppose $\vec{u}$ is the displacement vector from $P = (3, 0, -1)$ to $Q = (0, 4, -3)$, and $\vec{v}$ is the displacement vector from $S = (4, 2, 1)$ to $T = (0, 0, 3)$. Calculate the following quantities:

(i) $\vec{u} + \vec{v}$

(ii) $\vec{u} - 3\vec{v}$

(iii) Find the value of the vector $\vec{w}$ that satisfies the equation

$$\vec{u} - 3(4\vec{v} - 2\vec{w}) = \vec{w} + 2\vec{u}.$$

The first step, of course, is determining the components of $\vec{u}$ and $\vec{v}$.

$$\begin{aligned}\vec{u} =& \vec{PQ}\\ =& \langle 0-3, 4-0, -3--1\rangle\\ =& \langle -3, 4, -2\rangle\\ \vec{v} =& \vec{ST}\\ =& \langle 0-4, 0-2, 3-1\rangle\\ =& \langle -4, -2, 2\rangle\end{aligned}$$

Now we just use our algebraic properties to solve each problem.

(i)

$$\begin{aligned}\vec{u} + \vec{v} =& \langle -3, 4, -2\rangle + \langle -4, -2, 2\rangle\\ =& \langle -3 + -4, 4 + -2, -2 + 2\rangle\\ =& \langle -7, 2, 0\rangle\end{aligned}$$

(ii)

$$\begin{aligned}\vec{u} - 3\vec{v} &= \langle -3, 4, -2\rangle - 3\langle -4, -2, 2\rangle\\ &= \langle -3, 4, -2\rangle - \langle -12, -6, 6\rangle\\ &= \langle -3 - -12, 4 - -6, -2 - 6\rangle\\ &= \langle -3 + 12, 4 + 6, -2 - 6\rangle\\ &= \langle 9, 10, -8\rangle\end{aligned}$$

(iii) Since we're solving for $\vec{w}$, our first goal is to get $\vec{w}$ by itself on one side of the equation.

$$\begin{aligned} \vec{u} - 3(4\vec{v} - 2\vec{w}) &= \vec{w} + 2\vec{u} \\ \implies \vec{u} - 12\vec{v} + 6\vec{w} &= \vec{w} + 2\vec{u} \\ \implies 5\vec{w} &= \vec{u} + 12\vec{v} \\ \implies \vec{w} &= \frac{1}{5}\left(\vec{u} + 12\vec{v}\right) \end{aligned}$$

Now we can just plug in the components for $\vec{u}$ and $\vec{v}$ and work through the arithmetic.

$$\begin{aligned} \vec{w} =& \frac{1}{5}\left(\langle 3,4,-2\rangle + 12\langle -4,-2,2\rangle\right) \\ =& \frac{1}{5}\left(\langle 3,4,-2\rangle + \langle -48,-24,24\rangle\right) \\ =& \frac{1}{5}\langle 3-48, 4-24, -2+24\rangle \\ =& \frac{1}{5}\langle -45,-28,22\rangle \\ =& \left\langle -9, \frac{-28}{5}, \frac{22}{5}\right\rangle \end{aligned}$$

Application: velocity and acceleration

We very quickly recall what is probably the most basic application of vectors to physics: velocity and acceleration. Recall that *velocity* denotes a change in distance over time, and *acceleration* is a change in velocity over time. A *velocity vector* tells us how quickly the object is moving and is defined to be the displacement vector divided by the amount of time it took that for that displacement to happen.

Let's say we have some particle, in two dimensions, which at some point in time is at the position $P_0 = (1, 2)$, and then two seconds later is at the position $P_1 = (3, -2)$. The velocity vector of this particle is then

$$\vec{v} = \frac{1}{2}\overrightarrow{P_0P_1} = \frac{1}{2}\langle 2,-4\rangle = \langle 1,-2\rangle .$$

The *speed* of the particle (a scalar quantity) is the magnitude of the displacement, divided by the length of time: equivalently, the speed is the same thing as the magnitude of the velocity vector. So the speed of the particle describe above is

$$\|\vec{v}\| = \| \langle 1, -2 \rangle \| = \sqrt{1+4} = \sqrt{5}.$$

The *acceleration vector* is defined as a change in velocity (so the difference of two vectors) divided by a change in time (how long it took the velocity to change). Let's suppose that our particle, whose initial velocity was

$$\vec{v} = \langle 1, -2 \rangle$$

over the course of three seconds changed to

$$\vec{v_1} = \langle 7, 3 \rangle .$$

The corresponding acceleration vector is then

$$\vec{a} = \frac{1}{3}(\vec{v_1} - \vec{v}) = \frac{1}{3}\langle 6, 5 \rangle = \langle 2, {}^5\!/\!_3 \rangle.$$

Application: tension

Finally we put all of the above ideas together and present one important, physical application of vectors: tension. Before we get started, let's review a few very basic concepts in physics.

Units

The units that are typically used in scientific calculations are called the *SI units* which stands for the *international system*[3]. The SI unit for mass is the *kilogram* (abbreviated kg); the SI unit for distance is the *meter* (m); and the SI unit for time is the second (s). Since velocity is a change in distance over a change in time, the SI units for velocity are *meters per second* ($\frac{\mathrm{m}}{\mathrm{s}}$). Since acceleration is a change in velocity over time, the SI units for acceleration are (*meters per second*) *per second*, which is usually just pronounced *meters per second squared* and denoted $\frac{\mathrm{m}}{\mathrm{s}^2}$.

Newton's laws

Newton's second law of motion can be stated as *force equals mass times acceleration*:

$$\vec{F} = m\vec{a}.$$

[3]The letters *SI* are backwards because the abbreviation is French, and most adjectives in French come after the noun they modify: *système international*.

That is, this push or pull (the force) required to make an object of mass m have acceleration $\vec{a}$ is given by $\vec{F}$.

One quick comment about mass and weight: *mass* is an intrinsic property of matter that tells you how difficult it is to make an object accelerate; *weight* means how much force gravity is exerting on an object. That is, your weight on the Earth and your weight on the moon are different values because the pull of gravity is different on the Earth than it is on the moon. Your mass, however, stays the same. Newton's second law tells us precisely how mass and weight are related: if we know the mass of an object, to find the weight we just multiply the mass by the acceleration due to gravity. On the surface of the Earth, the acceleration due to gravity is about $9.8\frac{\text{m}}{\text{s}^2}$.

Because force is mass times acceleration, the SI unit for force is *kilogram meters per second squared*: $\frac{\text{kg m}}{\text{s}^2}$. This unit is a bit cumbersome to say and write, so a *Newton* (N) is defined to be a kilogram meter per second squared. That is, a force of 10 N (pronounced "ten Newtons") is just short for $10\frac{\text{kg m}}{\text{s}^2}$.

Just to compare these to quantities you're more familiar with, another unit for force (the English unit) is the *pound* (lb). Thus if you say that you weigh 150lb, what you mean is that the force of gravity the Earth exerts on you is 150 of these particular force units. A Newton is equal to about 4.49lb; so saying that you weigh 150 lb is the same as saying gravity is pulling you down with a force of 673.5 N.

In general there may be several difference forces acting on an object all at once. The *net force* is defined as the sum total of all of these forces.

Tension

Let's suppose you have some object of mass $m = 10$kg which is suspended by a rope. Gravity is pulling that object straight down by a force with magnitude 98N.

However, the object isn't moving: in particular, it's not accelerating. Since there's no acceleration, the net force must be zero. Since gravity is pulling the object downwards with a force of 98N, something must be pulling the object upwards with a force of 98N is well, and the forces balance out. This force is supplied by the rope, and is called *tension*. This is shown in Figure 1.30.

Now, most of the information in this picture is irrelevant. The width of the rope and the dimensions of the suspended object don't matter, so we don't need to draw them in our picture. If we replace our object with

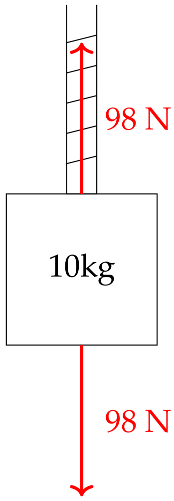


Figure 1.30: An object suspended by a rope. Forces are shown in red.

a single point, and then only draw the force vectors, the picture we get is called a *force diagram* and is shown in Figure 1.31.

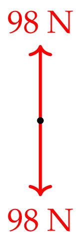
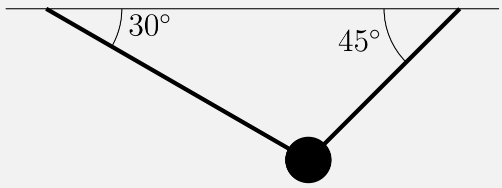


Figure 1.31: The force diagram for the setup from Figure 1.30.

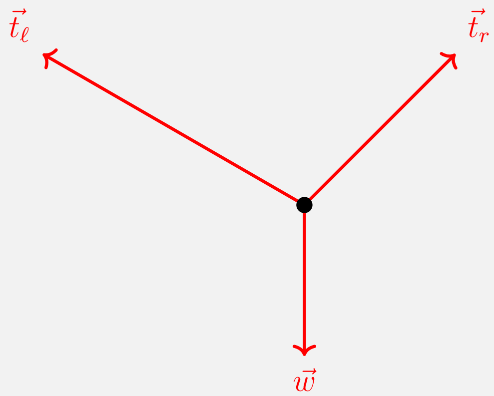
Let's say that $\vec{w}$ is the weight of the object: the force of gravity pulling the object down. Then $\vec{w} = \langle 0, -98 \rangle$. If $\vec{t}$ is the tension in the rope pulling the object up, then $\vec{t} = \langle 0, 98 \rangle$. Of course, $\vec{w} + \vec{t} = 0$.

Example 1.6.
Now let's consider a more complicated example. Suppose there are two ropes holding the object up, and that the ropes are hung from the ceiling at different angles, as indicated below. Suppose the weight of our object is 20 N. What is the magnitude of the tension vectors in each of these ropes? Phrased another way, how strong do the ropes need to hold up the object? (The "strength" of the ropes is the

amount of tension they can supply before breaking, and is called the *tensile strength* of the rope.)



Newton's laws tell us that if the object is going to be suspended, then the net force acting on the object is zero. There are three forces involved here: weight (gravity), the tension in the left-hand rope, and the tension in the right-hand rope. Let's call the corresponding force vectors $\vec{w}$, $\vec{t_\ell}$, and $\vec{t_r}$, respectively. Then our force diagram appears in the image below.



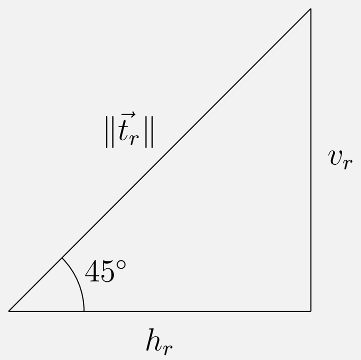
We know that the forces balance out,

$$\vec{t_\ell} + \vec{t_r} + \vec{w} = 0,$$

and also, since the object has weight 20 N, we know $\vec{w} = \langle 0, -20 \rangle$. Now we have to figure out what $\vec{t_\ell}$ and $\vec{t_r}$ look like. As a preliminary step, let's write

$$\begin{aligned} \vec{t_r} &= \langle h_r, v_r \rangle \\ \vec{t_\ell} &= \langle h_\ell, v_\ell \rangle . \end{aligned}$$

So our goal is to figure out what these numbers, h_r, v_r, h_ℓ, v_ℓ, are. Let's first notice that these values are related to $\|\vec{t_r}\|$ and $\|\vec{t_\ell}\|$ by some basic geometry. In particular, if we make a right triangle whose legs have lengths h_r and v_r, then $\|\vec{t}_r\|$ is the hypotenuse of the triangle. See the image below.



Notice that the angle in the lower left-hand corner of this triangle is $45°$. (This is by a basic theorem in geometry. Google *alternate interior angles* it if you don't understand why this angle should be $45°$.) Because of this, we know

$$\begin{aligned} h_r =& \|\vec{t}_r\| \cos(45°) = \frac{\sqrt{2}}{2}\|\vec{t}_r\|, \text{ and} \\ v_r =& \|\vec{t}_r\| \sin(45°) = \frac{\sqrt{2}}{2}\|\vec{t}_r\|. \end{aligned}$$

We could do the same thing to relate h_ℓ and v_ℓ to $\|\vec{t}_\ell\|$ to get

$$\begin{aligned} h_\ell =& -\|\vec{t}_\ell\| \cos(30°) = -\frac{\sqrt{3}}{2}\|\vec{t}_\ell\| \\ v_\ell =& \|\vec{t}_\ell\| \sin(30°) = \frac{1}{2}\|\vec{t}_\ell\|. \end{aligned}$$

(The h_ℓ is negative because our vector is pointing to the left.)

So, plugging these values into our equation $\vec{t}_\ell + \vec{t}_r + \vec{w} = 0$ we have

$$\left\langle -\frac{\sqrt{3}}{2}\|\vec{t}_\ell\|, \frac{1}{2}\|\vec{t}_\ell\| \right\rangle + \left\langle \frac{\sqrt{2}}{2}\|\vec{t}_r\|, \frac{\sqrt{2}}{2}\|\vec{t}_r\| \right\rangle + \langle 0, -20\rangle = \vec{0}$$

Adding these vectors together and equating the components, we get a system of equations.

$$\begin{aligned} -\sqrt{3}\|\vec{t_\ell}\| + \sqrt{2}\|\vec{t_r}\| =& 0 \\ \|\vec{t_\ell}\| + \sqrt{2}\|\vec{t_r}\| =& 40 \end{aligned}$$

Subtracting the top equation from the bottom we have

$$\begin{aligned} (1+\sqrt{3})\|\vec{t_\ell}\| &= 40 \\ \Longrightarrow \|\vec{t_\ell}\| &= \frac{40}{1+\sqrt{3}} \approx 14.641 \end{aligned}$$

Now we can take the second equation and solve for the magnitude of $\vec{t_r}$:

$$\begin{aligned} \|\vec{t_\ell}\| + \sqrt{2}\|\vec{t_r}\| &= 40 \\ \Longrightarrow \sqrt{2}\|\vec{t_r}\| &= 40 - \|\vec{t_\ell}\| = 40 - \frac{40}{1+\sqrt{3}} \\ \Longrightarrow \sqrt{2}\|\vec{t_r}\| &= 40 \cdot \frac{1+\sqrt{3}}{1+\sqrt{3}} - \frac{40}{1+\sqrt{3}} \\ \Longrightarrow \sqrt{2}\|\vec{t_r}\| &= \frac{40\sqrt{3}}{1+\sqrt{3}} \\ \Longrightarrow \|\vec{t_r}\| &= \frac{40\sqrt{3}}{\sqrt{2}+\sqrt{6}} \approx 17.9312 \end{aligned}$$

So, the tension in the left-hand rope is about 16.641 N, and the tension in the right-hand rope is about 17.931 N. (Just to help put this in perspective, if we convert everything to English units we get the following. Our object weights about 4.5 lb; the tension in the left-hand rope is 3.75 lb; and the tension in the right-hand rope is about 4.03 lb.)

You'll notice that in solving this problem, we stopped after finding the magnitude of the force (tension) vectors and didn't bother to write out the actual vectors. (Of course, now that we know the magnitudes, our relations above let us easily figure out the components of the vectors.) This is pretty typical for these sorts of problems: the magnitude is what we really care about. In a practical application, what we want to know is how strong the rope needs to be to hold something up, and the magnitude is all we need to answer that question.

Higher-dimensional vectors

As we had mentioned above, we can interpret a vector in two or three dimensions either geometrically as an arrow, or algebraically a finite ordered list of numbers. While arrows are nice to visualize when we want a picture, the algebraic point of view is usually much easier from the point of view of doing calcultions. It is also very easy to generalize this algebraic interpretation to higher dimensions. That is, we can define a four-dimensional vector as simply an ordered list of four numbers, such as

$$\vec{v} = \begin{pmatrix} 2 \\ 7 \\ -3 \\ 12 \end{pmatrix}.$$

Of course, there are several ways we could write down this information: we could write it as a column vector, or as a row vector, or a list of numbers in the angle brackets, $\langle$ and $\rangle$, or many other ways. All of these different ways of writing the list of numbers carry the same information, so we won't always make a big deal to distinguish the different ways of writing down a vector. The main thing is that we can easily write down vectors in four dimensions, five dimensions, six dimensions, and so on, just by generalizing what we can do in two or three dimensions, even if we can't easily visualize what a "five-dimensional arrow" might look like.

When considering n-dimensional vectors, we'll usually denote the components of the vector as v_1, v_2, v_3, ..., v_{n-1}, and v_n, so our vector may be written as

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}.$$

Our two operations of scalar multiplication and vector addition are just as easy to perform for n-dimensional vectors, regardless of how big

n is, as they are for two- and three-dimensional vectors:

$$\vec{v}+\vec{u}=\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}+\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix}=\begin{pmatrix} v_1+u_1 \\ v_2+u_2 \\ v_3+u_3 \\ \vdots \\ v_{n-1}+u_{n-1} \\ v_n+u_n \end{pmatrix}$$

$$\lambda\vec{v}=\lambda\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}=\begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \lambda v_3 \\ \vdots \\ \lambda v_{n-1} \\ \lambda v_n \end{pmatrix}$$

We will let $\mathbb{R}^n$ denote the set of all n-dimensional vectors.

Remark.
Notice that we think of $\mathbb{R}^n$ in two different ways: the set of all n-dimensional points, or the set of all n-dimensional vectors. Both interpretations carry the same amount of information: a point in n-dimensional space and an n-dimensional vector are both determined by an ordered list of n real numbers. It may seem odd that sometimes we'll interpret that information as a point and sometimes as a vector, but in practice it's usually not a problem to switch back and forth between these interpretations, even if it seems a little strange when you first learn the material.

1.4 Linear transformations and matrices

In the previous section we saw that elements of n-dimensional space, $\mathbb{R}^n$, can be thought of as vectors, and that there are two primary operations that we can perform on vectors: scalar multiplication and vector addition. It's often in the case in mathematics that once you have operations defined on a set, studying functions which preserve those operations can reveal important information. We now quickly discuss the functions that preserve vector addition and scalar multipication, which are called *linear transformations*.

We will see also see that linear transformations are very closely related to rectangular arrays of numbers called *matrices*; in fact, they are really the same thing in disguise. After explaining this and seeing a few examples, we will introduce an operation called the *determinant* which associates a number to a matrix/linear transformation that carries important geometric information.

Warning.
The material in this section is more abstract than what you may have seen in mathematics courses up to this point, and it may seem odd and confusing at first. Don't worry too much if you don't feel completely comfortable with everything discussed here, especially the first time you read this section. As we look at more material and see more examples throughout the course the material here will gradually become clearer and more comfortable.

Although this material may seem a bit odd and out of place right now, taking the time to discuss linear transformations will make certain operations later in the course simpler to describe, so we're trading some confusion now for clarity later.

It is possible to described multivariable calculus without explicitly discussing linear transformations, matrices, and determinants (and this is the approach many books take), however they are always hiding in the background. Certain formulas that are essential to later material become extremely "clunky" if they are not described in terms of matrices and determinants, and so we're investing some time discussing this material now to make other material later more straight-foward.

Linear transformations

To be a little bit more precise, a ***linear transformation*** is a function T that takes n-dimensional vectors and converts them into m-dimensional vectors (m and n may be the same, but they don't need to be), and such that for all pairs of n-dimensional vectors $\vec{u}$ and $\vec{v}$ and all scalars λ we have the following two identities:

$$\begin{aligned} T(\vec{u}+\vec{v}) &= T(\vec{u})+T(\vec{v}) \\ T(\lambda\vec{v}) &= \lambda T(\vec{v}) \end{aligned}$$

That is, a linear transformation allows us to split up sums and factor out scalars.

For example, consider the following function that converts three-dimensional vectors into two-dimensional vectors:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x+y \\ z-5y \end{pmatrix}. \tag{1.1}$$

Here we are specifying T by describing how it uses the components of its input vector (which is three-dimensional) to build its output vector (which is two-dimensional). When this function is applied to the vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, for example, the result is

$$T\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2\cdot 1+2 \\ 3-5\cdot 2 \end{pmatrix} = \begin{pmatrix} 4 \\ -7 \end{pmatrix}. \tag{1.2}$$

Sometimes in order to say that T takes n-dimensional vectors and converts them into m dimensional vectors we will write $T:\mathbb{R}^n \to \mathbb{R}^m$.

It is important to notice that for a function T to be linear it *must* satisfy $T(\vec{u}+\vec{v}) = T(\vec{u})+T(\vec{v})$ and $T(\lambda\vec{v}) = \lambda T(\vec{v})$ *for all* input vectors $\vec{u}$ and $\vec{v}$, and all scalars λ. In order to check that this is the case, we have to leave the components of $\vec{u}$ and $\vec{v}$ as variables and check if the two sides of the equalities mentioned above really are equal.

In the case of the map T described above, we must compute the following:

$$\begin{aligned}
T(\vec{u}+\vec{v}) &= T\begin{pmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{pmatrix} \\
&= \begin{pmatrix} 2(u_1+v_1)+u_2+v_2 \\ u_3+v_3-5(u_2+v_2) \end{pmatrix} \\
&= \begin{pmatrix} 2u_1+u_2+2v_1+v_2 \\ u_3-5u_2+v_3-5v_2 \end{pmatrix} \\
&= \begin{pmatrix} 2u_1+u_2 \\ u_3-5u_2 \end{pmatrix} + \begin{pmatrix} 2v_1+v_2 \\ v_3-5v_2 \end{pmatrix} \\
&= T\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + T\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\
&= T(\vec{u}) + T(\vec{v})
\end{aligned}$$

Since we left the components of our vectors as arbitrary values (the variables u_1, u_2, and so on), the calculation above shows that *for all* three-dimensional vectors $\vec{u}$ and $\vec{v}$, we have that $T(\vec{u}+\vec{v})$ equals $T(\vec{u})+T(\vec{v})$. If we had instead plugged in two particular choices for $\vec{u}$ and $\vec{v}$, our calcuation would have only shown that we can break up the sum of those particular vectors, but we need to show we can break up *all* sums of vectors, hence the need to leave the components as variables.

We can similarly check that scalars can be factored out of the function above:

$$
\begin{aligned}
T(\lambda\vec{v}) &= T\left(\lambda\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\right)\\
&= T\begin{pmatrix}\lambda v_1\\\lambda v_2\\\lambda v_3\end{pmatrix}\\
&= \begin{pmatrix}2\lambda v_1+\lambda v_2\\\lambda v_3-5\lambda v_2\end{pmatrix}\\
&= \begin{pmatrix}\lambda(2v_1+v_2)\\\lambda(v_3-5v_2)\end{pmatrix}\\
&= \lambda\begin{pmatrix}2v_1+v_2\\v_3-5v_2\end{pmatrix}\\
&= \lambda T\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\\
&= \lambda T(\vec{v}).
\end{aligned}
$$

Again, because the components of $\vec{v}$ and the scalar λ were left as variables, the above calculation shows that $T(\lambda\vec{v}) = \lambda T(\vec{v})$ *for all* choices of $\vec{v}$ and scalar λ.

Notice that not every function converting three-dimensional vectors into two-dimensional vectors is necessarily a linear transformation. For example, if we replace the function T above with the following,

$$
T\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x^2+y\\1+z\end{pmatrix},
$$

then we will not have a linear transformation. To see this, we just need to find a single example where either of the inequalities $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ or $T(\lambda\vec{v}) = \lambda T(\vec{v})$ fails. Consider, for example, the following:

$$
\begin{aligned}
T\left(4\begin{pmatrix}1\\2\\3\end{pmatrix}\right) &= T\begin{pmatrix}4\\8\\12\end{pmatrix}\\
&= \begin{pmatrix}4^2+8\\1+12\end{pmatrix}\\
&= \begin{pmatrix}24\\13\end{pmatrix}
\end{aligned}
$$

However, we also have

$$\begin{aligned} 4T\begin{pmatrix}1\\2\\3\end{pmatrix} &= 4\begin{pmatrix}1^2+2\\1+3\end{pmatrix}\\ &= 4\begin{pmatrix}3\\4\end{pmatrix}\\ &= \begin{pmatrix}12\\16\end{pmatrix} \end{aligned}$$

Since this does not equal our earlier calculation, we have

$$T\left(4\begin{pmatrix}1\\2\\3\end{pmatrix}\right) \neq 4T\begin{pmatrix}1\\2\\3\end{pmatrix}.$$

Since we have one instance where $T(\lambda\vec{v}) \neq \lambda T(\vec{v})$, the function T above is not linear.

Remark.
Notice that to be a linear transformation, we must have $T(\vec{u}+\vec{v}) = T(\vec{u})+T(\vec{v})$ and $T(\lambda\vec{v}) = \lambda T(\vec{v})$ *for all* choices of $\vec{u}$, $\vec{v}$, and λ. Once you find one instance where either of these inequalities aren't satisfied, the transformation can not be linear.

One important property of linear transformations is that their composition is a linear transformation. That is, suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation which converts n-dimensional vectors into m-dimensional vectors, and also suppose $S : \mathbb{R}^m \to \mathbb{R}^p$ is a linear transformation converting m-dimensional vectors into p-dimensional vectors. We can then take an n-dimensional vector $\vec{v}$, apply T to obtain an m-dimensional vector $T(\vec{v})$, and then apply S to obtain a p-dimensional vector $S(T(\vec{v}))$. This operation of applying T and then applying S is called ***composition*** and is denoted $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$.

For example, let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation from above

$$T\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}2x+y\\z-5y\end{pmatrix},$$

and suppose $S : \mathbb{R}^2 \to \mathbb{R}^4$ is the function

$$S\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ 2x \\ x-3y \end{pmatrix}. \tag{1.3}$$

Exercise 1.1.
Verify that the function S described in Equation 1.3 above is in fact a linear transformation.

The composition $S \circ T : \mathbb{R}^3 \to \mathbb{R}^4$ is then given by

$$\begin{aligned} S \circ T \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= S\left(T\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) \\ &= S\begin{pmatrix} 2x+y \\ z-5y \end{pmatrix} \\ &= \begin{pmatrix} 2x+y+z-5y \\ 2x+y-(z-5y) \\ 2(2x+y) \\ 2x+y-3(z-5y) \end{pmatrix} \\ &= \begin{pmatrix} 2x-4y+z \\ 2x+6y-z \\ 4x+2y \\ 2x+6y-3z \end{pmatrix} \end{aligned}$$

We would like to see if $S \circ T$ is a linear transformation. We can do this by manually checking that our equalities hold in each example, but this is tedious. Thus we like the following theorem which tells us that provided we already know S and T are both linear transformations, their composition must be a linear transformation as well.

Theorem 1.5.
If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ are both linear transformations, then

their composition $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$ is also a linear transformation.

Notice that we need to already now S and T are linear transformations for Theorem 1.5 to be helpful, but once we've shown this for S and T, we get that $S \circ T$ is linear for free.

Matrices

A ***matrix*** is simply a rectangular array of numbers, such as

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

or

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \\ 1 & -3 \end{pmatrix}$$

When referring to a matrix we specify its number of rows and columns by saying the matrix is $m \times n$ (pronounced "m by n") if it has m rows and n columns. The matrices above, for example, are 2×3 and 4×2.

It is often convenient to give matrices a name to save ourselves some writing. For example, if we write

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix}$$

then we will refer to this particular 2×3 matrix by A. Once we given a matrix a name, we will sometimes refer to the the individual entries in the matrix by giving the lowercase name of the matrix with subscripts indicating the row and column. For instance, we write a_{13} to mean the entry of matrix A in the first row and third column. For our matrix A above this would be $a_{13} = 0$; and a_{22} would be -5.

In general, when referring to an arbitrary $m \times n$ matrix we will leave the individual entries as variables a_{ij}; these act as a placeholder for the entry in the i-th row and j-th column,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

Matrices are used in many different areas of mathematics and have lots of applications; you could easily spend an entire course discussing matrices and their applications (and if you take Linear Algebra, that's basically what you'll be doing). There are lots of things we could say about matrices, and several different operations, but there are three main things that will be important for us in this course. Two operations we'll describe now, and one we'll discuss later.

The first operation we will consider will allow us to "multiply" a matrix by a vector, the result of which is a new vector. In particular, we will multiply an $m \times n$ matrix by an n-dimensional vector, and the result will be an m-dimensional vector. We are explicitly defining this operation only when the number of columns in the matrix equals the number of entries in the vector. For example, we can multiply a 2×3 matrix by a 3-dimensional vector, but we *can not* multiply a 2×3 matrix by a 4-dimensional vector.

This operation is rather tedious to describe in words, but actually very easy to compute in practice. We will write down the wordy description first, but just bear with it for a moment, and then we will do an example and the example should be easy to follow.

Let's suppose that A is an $m \times n$ matrix and $\vec{v}$ is an n-dimensional vector,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \qquad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

We will treat each column of the matrix A as an m-dimensional vector,

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \begin{pmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

We will then multiply the first column (thought of as a vector) by the first entry of the vector, v_1, multiply the second column by the second entry in the vector, v_2, multiply the third column by the third entry, v_3, and so on, then add these vectors together. The result is the product of A and $\vec{v}$, denoted $A\vec{v}$:

$$A\vec{v} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

$$= v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + v_3 \begin{pmatrix} a_{13} \\ a_{23} \\ \vdots \\ a_{m3} \end{pmatrix} + \cdots + v_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Notice that since we are adding together m-dimensional vectors (each column has m entries since the original matrix A has m rows), the result is an m-dimensional vector.

As a concrete example, let's multiply the matrix

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix}$$

by the vector

$$\vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

We will multiply the first column of the matrix by 1, the second column by 2, and the third column by 3, then add the result together:

$$\begin{aligned} A\vec{v} &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= 1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ -5 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ -10 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ -7 \end{pmatrix} \end{aligned}$$

One of the main reasons we care about matrices is that this operation of multiplying a matrix and a vector defines a linear transformation! That is, we claim that matrix multiplication distributes over vector addition,

$$A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v},$$

and commutes with scalar multiplication,

$$A(\lambda\vec{v}) = \lambda(A\vec{v}).$$

Assuming these two properties are true (which is proven in Appendix A.1), this means we can use an $m \times n$ matrix A to define a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ by defining $T(\vec{v})$ to be the product the product $A\vec{v}$. The two properties mentioned above, and restated in Theorem 1.6 below, are exactly what we need to know that $T(\vec{v}) = A\vec{v}$ defines a linear transformation.

Theorem 1.6.
If A is an $m \times n$ matrix, $\vec{u}$ and $\vec{v}$ are n-dimensional vectors and λ is a scalar, then

$$A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}, \text{ and}$$
$$A(\lambda\vec{v}) = \lambda(A\vec{v}).$$

As a consequence, the function $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\vec{v}) = A\vec{v}$ is a linear transformation.

Thus every matrix determines a linear transformation. Moreover, every linear transformation can be written as matrix multiplication. You may have noticed, for example, that the product we computed above

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ -7 \end{pmatrix}$$

was the same as the result of the linear transformation described in Equation 1.2 on page 48. In fact, if we leave the components of the vector as variables, we compute

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\begin{pmatrix} 2 \\ 0 \end{pmatrix} + y\begin{pmatrix} 1 \\ -5 \end{pmatrix} + z\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2x+y \\ -5y+z \end{pmatrix} = \begin{pmatrix} 2x+y \\ z-5y \end{pmatrix}$$

gives us the exact same result as the linear transformation described in Equation 1.1. This is not a coincidence: all linear transformations are really just multiplication of a matrix and a vector. We can even explicitly compute what the matrix should be for any given linear transformation, as described in Theorem 1.7.

Theorem 1.7.
If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then T has the form

$$T(\vec{v}) = A\vec{v}$$

where A is the $m \times n$ matrix whose first column contains the entries in the vector

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

whose second column contains the entries in the vector

$$T\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

whose third column contains the entries in the vector

$$T\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and so on.

In the case of the example described by Equation 1.1, notice that we compute

$$T\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}2\cdot 1+0\\0-5\cdot 0\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix}$$
$$T\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}2\cdot 0+1\\0-5\cdot 1\end{pmatrix} = \begin{pmatrix}1\\-5\end{pmatrix}$$
$$T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}2\cdot 0+0\\1-5\cdot 0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$$

and these are the columns of the matrix

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

previously described.

Example 1.7.
Determine the matrix associated to the linear transformation S that was described in Equation 1.3 on page 52.

Recall $S : \mathbb{R}^2 \to \mathbb{R}^4$ was defined by

$$S\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+y\\x-y\\2x\\x-3y\end{pmatrix}.$$

Since this transformation takes two-dimensional vectors and converts them into four-dimensional vectors, we expect its representative matrix to be 4×2. To find our two columns, we need to apply S to the vectors

$$\begin{pmatrix}1\\0\end{pmatrix} \text{ and } \begin{pmatrix}0\\1\end{pmatrix}.$$

We simply compute

$$S\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\2\\1\end{pmatrix}$$
$$S\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\-1\\0\\-3\end{pmatrix}.$$

Putting these together, we see that our matrix is

$$\begin{pmatrix}1 & 1\\1 & -1\\2 & 0\\1 & -3\end{pmatrix}.$$

Just to double-check this really is the correct matrix, we can multiply this 4×2 matrix by some arbitrary two-dimensional vector and verify that the result is the same as the transformation S applied to that vector,

$$\begin{pmatrix}1 & 1\\1 & -1\\2 & 0\\1 & -3\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = x\begin{pmatrix}1\\1\\2\\1\end{pmatrix} + y\begin{pmatrix}1\\-1\\0\\-3\end{pmatrix} = \begin{pmatrix}x+y\\x-y\\2x\\x-3y\end{pmatrix}.$$

Notice this is exactly the value of $S\begin{pmatrix}x\\y\end{pmatrix}$.

Exercise 1.2.
Consider the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}3x - y + 2z\\-x + z\\6y - 3z\end{pmatrix}.$$

Compute the 3×3 matrix that represents this transformation.

Multiplying matrices

Suppose now that we have two linear transformations which we can compose; say $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$. We had previously seen in Theorem 1.5 that their composition $S \circ T$ is a linear transformation which takes n-dimensional vectors, applies T to convert them into m-dimensional vectors, and then applies S to finally convert them into p-dimensional vectors. Since $S \circ T$ is linear, it should be represented by a matrix. Is there an "easy" way for us to compute this matrix if we already have the matrices for S and T?

We have seen that we can multiply a matrix and a vector, and this was tantamount to applying a linear transformation. To describe the composition of two linear transformations we will extend our notion of multiplication to allow us to multiply two matrices together. Our goal here is to define matrix multiplication in such a way that multiplying two matrices is the same as composing linear transformations. We can only compose linear transformations, though, when the dimensions "match up" appropriately. That is, the composition $S \circ T$ is only defined if the output of T has the same dimension as the input of S. In terms of our matrices, the dimension of the output of a linear transformation is the number of columns of the matrix, and the dimension of the number is the number of rows. Thus we will only define matrix multiplication when the number of rows of the right-hand matrix equals the number of columns of the left-hand matrix.

Before describing the general procedure, let's consider a concrete example. We had previously seen linear transformations $T : \mathbb{R}^3 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^4$ given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ -5y + z \end{pmatrix} \qquad S\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \\ 2x \\ x - 3y \end{pmatrix}.$$

The composition of these matrices, $S \circ T$, will take three-dimensional vectors and ultimately convert them into four-dimensional vectors. Thus the matrix representing $S \circ T$ will be 4×3. Our goal will be to determine this 4×3 matrix just from the matrices that represent T and S.

Let's denote the 2×3 matrix representing T by A, and let B denote the 4×2 matrix representing S. We had calculated above that

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \\ 1 & -3 \end{pmatrix}.$$

Now let's imagine that we apply $S \circ T$ to some arbitrary three-dimensional vector $\vec{v}$. Since T is applied first and S is applied second, the matrix A (representing T) should be multiplied with $\vec{v}$ first, and then afterwards we should multiply the result by the matrix B (representing S). That is, we want to compute

$$BA\vec{v} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & -5 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Multiplying A and $\vec{v}$ together first, leaving B alone, this becomes

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 2x + y \\ -5y + z \end{pmatrix}$$

We can multiply this matrix and vector together to obtain

$$\begin{pmatrix} 2x + y + (-5y + z) \\ 2x + y - (-5y + z) \\ 2(2x + y) \\ 2x + y - 3(-5y + z) \end{pmatrix} = \begin{pmatrix} 2x - 4y + z \\ 2x + 6y - z \\ 4x + 2y \\ 2x + 6y - 3z \end{pmatrix}$$

Notice this is the same as the product

$$\begin{pmatrix} 2 & -4 & 1 \\ 2 & 6 & -1 \\ 4 & 2 & 0 \\ 2 & 6 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

That is, if our definition of matrix multiplication is going to agree with our composition of linear transformations, we will need to define matrix multiplication in such a way that

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

To see exactly what's going on, let's replace the numbers in our example above with variables:

$$A = \begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{pmatrix} \qquad B = \begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\\ b_{41} & b_{42}\end{pmatrix}.$$

Now let's again consider the product $BA\vec{v}$,

$$\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\\ b_{41} & b_{42}\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}.$$

We first multiply $A\vec{v}$, leaving B alone for the moment, to obtain

$$\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\\ b_{31} & b_{32}\\ b_{41} & b_{42}\end{pmatrix}\begin{pmatrix}x\,a_{11} + y\,a_{12} + z\,a_{13}\\ x\,a_{21} + y\,a_{22} + z\,a_{23}\end{pmatrix}$$

Now we multiply the B matrix with this vector which gives us

$$\begin{pmatrix}(x\,a_{11} + y\,a_{12} + z\,a_{13})\,b_{11} + (x\,a_{21} + y\,a_{22} + z\,a_{23})\,b_{12}\\ (x\,a_{11} + y\,a_{12} + z\,a_{13})\,b_{21} + (x\,a_{21} + y\,a_{22} + z\,a_{23})\,b_{22}\\ (x\,a_{11} + y\,a_{12} + z\,a_{13})\,b_{31} + (x\,a_{21} + y\,a_{22} + z\,a_{23})\,b_{32}\\ (x\,a_{11} + y\,a_{12} + z\,a_{13})\,b_{41} + (x\,a_{21} + y\,a_{22} + z\,a_{23})\,b_{42}\end{pmatrix}$$

We can distribute and rewrite this as

$$\begin{pmatrix}x\,(b_{11}a_{11} + b_{12}a_{21}) + y\,(b_{11}a_{12} + b_{12}a_{22}) + z\,(b_{11}a_{13} + b_{12}a_{23})\\ x\,(b_{21}a_{11} + b_{22}a_{21}) + y\,(b_{21}a_{12} + b_{22}a_{22}) + z\,(b_{21}a_{13} + b_{22}a_{23})\\ x\,(b_{31}a_{11} + b_{32}a_{21}) + y\,(b_{31}a_{12} + b_{32}a_{22}) + z\,(b_{31}a_{13} + b_{32}a_{23})\\ x\,(b_{41}a_{11} + b_{42}a_{21}) + y\,(b_{41}a_{12} + b_{42}a_{22}) + z\,(b_{41}a_{13} + b_{42}a_{23})\end{pmatrix}.$$

This is the same as the following multiplication of a matrix and a vector:

$$\begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} + b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} + b_{21}a_{13} + b_{22}a_{23} \\ b_{31}a_{11} + b_{32}a_{21} & b_{31}a_{12} + b_{32}a_{22} + b_{31}a_{13} + b_{32}a_{23} \\ b_{41}a_{11} + b_{42}a_{21} & b_{41}a_{12} + b_{42}a_{22} + b_{41}a_{13} + b_{42}a_{23} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

That is, the product of our two matrices should be defined to be

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} & b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} & b_{21}a_{13} + b_{22}a_{23} \\ b_{31}a_{11} + b_{32}a_{21} & b_{31}a_{12} + b_{32}a_{22} & b_{31}a_{13} + b_{32}a_{23} \\ b_{41}a_{11} + b_{42}a_{21} & b_{41}a_{12} + b_{42}a_{22} & b_{41}a_{13} + b_{42}a_{23} \end{pmatrix}$$

This is quite an ugly expression, but it is the "right" expression, the right way to define matrix multiplication, if we want matrix multiplication to be the same thing as composition of linear transformations.

Even though this expression is rather ugly at first glance, if you looked at it for a moment you might realize there are some nice patterns. In particular, you might observe that the first column of our product matrix, is obtained by multiplying the B matrix with the first column of the A matrix, thought of as a vector:

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} \\ b_{21}a_{11} + b_{22}a_{21} \\ b_{31}a_{11} + b_{32}a_{21} \\ b_{41}a_{11} + b_{42}a_{21} \end{pmatrix}$$

Similarly, the second column of the product matrix is the product of the entire B matrix with the second column of the A matrix; the third column of the product is the same as the product of the entire B matrix with the third column of the A matrix. This, of course, is not simply a coincidence: what's really happening here is that Theorem 1.7 is being applied.

We know from Theorem 1.7 that the matrix representing $S \circ T$ is given by applying $S \circ T$ to the vector with a 1 as its first component and all zeros otherwise, this gives the first column of the matrix. The second column is obtained by applying $S \circ T$ to the vector with a 1 in its second component and all zeros otherwise, and so on. When we actually do this, however, we apply T first, and we know tha T applied to the vector with first coordinate 1 and all zeros otherwise gives us the first column of of the matrix representing T. When we then apply S, we are thus multiplying the matrix representing S by the first column of the matrix representing T, and this product gives us the first column of the matrix representing $S \circ T$. It's all a bit tedious to write out in detail, but that's all that's happening with matrix multiplication.

Example 1.8.
Compute the following product of a 3×2 matrix A with a 2×4 matrix B,

$$A = \begin{pmatrix} 1 & 1 \\ 3 & -5 \\ -1 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & 3 & 3 & 4 \\ 1 & -4 & 0 & 2 \end{pmatrix}$$

We compute the product AB one column at a time by multiplying A by the first column of B, then multiplying A by the second column of B, and so on. This gives us the following:

$$\begin{pmatrix} 1 & 1 \\ 3 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -11 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 3 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ -4 \end{pmatrix} = \begin{pmatrix} -1 \\ 29 \\ -15 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 3 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ -3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 3 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix}$$

We then put these four vectors together as the columns of the product matrix AB,

$$AB = \begin{pmatrix} -1 & -1 & 3 & 6 \\ -11 & 29 & 9 & 2 \\ 5 & -15 & -3 & 2 \end{pmatrix}$$

Determinants

Given an $n \times n$ matrix A, we will associate to A a number called the ***determinant*** of A and denoted $\det(A)$. The general formula for computing determinants is a bit hairy, so let's momentarily only consider the case of a 2×2 matrix where the formula is a little bit simpler.

If A is a 2×2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then we define the determinant of A to be

$$\det(A) = ad - bc.$$

One way to interpret the what this number $\det(A)$ means is that it tells us how the linear transformation associated to the matrix changes the size of subsets.

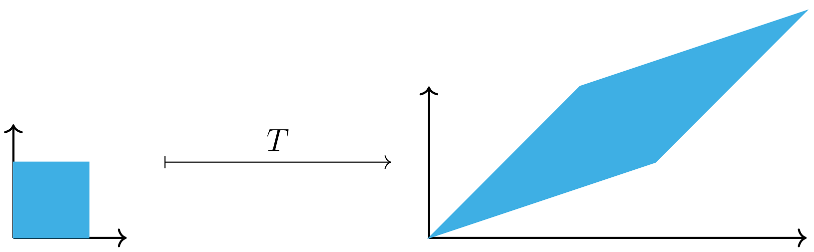
To be concrete, consider the matrix

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

Recall that this matrix determines a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which transforms 2-dimensional vectors:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x + 2y \\ x + 2y \end{pmatrix}$$

Now suppose that we applied this transformation to every point in the unit square of the plane. (Here we are taking points (x, y) in the square, interpreting them as vectors from the origin to the point which is $\begin{pmatrix} x \\ y \end{pmatrix}$, and then applying the transformation above.) This transformation takes the unit square and converts it into some parallelogram of area 4.



Notice the determinant of our matrix is

$$\det(A) = 3 \cdot 2 - 2 \cdot 1 = 6 - 2 = 4$$

and our transformation took a region of area 1 and transformed it into a region of area 4.

In general, if we have a set $S \subseteq \mathbb{R}^2$ and we apply a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ to each point of S to get a new set $T(S)$. If A is the matrix representing the transformation, then the absolute value of $\det(A)$ is the ratio of the areas of S and $T(S)$:

$$\text{Area}(T(S)) = |\det(A)| \cdot \text{area}(S). \tag{1.4}$$

All we're saying here is that determinants of a 2×2 matrix measure how area changes when we apply a linear transformation.

Exercise 1.3.
Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - 3y \\ 4x + y \end{pmatrix}$$

How does applying this linear transformation to a region S in the plane affect that region's area?

Something similar happens in three dimensions: the determinant of a 3×3 matrix tells us how volume changes when we apply a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$. The formula for computing the determinant of a 3×3 matrix is a bit more involved, though. Given a 3×3 matrix A,

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

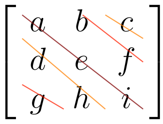
the determinant $\det(A)$ is computed by

$$\det(A) = aei + bfg + cdh - ceg - bdi - afh.$$

This looks like a complicated formula, but luckily there's a nice way to remember it. First, let's rearrange our formula just a tiny little bit.

$$\begin{aligned} &aei - afh - bei + bfg + ceh - ceg \\ =&aei + bfg + cdh - ceg - bdi - afh \end{aligned}$$

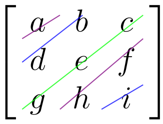
Now the way we remember this formula is that we look at lines through the matrix which go down and to the right (wrapping around if you hit the "edge" of the matrix):



We multiply the entries on each line, and then add them all up:

$$aei + bfg + bdh$$

Now to get the other entries, we draw lines through the matrix which go down and left:



Now we multiply the entries on each line, make them negative, and add them all up:

$$-ceg - bdi - afh$$

Now adding up these two quantities (the positives and the negatives), we have our formula for the determinant.

$$\det(M) = aei + bfg + bdh - ceg - bdi - afh$$

As one very easy example, consider the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ determined by the following matrix,

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

The linear transformation determined by this matrix has the following simple form:

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ y \\ 3z \end{pmatrix}$$

Notice that this simply stretches the x-coordinate by 2, leaves the y-coordinate alone, and stretches the z-coordinate by 3. If we applied this transformation to the unit $1 \times 1 \times 1$ cube (which has volume 1) in 3-space, the result would be a $2 \times 1 \times 3$ prism. Thus we take a cube of volume 1 to a prism of volume 6. Notice the determinant of our matrix is

$$\begin{aligned}\det(A) &= 2 \cdot 1 \cdot 3 + 0 \cdot 0 \cdot 0 + 0 \cdot 0 \cdot 0 - 0 \cdot 1 \cdot 0 - 0 \cdot 0 \cdot 3 - 2 \cdot 0 \cdot 0 \\ &= 6\end{aligned}$$

In general, if $T : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation by the 3×3 matrix A, for any subset S of 3-space,

$$\text{Volume}(T(S)) = |\det(A)| \cdot \text{volume}(S). \tag{1.5}$$

You may notice that we have absolute values around the determinant in Equation 1.4 and Equation 1.5. We need to include absolute values because areas and volumes should never be negative, but the determinant *can* be negative. For instance, we can easily calculate that the determinant of the following 2×2 matrix,

$$\begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$$

is -1.

The fact that determinants can be negative actually carries geometric meaning. To see this, the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by the matrix above,

$$\begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$$

This transformation takes a region in the plane, "shears" it to the right, but then flips the region over, as indicated in Figure 1.32. In Figure 1.32 notice the shading in the original square goes from white on the left to blue on the right. When we apply the transformation we will preserve color; white points will be sent to white points, and blue points sent to blue points. Notice, though, instead of going white-to-blue as we go left-to-right, we have to go right-to-left. That is, point P was to the left of point Q in the original square, then in the parallelogram we have $T(Q)$ is to the left of $T(P)$.

The sign of the determinant tells us whether the linear transformation "flips" regions over or not. If the determinant is positive, then there is no "flip" and we say the linear transformation is ***orientation preserving***. If the determinant is negative, then a "flip" occurs and we say the linear transformatioin is ***orientation reversing***.

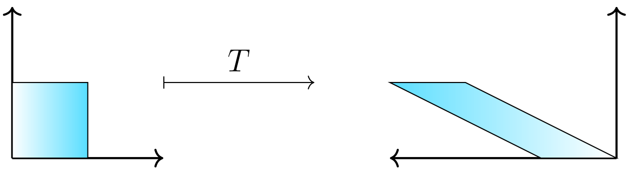


Figure 1.32: A linear transformation may "flip" a region.

Remark.
If the discussion above about determinants, and in particular the issue about orientation preserving and orientation reversing, seems confusing, that's okay! This material takes a little while to come to terms with and will gradually make more sense over time.

We have only described determinants for 2×2 and 3×3 matrices since those are the only ones we will need for calculations in this class. It is possible to compute determinants for any *square* $n \times n$ matrix, but the formulas are a bit more involved. In higher dimensions we actually define the notion of "size" (e.g., the n-dimensional version of area or volume) by using determinants, so determinants are essential for geometry in higher dimensions. In this class we'll primarily be interested in the 2-dimensional and 3-dimensional settings, but this is mostly for convenience and not mathematical necessity. Still, it's good to know that the mathematics we will develop in this course extends to even higher dimensions.

1.5 The dot product

In previous sections we saw how to add and subtract vectors, and we also saw how to multiply a scalar and a vector. However, we never mentioned anything about "multiplying" two vectors together. In this section we'll discuss one way to multiply two vectors, called the *dot product* of the vectors, which results in a scalar. This scalar carries a lot of interesting geometric information about the vectors, and in particular can tell us a little bit about the magnitude of a vector, and the angle between two vectors.

We will begin by first defining the basic algebraic procedure for calculating dot products, list some of the nice properties the dot product has, and finally show how to use the dot product to read off the geometric information described above.

Definition and Basic Properties

Let $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$ be two two-dimensional vectors. We define the *dot product* of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \cdot \vec{v}$ as

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

Example 1.9.
We do a couple of simple examples.

(i) $\langle 2, 3 \rangle \cdot \langle 4, 7 \rangle = 2 \cdot 4 + 3 \cdot 7 = 29$

(ii) $\langle 2, -4 \rangle \cdot \langle 3, 3 \rangle = 2 \cdot 3 + (-4) \cdot 3 = -6$

(iii) $\langle 1, -2 \rangle \cdot \langle 3, 3 \rangle = 1 \cdot 3 + (-2) \cdot 3 = -3$

(iv) $\langle 1, 2 \rangle \cdot \langle 1, 2 \rangle = 1 \cdot 1 + 2 \cdot 2 = 5$

If $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ are two three-dimensional vectors, then their dot product is defined as

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Notice that if we think about our vectors as 3×1 matrices,

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

then the dot product can be expressed in terms of matrix multiplication:

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = [u_1, \quad u_2, \quad u_3] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Theorem 1.8.
Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be three vectors of the same dimension and let $\lambda \in \mathbb{R}$ be a scalar. Then the following properties are satisfied.

(i) $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$

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

(iii) $\vec{0} \cdot \vec{v} = 0$

(iv) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$

(v) $(\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (\lambda \vec{v})$

The angle between two vectors

One of the most important properties of dot products is that they let us calculate the angle between two vectors. Say $\vec{u}$ and $\vec{v}$ are two two-dimensional vectors. By the angle between $\vec{u}$ and $\vec{v}$ we mean the shortest angle when the two vectors are placed tail-to-tail, as indicated in Figure 1.33.

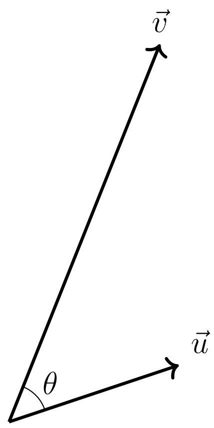


Figure 1.33: The angle between $\vec{u}$ and $\vec{v}$ is θ.

To calculate this angle θ, let's use our vectors $\vec{u}$ and $\vec{v}$ to create a triangle. If this was a right triangle we could use a rule like $\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}}$, but in general this triangle we get *won't* be a right triangle. See Figure 1.34, where a right triangle is given by the dashed line in red, but the triangle we have is a little bit "shorter."

Instead we'll use another theorem from geometry: the law of cosines.

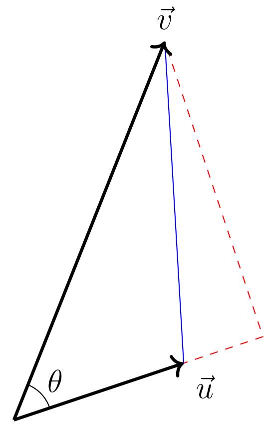
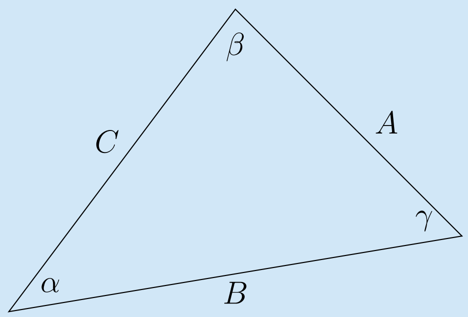


Figure 1.34: The triangle with our vectors.

Theorem 1.9 (The law of cosines)**.**
Suppose that we have a triangle whose sides have lengths A, B, and C, and the angles "opposite" these sides are α, β, and γ, as in the figure below. Let D denote the length of this line segment we have drawn, and let B_1 and B_2 denote the lengths of the two segments into which the side of length B has been cut. Notice $B_1 + B_2 = B$.



Then the lengths of the sides and the angles of the triangle are related by the following equations:

$$\begin{aligned} A^2 =& B^2 + C^2 - 2BC\cos\alpha, \\ B^2 =& A^2 + C^2 - 2AC\cos\beta, \\ C^2 =& A^2 + B^2 - 2AB\cos\gamma. \end{aligned}$$

For the triangle we care about, the side opposite of our angle θ is given by the vector $\vec{v} - \vec{u}$. (To see this, call this vector $\vec{w}$. Then $\vec{u} + \vec{w} = \vec{v}$, and moving $\vec{u}$ to the other side gives us $\vec{w} = \vec{v} - \vec{u}$.) See Figure 1.35.

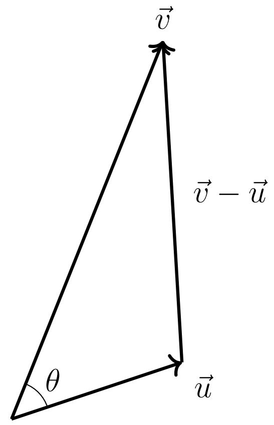


Figure 1.35: The side opposite the angle θ is $\vec{v} - \vec{u}$.

The sides of this triangle obviously have lengths $\|\vec{v}\|$, $\|\vec{u}\|$, and $\|\vec{v}-\vec{u}\|$. Using the law of cosines we have the following:

$$\|\vec{v} - \vec{u}\|^2 = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\|\vec{v}\|\|\vec{u}\| \cos\theta.$$

We can use our properties of dot products to rewrite the left-hand side. First,

$$\|\vec{v} - \vec{u}\|^2 = (\vec{v} - \vec{u}) \cdot (\vec{v} - \vec{u})$$

Now we can use the distributive property of dot products to rewrite this:

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

Using our property that $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$ and $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$, we can rewrite this as

$$\vec{v} \cdot \vec{v} - \vec{v} \cdot \vec{u} - \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{u} = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\vec{v} \cdot \vec{u}$$

Thus our equations becomes

$$\|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\vec{v} \cdot \vec{u} = \|\vec{v}\|^2 + \|\vec{u}\|^2 - 2\|\vec{v}\|\|\vec{u}\| \cos\theta.$$

This can be rewritten as

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

And, of course, putting the $\cos\theta$ by itself we have

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

Now if we want to solve for θ we have

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

Notice that this same argument would work for three-dimensional vectors. We summarize this with the following theorem.

Theorem 1.10.
Let $\vec{u}$ and $\vec{v}$ be two vectors, and θ the angle between them. Then we have the following:

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

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

(*iii*) $\theta = \arccos\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)$

A consequence of this relation between the dot product of two vectors and the angle between them is that $\vec{u}$ and $\vec{v}$ are orthogonal (meet at a $90°$ angle) if and only if $\vec{u}\cdot\vec{v} = 0$. (If they meet at a $90°$ angle, then the cosine of the angle is $\cos(90°) = 0$ which by property (i) in the above theorem means their dot product is zero.)

This observation that $\vec{u}\cdot\vec{v} = 0$ only if the vectors are orthogonal may seem trivial, but it's actually a pretty nice thing to notice. Mainly because if we care about orthogonality, we don't have to go through the process of trying to find the arccosine of some weird value: we can just do the dot product (which is easy to calculate), and see if it's zero or not.

Example 1.10.
Find the angle between the vectors $\vec{u} = \langle 3, 4, -1\rangle$ and $\vec{v} = \langle -2, 0, 2\rangle$.

Call the angle between these vectors θ, and we know

$$\begin{aligned}\cos\theta =& \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}\\ =& \frac{\langle 3,4,1\rangle\cdot\langle -2,0,2\rangle}{\sqrt{3^2+4^2+(-1)^2}\sqrt{(-2)^2+0^2+2^2}}\\ =& \frac{3\cdot(-2)+4\cdot 0+1\cdot 2}{\sqrt{9+16+1}\sqrt{4+4}}\\ =& \frac{-6+2}{\sqrt{26}\sqrt{8}}\\ =& \frac{-4}{\sqrt{208}}\end{aligned}$$

Thus

$$\theta = \cos^{-1}\left(\frac{-4}{\sqrt{208}}\right) \approx 106.1^\circ.$$

Example 1.11.
Find the angle between the vectors $\vec{u} = \left\langle \sqrt{27}-3, \sqrt{27}+3\right\rangle$ and $\vec{v} = \left\langle \sqrt{3}+1, \sqrt{3}-1\right\rangle$.

Calling the angle between these vectors θ, we know

$$\begin{aligned}
\cos\theta &= \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|} \\
&= \frac{\langle\sqrt{27}-3, \sqrt{27}+3\rangle \cdot \langle\sqrt{3}+1, \sqrt{3}-1\rangle}{\sqrt{(\sqrt{27}-3)^2+(\sqrt{27}+3)^2}\sqrt{(\sqrt{3}+1)^2+(\sqrt{3}-1)^2}} \\
&= \frac{(\sqrt{27}-3)\cdot(\sqrt{3}+1)+(\sqrt{27}+3)\cdot(\sqrt{3}-1)}{\sqrt{(36-18\sqrt{3})+(36+18\sqrt{3})}\sqrt{(4+2\sqrt{3})+(4-2\sqrt{3})}} \\
&= \frac{6+6}{\sqrt{72}\sqrt{8}} \\
&= \frac{12}{\sqrt{576}} \\
&= \frac{12}{24} \\
&= \frac{1}{2}
\end{aligned}$$

Since $\cos\theta = {}^1\!/_2$,

$$\theta = \cos^{-1}({}^1\!/_2) = 60^\circ.$$

Projections

To end our introduction to dot products, we will show how to "project" one vector onto another vector. To see why this is something you might care about doing, consider the following scenario. Suppose that an airplane is flying over an upward-sloping hill, and that the airplane's velocity vector is not parallel to the hill. If we know the velocity of the airplane, can we determine the velocity of its shadow on the hill? Notice that we already know the direction the shadow is moving in (since it's just moving up the hill), so all we really need to figure out is the speed.

In general, let's say that we have two vectors, $\vec{u}$ and $\vec{v}$, and that we want to look at the "shadow" of $\vec{v}$ in the direction of $\vec{u}$. (So, comparing to the scenario above, $\vec{v}$ is the velocity of the airplane, and $\vec{u}$ tells us the slope of the hill.) The "shadow" of $\vec{v}$ in the direction of $\vec{u}$ is called the ***(vector) projection of $\vec{v}$ onto $\vec{u}$***. This is a vector denoted $\text{proj}_{\vec{u}}\vec{v}$. If we could just figure out how long this vector is supposed to be, we'd be done:

if $\|\text{proj}_{\vec{u}}\vec{v}\| = \ell$, then we'd know

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

The reason for this is that $\frac{\vec{u}}{\|\vec{u}\|} = \frac{1}{\|\vec{u}\|}\vec{u}$ is a unit vector (a vector of length one) that points in the direction of $\vec{u}$. We then just stretch that unit vector out so that it has the length we want.

Since this length $\|\text{proj}_{\vec{u}}\vec{v}\|$ is the really important piece of information, we give it a special name. We call it the ***scalar projection of $\vec{v}$ onto $\vec{u}$***, and denote this scalar by $\text{comp}_{\vec{u}}\vec{v} = \|\text{proj}_{\vec{u}}\vec{v}\|$. To figure out what $\text{comp}_{\vec{u}}\vec{v}$ is, we consider making a right triangle with the vectors $\vec{u}$ and $\vec{v}$. See Figure 1.36.

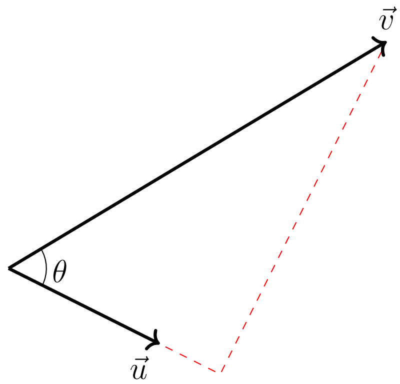


Figure 1.36: The first step in finding the projection is figuring out the lengths of the sides of this triangle.

Notice that we're able to figure out the angle θ between $\vec{u}$ and $\vec{v}$ using our properties of dot products listed above. We want to figure out the length of the side of the triangle adjacent to the vertex with angle θ, and we already know the hypotenuse of this triangle (it's just $\|\vec{v}\|$). So, we use cosine to relate these quantities:

$$\cos\theta = \frac{\text{adj}}{\text{hyp}}$$
$$\implies \text{adj} = \text{hyp} \cdot \cos\theta$$

Plugging in what we know about the angle θ and the lengths of the sides we have

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

And of course the $\|\vec{v}\|$'s cancel out to give us

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

This is the length of the projection vector. To get the projection vector we then multiply this length by $\frac{\vec{u}}{\|\vec{u}\|}$ to get

$$\text{proj}_{\vec{u}}\vec{v} = \text{comp}_{\vec{u}}\vec{v}\frac{\vec{u}}{\|\vec{u}\|} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2}\vec{u}$$

Notice in the above we never made any assumptions about the dimension of the vectors $\vec{u}$ or $\vec{v}$: this construction works both in two dimensions and in three dimensions.

Example 1.12.
Calculate the scalar and vector projections of $\vec{v} = \langle 4, 1\rangle$ onto $\vec{u} = \langle 2, -3\rangle$.

First we'll calculate the scalar projection:

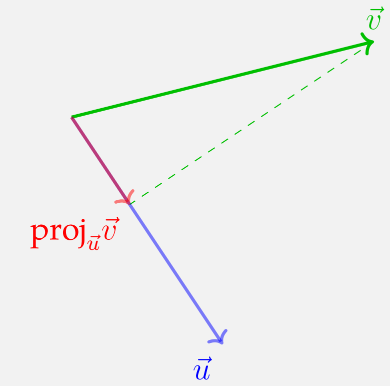
$$\begin{aligned}\text{comp}_{\vec{u}}\vec{v} =& \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|}\\ =& \frac{\langle 2, -3\rangle \cdot \langle 4, 1\rangle}{\|\langle 2, -3\rangle\|}\\ =& \frac{2 \cdot 4 - 3 \cdot 1}{\sqrt{2^2 + 3^2}}\\ =& \frac{5}{\sqrt{13}}\end{aligned}$$

Notice that in calculating this quantity we found $\|\vec{u}\| = \sqrt{13}$.

Now to find the vector projection:

$$\begin{aligned}\text{proj}_{\vec{u}}\vec{v} =& \frac{5}{\sqrt{13}}\frac{\vec{u}}{\sqrt{13}}\\ =& \frac{5}{13}\vec{u}\\ =& \frac{5}{13}\langle 2, -3\rangle\\ =& \langle {}^{10}/_{13}, -{}^{15}/_{13}\rangle\end{aligned}$$

These vectors are graphed below.



Again, notice that the projection is like the "shadow" of $\vec{v}$ onto $\vec{u}$, as indicated by the dashed line.

Application: Work Done by a Constant Force

In physics, *work* is defined as force applied over a distance. The SI units for work, thus, are Newton-meters (Nm) which are usually referred to as *joules* (J). (The English units of work are foot-pounds.) This is all well and good if the force is in the same direction as the movement, but suppose the force vectors points in a direction different from the direction of motion. For example, imagine you want to move a heavy weight on the ground. If the weight is too heavy for you to lift with your arms, you might tie a rope to it, throw the rope over your shoulder, and then use your legs to pull the weight by walking. The force pulling the weight (the tension in the rope) pulls at some angle, but the weight only moves horizontally. How can we determine the work done in this case?

What we need to do is determine how much of the force points in the horizontal direction. To do this we can use projections: we project the force vector onto the horizontal vector $\vec{\imath} = \langle 1, 0 \rangle$, and this will tell us how much force is applied in the horizontal direction. Finally we multiply this by the distance moved.

In general, let's suppose we have a force vector $\vec{F}$ that we apply across a displacement vector $\vec{d}$. We want to know how much of the force points in direction $\vec{d}$, which we can figure out with projections, and then multiply this value by $\|\vec{d}\|$ to get the work done.

$$
\begin{aligned}
W =& \|\text{proj}_{\vec{d}}\vec{F}\|\|\vec{d}\| \\
=& \text{comp}_{\vec{d}}\vec{F}\,\|\vec{d}\| \\
=& \frac{\vec{F}\cdot\vec{d}}{\|\vec{d}\|}\,\|\vec{d}\| \\
=& \vec{F}\cdot\vec{d}
\end{aligned}
$$

Example 1.13.
Suppose you pull a heavy weight 20 feet by tying a rope to it, throwing the rope over your shoulder, and walking. Say the rope makes a 30° angle with the ground, and there's 100 pounds of tension in the rope. How much work is done?

Notice that the components of the force and displacement vectors aren't given. We have enough information to find the components, but we don't actually need to: we can do this problem, using properties of dot products, without writing down the components.

$$
\begin{aligned}
W =& \vec{F}\cdot\vec{d} \\
=& \|\vec{F}\|\|\vec{d}\|\cos(30^\circ) \\
=& 100\cdot 20\cdot\frac{\sqrt{3}}{2} \\
=& 1000\sqrt{3}\text{ft lb}
\end{aligned}
$$

1.6 The cross product

In the last section we introduced the dot product of two vectors, which was a way of associating a scalar to a given pair of vectors, and this scalar told us some geometric information about those vectors. In particular, the scalar gave us some information about the angle between the vectors. In this lecture we'll introduce an operation where we take two vectors and associate a third vector that, again, tells us some geometric information about those vectors.

Let's say that $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$ are two vectors in 3-space. The *cross product* of $\vec{u}$ and $\vec{v}$, denoted $\vec{u} \times \vec{v}$, is defined as follows:

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

This probably looks strange, but here's the important thing: this is actually a vector. To see this, let's just calculate what the determinant of the above matrix should be:

$$\begin{aligned} \vec{u} \times \vec{v} =& \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{bmatrix} \\ =& \vec{\imath}(u_2 v_3 - u_3 v_2) - \vec{\jmath}(u_1 v_3 - u_3 v_1) + \vec{k}(u_1 v_2 - u_2 v_1) \\ =& \vec{\imath}(u_2 v_3 - u_3 v_2) + \vec{\jmath}(u_3 v_1 - u_1 v_3) + \vec{k}(u_1 v_2 - u_2 v_1) \\ =& \langle u_2 v_3 - u_3 v_2,\ u_3 v_1 - u_1 v_3,\ u_1 v_2 - u_2 v_1 \rangle \end{aligned}$$

Example 1.14.
Calculate the cross product $\vec{u} \times \vec{v}$ of $\vec{u} = \langle 1, 0, 2 \rangle$ and $\vec{v} = \langle 3, -2, -4 \rangle$.

$$\begin{aligned} \vec{u} \times \vec{v} =& \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 0 & 2 \\ 3 & -2 & -4 \end{bmatrix} \\ =& \vec{\imath}(0 \cdot (-4) - 2 \cdot (-2)) - \vec{\jmath}(1 \cdot (-4) - 2 \cdot 3) + \vec{k}(1 \cdot (-2) - 0 \cdot 3) \\ =& \langle 4, 10, -2 \rangle \end{aligned}$$

Recall that one of the properties of dot products is that the dot product of $\vec{u}$ and $\vec{v}$ told us some information about the angle between $\vec{u}$ and $\vec{v}$. In particular, if the angle between $\vec{u}$ and $\vec{v}$ is θ, then

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

There's a similar relation between cross products and angles.

Theorem 1.11.
Let $\vec{u}$ and $\vec{v}$ be two three-dimensional vectors. Then

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

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

A useful corollary of this theorem is that two vectors are parallel if and only if their cross product is zero.

One of the important things about this operation, the cross product of two vectors, is that it's only defined for three-dimensional vectors.[4] If you wanted to take the cross product of two two-dimensional vectors, $\vec{u} = \langle u_1, u_2 \rangle$ and $\vec{v} = \langle v_1, v_2 \rangle$, you'd first have to stick them into three-dimensional space as $\langle u_1, u_2, 0 \rangle$ and $\langle v_1, v_2, 0 \rangle$. Before we describe various properties of the cross product in general, let's consider the cross product of two vectors in the xy-plane:

Example 1.15.
Calculate the cross product $\vec{u} \times \vec{v}$ of $\vec{u} = \langle 1, 2, 0 \rangle$ and $\vec{v} = \langle 2, 3, 0 \rangle$.

$$\begin{aligned} \vec{u} \times \vec{v} = &\det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 2 & 0 \\ 2 & 3 & 0 \end{bmatrix} \\ =& \vec{\imath}(0-0) - \vec{\jmath}(0-0) + \vec{k}(3-4) \\ =& \langle 0, 0, -1 \rangle \end{aligned}$$

Let's notice two things about this vector we've calculated: it's orthogonal to the two vectors we started with (they were in the xy-plane, but this points straight along the z-axis), and its magnitude is equal to the area of the parallelogram determined by our vectors. This is true in general, not just for vectors in the xy-plane.

[4] This is a little bit of a white lie. There is a way to generalize the cross product to vectors of n-dimensional space for $n > 3$, but the result isn't a normal vector. We will never need the higher-dimensional analogue of a cross product in this class, so we won't take the time to go through the details of that operation.

Theorem 1.12.
Let $\vec{u}$ and $\vec{v}$ be two three-dimensional vectors. Then

(i) *$\vec{u} \times \vec{v}$ is orthogonal to both $\vec{u}$ and $\vec{v}$.*

(ii) *$\|\vec{u} \times \vec{v}\|$ is the area of the parallelogram with sides $\vec{u}$ and $\vec{v}$.*

Let's notice two things about these properties: they *almost* give us a geometric definition of the cross product. If we're given two vectors, $\vec{u}$ and $\vec{v}$, then we know that $\vec{u} \times \vec{v}$ is orthogonal to both of these vectors. This tells us that $\vec{u} \times \vec{v}$ points along some line. Furthermore, we know the magnitude of $\vec{u} \times \vec{v}$ has to equal the area of the parallelogram with sides $\vec{u}$ and $\vec{v}$. This limits us to two particular vectors.

Just to make things clear, consider Figure 1.37. Our two possible vectors are marked $\vec{w}_1$ and $\vec{w}_2$.

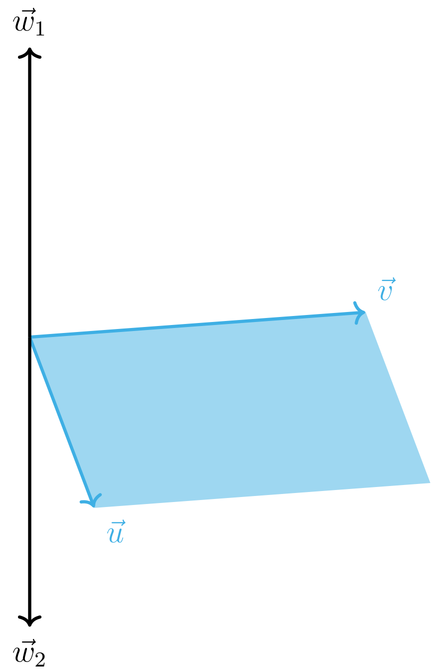


Figure 1.37: There are two vectors, $\vec{w}_1$ and $\vec{w}_2$, which satisfy the properties of the cross product from Theorem 1.12.

We need some convention for determining whether $\vec{u} \times \vec{v}$ should be $\vec{w}_1$ or $\vec{w}_2$. This convention is known as the ***right-hand rule*** and says that if you take your right hand and point your fingers in the direction of $\vec{u}$ so that your fingers can curl towards $\vec{v}$, then your thumb points in the direction of $\vec{u} \times \vec{v}$. So in Figure 1.37, the vector labelled $\vec{w}_1$ points is $\vec{u} \times \vec{v}$.

Notice that, because of this convention, $\vec{v} \times \vec{u} = -\vec{u} \times \vec{v}$. This property is sometimes called the *anti-commutativity* of the cross product. (Operations like vector addition or the dot product where you can change the order of the things you're adding or dotting together and not change the outcome – i.e., $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ and $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ – are called *commutative*. The cross product is almost commutative, except that the sign changes, so we call it *anti-commutative*.)

We of course want to know how our cross product operation interacts with the other operations we've introduced for vectors.

Theorem 1.13.
Let $\vec{u}$, $\vec{v}$, and $\vec{w}$ be three-dimensional vectors, and let $\lambda \in \mathbb{R}$ be a scalar. The following properties hold.

(*i*) $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$

(*ii*) $(\lambda \vec{u}) \times \vec{v} = \lambda(\vec{u} \times \vec{v}) = \vec{u} \times (\lambda \vec{v})$

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

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

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

(*vi*) $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}$

Notice that property (vi) in particular tells us that $\vec{u} \times (\vec{v} \times \vec{w}) \neq (\vec{u} \times \vec{v}) \times \vec{w}$. So, you can't just assume that the properties you expect "multiplication" to have work for the cross product!

The scalar triple product

If we combine the dot product and scalar product as in property (v) of the theorem above, we have what's called the *scalar triple product*, which is really just the determinant of a certain 3×3 matrix:

$$\begin{aligned}\vec{u} \cdot (\vec{v} \times \vec{w}) =& \langle u_1, u_2, u_3 \rangle \cdot \langle v_2 w_3 - v_3 w_2, w_3 v_1 - v_1 w_3, v_1 w_2 - v_2 w_1 \rangle \\ =& u_1(v_2 w_3 - v_3 w_2) + u_2(w_3 v_1 - v_1 w_3) + u_3(v_1 w_2 - v_2 w_1) \\ =& u_1(v_2 w_3 - v_3 w_2) - u_2(v_1 w_3 - w_3 v_1) + u_3(v_1 w_2 - v_2 w_1) \\ =& \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}\end{aligned}$$

This tells us that the absolute value of $\vec{u} \cdot (\vec{v} \times \vec{w})$ is the volume of the parallelepiped whose sides are the vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$.

In fact, you can use our properties of cross products above to prove our earlier claim that the determinant of a 3×3 matrix gives you the (signed) volume of certain parallelepiped. I won't go through that in these notes, but you should try to prove that for yourself to make sure you understand everything. The key is to use Cavalieri's principle to say that the volume of the parallelepiped is equal to the area of the base times the height, and then use our property $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \|\vec{v}\| \sin\theta$ to help determine that height.

A nice consequence of this scalar triple product stuff is that it gives us a way to figure out if four points are coplanar (all live in the same plane or not). In two dimensions, any two points determine in a line. In three dimensions, any three non-colinear points determine a plane. (That is, as long as the three points you've been given don't live on a single line, they determine a unique plane.) You can pick one of these points to be the corner of a parallelepiped and use the displacement vectors to the other points for the sides of the parallelepiped, and then use the scalar triple product to find the volume of that parallelepiped. If the parallelepiped has zero volume, then the four points are coplanar.

Application: torque

We end this lecture by mentioning one physical application of cross products. Physics is all about studying natural phenomena, and trying to develop laws that explain these phenomena. One such phenomenon is *torque*, which just refers to the fact that if a (rigid) object is tethered or anchored down at some point and a force is then applied to that object, the object tends to rotate around the point where it's anchored. This point is usually called a *pivot*.

So torque is supposed to tell us how an object spins around a pivot. To describe this motion we'll construct a vector whose direction is orthogonal to the plane the object is spinning in and whose magnitude tells us how "easily" the object spends.

More specifically, suppose that a force of $\vec{F}$ is applied to the object at some point, and the displacement from the pivot to the point where the force is applied is given by $\vec{d}$. The torque, denoted $\vec{\tau}$, is then defined to be

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

The units of torque are those Newton-meters, which as we mentioned when we discussed work, are usually just called joules.

Example 1.16.
Imagine a wrench which is one foot long being used to tighten a bolt. If you apply a force of 3 pounds at a $45°$ angle from the end of the wrench, what is the torque vector?

In this problem, $\vec{d} = \langle 1, 0, 0\rangle$ and $\vec{F} = \left\langle \frac{3\sqrt{2}}{2}, \frac{3\sqrt{2}}{2}, 0\right\rangle$. Thus

$$\begin{aligned}\vec{\tau} =& \vec{F} \times \vec{d} \\ =& \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{3\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ =& \left\langle 0, 0, -\frac{3\sqrt{2}}{2}\right\rangle\end{aligned}$$

The magnitude of the torque is thus $\frac{3\sqrt{2}}{2}$ J.

Notice in the above problem that if we only cared about the magnitued, we could use the formula

$$\begin{aligned} \vec{\tau} &= \vec{F} \times \vec{d} \\ \implies \|\vec{\tau}\| &= \|\vec{F}\|\|\vec{d}\| \sin\theta \end{aligned}$$

and this would have told us $\|\vec{\tau}\| = 3 \cdot 1 \cdot \sin 45° = \frac{3\sqrt{2}}{2}$.

This formula also makes it clear that if you want to maximize the torque, then the force and displacement vectors need to be orthogonal: this forces $\sin\theta$ to be 1. In any other case, $\sin\theta < 1$.

1.7 Lines and planes

In this section we derive the equations for lines and planes living in 3-space, as well as define the angle between two non-parallel planes, and determine the distance from a point to a plane using properties of vector projections.

Lines

Vector, parametric, and symmetric equations

Recall that in two dimensions, to specify a line you need two pieces of information: a point the line passes through, and the slope of the line. The slope of the line really just tells us the direction the line points in. In three dimensions we also need two pieces of information to determine a line; if we want to give the equation of a line, then we need to know a point on the line and the direction of the line. While in two dimensions we could use a single number (the slope) to determine the direction of the line, in three dimensions we'll use vectors. (Really, the slope in two dimensions determines a vector: a slope of m is the same as a vector $\langle 1, m\rangle$. Since the first component of this vector is always 1, it's only the second component that matters.)

So let's say, to keep things easy, we want the equation of a line through the origin. Let's suppose the vector $\vec{v} = \langle a, b, c\rangle$ points along the line. This means that if a point $P = (x, y, z)$ lives on the line, we could perform the scalar multiplication $t\vec{v}$ to stretch $\vec{v}$ out enough so that its tip was at P. See Figure **??** on page ??.

That is, the displacement vector $\vec{r} = \overrightarrow{OP} = \langle x, y, z\rangle$ is related to the direction vector $\vec{v} = \langle a, b, c\rangle$ by the following equation:

$$\vec{r} = t\vec{v}$$

This is the ***vector equation*** of the line through the origin in direction $\vec{v}$.

Now let's suppose we want our line to pass through some point $P_0 = (x_0, y_0, z_0)$ instead of the origin. Let $\vec{r}_0$ denote the displacement vector $\vec{r}_0 = \overrightarrow{OP_0} = \langle x_0, y_0, z_0\rangle$. All we need to do is take our line through the origin, and translate it along the vector $\vec{r}_0$ so that the line passes through P_0. See Figure 1.39 on the next page.

So, if another point $P = (x, y, z)$ is to be on the line, letting $\vec{r}$ denote the vector $\langle x, y, z\rangle$, we have the following equation:

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

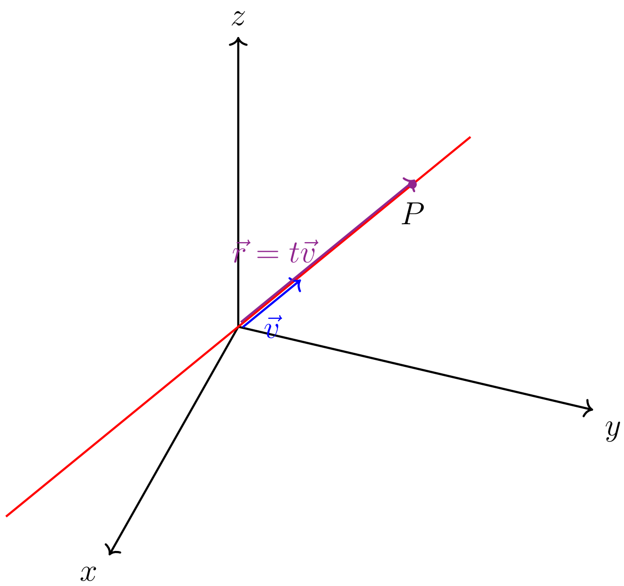


Figure 1.38: Every point on a line through the origin can be obtained by a vector which is a scalar multiple of the line's direction vector.

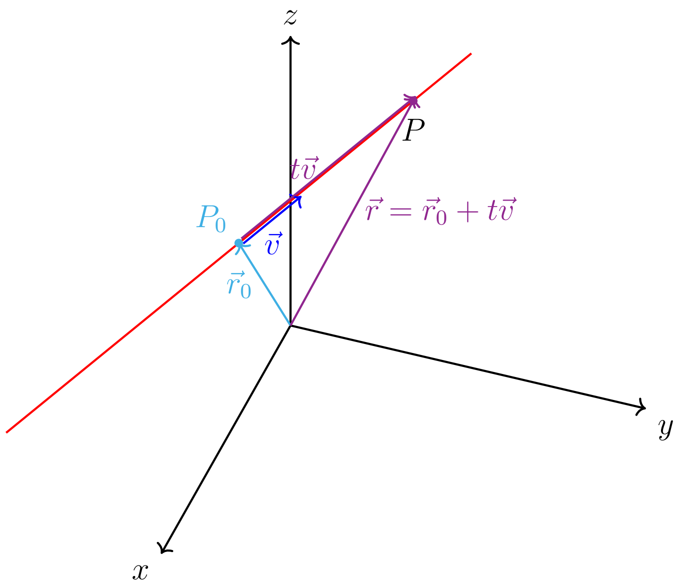


Figure 1.39: We can translate a line through the origin so that it goes through another point by adding a vector to each point on the line.

This is the ***vector equation*** of the line through P_0 in the direction of $\vec{v}$. If we write this out in terms of components we have

$$\begin{aligned}\langle x, y, z\rangle &= \langle x_0, y_0, z_0\rangle + \langle ta, tb, tc\rangle \\ &= \langle x_0 + ta, y_0 + ta, z_0 + ta\rangle .\end{aligned}$$

Two vectors are equal if and only if their components are equal, so by equating components we actually have three equations:

$$\begin{aligned}x =& x_0 + ta \\ y =& y_0 + tb \\ z =& z_0 + tc.\end{aligned}$$

Notice that on the right-hand side the values a, b, c, x_0, y_0, z_0 are all fixed: they don't change once we say what the point P_0 is and where the vector $\vec{v}$ points. The t is the only thing on the right-hand side that can change. Thus the values x, y, z on the left-hand side are functions of t. So really the above equations should be written as

$$\begin{aligned}x(t) =& x_0 + ta \\ y(t) =& y_0 + tb \\ z(t) =& z_0 + tc.\end{aligned}$$

These three equations form the ***parametric equations*** of the line through $P_0 = (x_0, y_0, z_0)$ in the direction of $\vec{v} = \langle a, b, c\rangle$.

Example 1.17.
Find the vector and parametric equations for the line through the point $(3, -4, 1)$ in the direction of $\langle 1/2, -1, 3\rangle$.

The vector equation is

$$\langle x, y, z\rangle = \left\langle 3 + \frac{t}{2}, -4 - t, 1 + 3t\right\rangle .$$

To get the parametric equations, just equate the components to get:

$$\begin{aligned}x(t) =& 3 + \frac{t}{2} \\ y(t) =& -4 - t \\ z(t) =& 1 + 3t.\end{aligned}$$

Now, assuming none of a, b, or c is zero, we could solve each of our equations above for t to get

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

Since each of the things on the right-hand side equals t, these quantities are all equal. Thus we have

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

These are the ***symmetric equations*** of the line through $P_0 = (x_0, y_0, z_0)$ in the direction of $\vec{v} = \langle a, b, c \rangle$. Again, notice these equations only make sense if none of a, b, or c equals zero. (If one of them did equal zero, we'd have division by zero.)

In our example above, the symmetric equations are

$$\frac{x - 3}{1/2} = \frac{y + 4}{-1} = \frac{z - 1}{3}.$$

These symmetric equations are nice because they give us an easy way to determine if a point is on the line or not.

Example 1.18.
Are the points $(4, -6, 7)$ and $(-5, 0, -14)$ on the line through $(3, -4, 1)$ in the direction of $\langle 1/2, -1, 3 \rangle$?

This is our line from the last example, so the symmetric equations are given above. We plug in the coordinates of each point to see if we have equality or not. In the case of $(4, -2, 7)$ we have:

$$\frac{4 - 3}{1/2} = \frac{-6 + 4}{-1} = \frac{7 - 1}{3}.$$

Each of these expressions is 2, so we have equality, and the point $(4, -2, 7)$ is on the line.

Plugging in $(-5, 0, -14)$ we have

$$\frac{-5-3}{1/2} = \frac{8+0}{-1} \neq \frac{-12-1}{3}.$$

The first two expressions equal -4, while the last one equals -5. Thus the point $(-5, 0, -14)$ is not on the line.

Parallel and skew lines

Just as in two dimensions, two lines are ***parallel*** if they point in the same direction. In the case of two dimensions this meant that the two lines had the same slope. In the case of three dimensions it means that the two direction vectors are scalar multiples of one another. For example, the lines

$$\frac{x-2}{3} = \frac{y+1}{1} = \frac{z-1}{-2}$$

and

$$\frac{x+1}{-9} = \frac{y}{-3} = \frac{z-2}{6}$$

are parallel. These lines have respective direction vectors $\langle 3, 1, -2\rangle$ and $\langle -9, -3, 6\rangle$. The second one is -2 times the first.

In the case of two dimensions, two lines are parallel if and only if they never touch. This is *not* the case in three dimensions. Consider the lines with the following parametric equations.

$$\begin{aligned} x_1(t) &= 1 - t \\ y_1(t) &= 1 + 2t \\ z_1(t) &= 0 \end{aligned}$$

$$\begin{aligned} x_1(t) &= 4 - t \\ y_1(t) &= 2 + t \\ z_1(t) &= 4 \end{aligned}$$

Let's call these lines L_1 and L_2. From the equations we can already tell that the first line passes through the point $(1, 1, 0)$ in the direction of $\vec{v}_1 = \langle -1, 2, 0\rangle$, and the second line passes through the point $(4, 2, 4)$ in the direction of $\vec{v}_2 = \langle -1, 1, 0\rangle$. Notice that L_1 is contained in the plane $z = 0$, and L_2 is in the plane $z = 4$. See Figure 1.40. (Because of perspective it may appear that the lines intersect in Figure 1.40, but they never actually intersect since they are contained in different planes parallel to the xy-plane.) Thus there's no possible way these lines can ever intersect. However, their direction vectors are not scalar multiples of one another: there is no $\lambda \in \mathbb{R}$ that makes $\vec{v}_1 = \lambda \vec{v}_2$. So these lines are not parallel, nor do they intersect. When this happens we say that the lines are ***skew***.

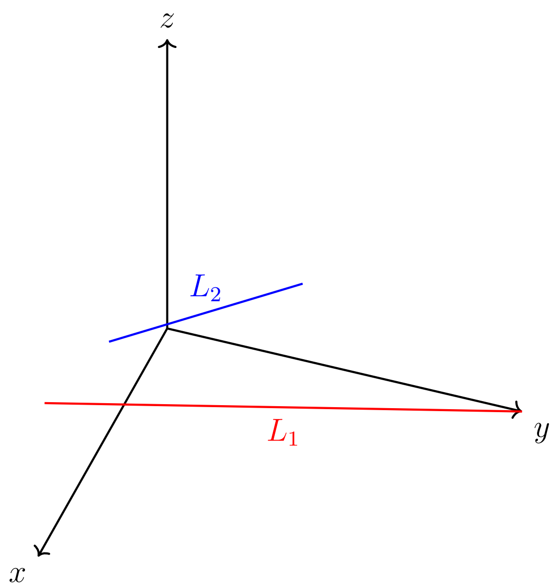


Figure 1.40: Lines are *skew* if they are never intersect but are not parallel.

Planes

Equations of planes

Now we consider planes. Of course we've already seen several examples of equations of planes in previous lectures, but everything we looked at before was a very special case (e.g., the coordinate planes). We want to be able to determine the equation of any plane inside of 3-space, even if it's "tilted" at a strange angle with respect to the coordinate planes.

Any plane is determined by two pieces of information: a point $P_0 = (x_0, y_0, z_0)$ contained in the plane and a vector $\vec{n} = \langle n_1, n_2, n_3 \rangle$ orthogonal to the plane. This vector is called a ***normal vector*** to the plane. To see why these two pieces of information are sufficient for specifying a plane, imagine that plane is at the tail of the normal vector. If you want to rotate the plane around, it's the same as rotating the normal vector around: if you know the "tilt" of the normal vector, then you know the "tilt" of the plane.

Let's first suppose our plane goes through the origin to make things easy. Then what we want is the collection of all points $P = (x, y, z)$ which are orthogonal to the given normal vector, $\vec{n} = \langle n_1, n_2, n_3 \rangle$. If we let $\vec{r} = \overrightarrow{OP}$ denote the displacement vector from the origin to P, then we have to have that $\vec{r} \cdot \vec{n} = 0$. (Recall that this was one of our properties of dot products: two vectors are orthogonal if and only if their dot product is zero.)

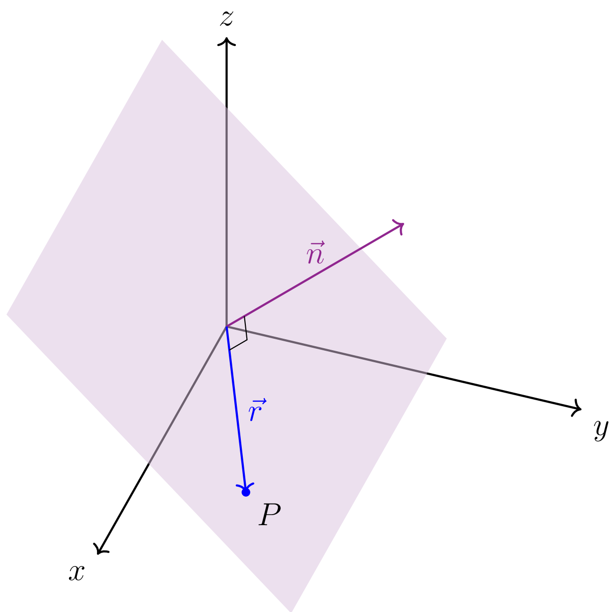


Figure 1.41: Every vector inside of a plane is orthogonal to the plane's normal vector.

Now suppose that instead of going through the origin, we want the plane to go through the point $P_0 = (x_0, y_0, z_0)$. Let $\vec{r}_0 = \overrightarrow{OP_0}$ and let $\vec{n} = \langle n_1, n_2, n_3 \rangle$ be the normal vector. We want to know if a point $P = (x, y, z)$ lives on this plane or not. This is not quite the same as saying $\vec{r} = \overrightarrow{OP} = \langle x, y, z \rangle$ is orthogonal to $\vec{n}$: the vector $\vec{r}$ points at some point in our plane, but the vector itself doesn't lie in the plane. However, the vector $\vec{r} - \vec{r}_0$ *does* live in the plane and point to our point P (the tail is at P_0 and the tip is at P_1). Thus what we really want is $\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0$. This is the ***vector equation of the plane containing P_0 with normal vector $\vec{n}$***.

Writing this vector equation out in components, we have the following:

$$
\begin{aligned}
&\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0 \\
\implies &\langle n_1, n_2, n_3 \rangle \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0 \\
\implies &n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0.
\end{aligned}
$$

This is the ***scalar equation of the plane through the point*** (x_0, y_0, z_0) ***with normal vector*** $\langle n_1, n_2, n_3 \rangle$.

Notice that we could distribute the n_i and simplify and our equation would become

$$n_1 x + n_2 y + n_3 z + d = 0$$

where $d = -(n_1 x_0 + n_2 y_0 + n_3 z_0)$. This is a ***linear equation***, and just like the symmetric equations for lines, they give us a really easy way to determine if a point is on a plane or not.

Example 1.19.
Find the linear equation of the plane which contains the point $(-2, 3, -5)$ and whose normal vector is $\langle 7, 2, 3 \rangle$. Is the point $(1, 3, -4)$ on this plane?

To get the linear equation:

$$
\begin{aligned}
&7(x + 2) + 2(y - 3) + 3(z + 5) = 0 \\
\implies &7x + 14 + 2y - 6 + 3z + 15 = 0 \\
\implies &7x + 2y + 3z + 23 = 0
\end{aligned}
$$

Now we check if $(1, 3, -4)$ is on the plane or not:

$$7 \cdot 1 + 2 \cdot 3 + 3 \cdot (-4) + 23 = 24 \neq 0$$

So the point is not on the plane.

Distance to a plane

Recall that if $P = (x, y, z)$ is some point in 3-space, we can measure the distance from that point to each of the coordinate planes by first projecting onto the plane, and then measuring the distance from the projection to our initial point. We can do the exact same thing but for other planes. Let's say that we're given a point $P_1 = (x_1, y_1, z_1)$ and we want to measure the distance from this point to the plane

$$ax + by + cz = d.$$

What we want to do is find the point $P_0 = (x_0, y_0, z_0)$ on the plane so that the displacement vector $\overrightarrow{P_0P_1}$ sticks orthogonally out of the plane. That is, $\overrightarrow{P_0P_1} = \lambda\vec{n}$ where $\vec{n} = \langle a, b, c\rangle$ is the normal vector of the plane. We then need to figure out how long this displacement vector is.

Let's suppose that $P = (x, y, z)$ is any other point in this plane. Then if $\vec{r} = \overrightarrow{OP}$ and $\vec{r}_0 = \overrightarrow{OP_0}$, we know

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0.$$

Let $\vec{s} = \overrightarrow{PP_1}$. What we want to do is project $\vec{s}$ onto $\vec{n}$, and measure the length of that vector – that is, we want the absolute value of $\text{comp}_{\vec{n}}\vec{s}$. In components, $\vec{s} = \langle x_1 - x, y_1 - y, z_1 - z\rangle$. Thus the distance from $P_1 = (x_1, y_1, z_1)$ to the plane $ax + by + cz = d$ is given by

$$\begin{aligned}
\text{distance} &= \left|\text{comp}_{\vec{n}}\vec{s}\right| \\
&= \left|\frac{\vec{s} \cdot \vec{n}}{\|\vec{n}\|}\right| \\
&= \left|\frac{\langle x_1 - x, y_1 - y, z_1 - z\rangle \cdot \langle a, b, c\rangle}{\sqrt{a^2 + b^2 + c^2}}\right| \\
&= \frac{|a(x_1 - x) + b(y_1 - y) + c(z_1 - z)}{\sqrt{a^2 + b^2 + c^2}} \\
&= \frac{|ax_1 + by_1 + cz_1 - (ax + by + cz)|}{\sqrt{a^2 + b^2 + c^2}}
\end{aligned}$$

Now, since our point $P = (x, y, z)$ lives on the plane, it satisfies the equation $ax + by + cz = d$, so the above becomes

$$\text{distance} = \frac{|ax_1 + by_1 + cz_1 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Caution: The formula that appears in your book is slightly different than what's written here because your book assumes the equation of the plane is written as $ax + by + cz + d = 0$, which becomes $ax + by + cz = -d$: our d is the negative of the d in the book. So you have to be a slightly cautious when using these formulas.

Example 1.20.
Find the distance from the point $(1, 2, -4)$ to the plane $3x - y + 5z = 6$.

Here the normal vector to plane is $\vec{n} = \langle 3, -1, 5 \rangle$, $d = 6$, and $(x_1, y_1, z_1) = (1, 2, -4)$. Plugging these values into our formula above,

$$\begin{aligned} \text{distance} =& \frac{|ax_1 + by_1 + cz_1 - (ax + by + cz)|}{\sqrt{a^2 + b^2 + c^2}} \\ =& \frac{|3 - 2 - 20 - 6|}{\sqrt{9 + 1 + 25}} \\ =& \frac{25}{\sqrt{35}} \\ \approx& 4.23. \end{aligned}$$

The angle between two planes

We'll say that two planes are ***parallel*** if their normal vectors are scalar multiples of one another.

Example 1.21.
Are the two planes below parallel?

$$\begin{aligned} 3x + 6y - 7z =& 13 \\ -x - 2y + \frac{7}{3}z =& 0 \end{aligned}$$

From the equations, we can easily pull of the normal vectors: $\langle 3, 6, -7 \rangle$ and $\langle -1, -2, 7/3 \rangle$. These vectors are obviously scalar multiples:

$$\langle 3, 6, -7 \rangle = -3 \langle -1, -2, 7/3 \rangle .$$

Thus the planes are parallel.

Example 1.22.
Are the two planes below parallel?

$$\begin{aligned} 2x - y + z =& 3 \\ 4x - y + 2x =& 6 \end{aligned}$$

Again, we pull off the normal vectors: $\langle 2, -1, 1\rangle$ and $\langle 4, -1, 2\rangle$. Notice these vectors can't be scalar multiples of one another: we'd need to multiply the first vector by 2 to get a 4 in the first component of the second vector, but then this would give a -2 in the second component. Hence the planes are not parallel.

If two planes are not parallel, we can talk about the angle at which those planes meet. Of course, geometrically, this is exactly what you think it should be, but how do you go about calculating this angle? It's relatively clear that the normal vectors to the planes meet at the same angle as the planes themselves, and we can calculate the angle between the vectors by using dot products. That is, if our planes have normal vectors $\vec{n} = \langle n_1, n_2, n_3\rangle$ and $\vec{m} = \langle m_1, m_2, m_3\rangle$, then the angle between the planes is the same as the angle between these vectors, which we know is just

$$\theta = \cos^{-1}\left(\frac{\vec{m}\cdot\vec{n}}{\|\vec{m}\|\|\vec{n}\|}\right).$$

Example 1.23.
What's the angle between the planes given by the equations below?

$$\begin{aligned} 2x + 3y - z =& 0 \\ 3x + y + 4z =& 12 \end{aligned}$$

Our normal vectors are $\langle 2, 3, -1\rangle$ and $\langle 3, 1, 4\rangle$. Notice these vectors aren't scalar multiples of one another, so the planes aren't parallel and must intersect. The angle between the planes is the same as the angle between these two vectors:

$$\begin{aligned}\theta =& \cos^{-1}\left(\frac{\langle 2,3,-1\rangle \cdot \langle 3,1,4\rangle}{\|\langle 2,3,-1\rangle\| \, \|\langle 3,1,4\rangle\|}\right)\\ =& \cos^{-1}\left(\frac{6+3-12}{\sqrt{4+9+1}\,\sqrt{9+1+16}}\right)\\ =& \cos^{-1}\left(\frac{-3}{\sqrt{364}}\right)\\ \approx& 99.05^\circ\end{aligned}$$

Curves 2

Mathematics is not about numbers, equations, computations or algorithms: it is about understanding.

BILL THURSTON

In the previous chapter we were essentially interested in "linear" quantities: vectors, linear transformations, lines, and planes. We made a couple of diversions to discuss some other things like spheres, but the bulk of the first chapter of the notes was really about things that were straight and flat. In this chapter we start to study more general objects by replacing lines with curves. As we will see, however, one of the main motifs of calculus is that we should approximate non-linear objects with linear ones. As the course progresses we will see this idea reappear in various ways, but to begin we first need to discuss how to represent curves in space.

2.1 Vector-valued functions and parametric curves

In this section we'll introduce the notion of a vector-valued function, discuss the relationship between vector-valued functions and curves, and finally start doing some calculus by talking about limits and continuity of vector-valued functions.

Review of parametric curves

In your second semester of calculus you learned about parametric curves, which were curves described by two functions, $x(t)$ and $y(t)$, which individually told you the x- and y-coordinates of points on the curve. Just to refresh your memory, let's work through one very simple example. Suppose that $x(t) = \cos(t)$ and $y(t) = \sin(t)$. At a given value of t, the point $(x(t), y(t)) = (\cos(t), \sin(t))$ is a point on the curve. For example, the points

$$\begin{aligned}(\cos(0),\sin(0)) =&(1,0)\\ (\cos(\pi/6),\sin(\pi/6)) =&\left(\frac{\sqrt{3}}{2},\frac{1}{2}\right)\\ (\cos(\pi/4),\sin(\pi/4)) =&\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)\\ (\cos(\pi/3),\sin(\pi/3)) =&\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)\\ (\cos(\pi/2),\sin(\pi/2)) =&(0,1)\end{aligned}$$

are all on the curve. If we take all of the points we could possibly get for all values of $t \in \mathbb{R}$, that collection of points gives us the circle of radius one centered at the origin.

In general, if $x, y : \mathbb{R} \to \mathbb{R}$ are continuous functions, then the set

$$\left\{(x(t), y(t)) \,\middle|\, t \in \mathbb{R}\right\}$$

forms a curve called a ***parametric curve***.

We want to do the exact same thing for curves in 3-space. So, of course, all we have to do is add in a function to tell us the z-coordinate of points on the curve. That is, we have three continuous functions $x, y, z : \mathbb{R} \to \mathbb{R}$, and consider the set of points

$$\left\{(x(t), y(t), z(t)) \,\middle|\, t \in \mathbb{R}\right\}.$$

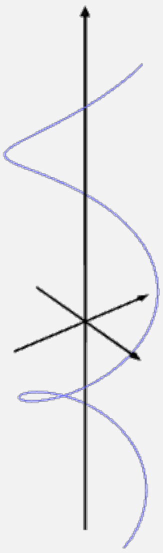
This is a parametric curve in three-dimensional space.

Example 2.1.
What does the curve given by

$$\begin{aligned}x(t) =& \cos(t)\\ y(t) =& \sin(t)\\ z(t) =& t\end{aligned}$$

look like?

If we ignored the z-coordinate, we'd have a unit circle in the xy-plane. The z-coordinate just changes as t changes, though: so as we increase t to move counter-clockwise around the circle, the z-value of our point is increasing. This gives us a curve that winds around the z-axis. Such a curve is called a ***helix***, and is plotted in the figure below.



Vector-valued functions

In the case of parametric curves, we take three continuous functions of a single variable, and glue them together to get a point in space. We could just as well think of this as a function which takes in a single real number, and spits out a point in space. That is, we could think of these three separate functions as being one function which takes an input t and spits out three different numbers, $x(t)$, $y(t)$, and $z(t)$.

In general, a function which takes a single value t, and associates to it three numbers is an example of a ***vector-valued function***: $\vec{r} : \mathbb{R} \to \mathbb{R}^3$. So, the input to this function is a real number, but the out is a 3-dimensional vector: hence the arrow over the function's name. Any such function can always be thought of as three separate functions which give the x-, y-, and z-components of the vector. These are the ***component functions*** of the vector valued function:

$$\begin{aligned}\vec{r}(t) &= \langle x(t), y(t), z(t) \rangle \\ &= x(t)\,\vec{\imath} + y(t)\,\vec{\jmath} + z(t)\,\vec{k}.\end{aligned}$$

In the case of curves we have to assume that each of these functions was continuous, but for a general vector valued function we don't need to make such an assumption.

A vector-valued function $\vec{r}(t)$ is defined only where all of its component functions are defined. This is, of course, called the ***domain*** of the function.

Example 2.2.
What's the domain of the following vector-valued function?

$$\vec{r}(t) = \left\langle \sqrt{9-t^2}, t, \sin^{-1}(2+t) \right\rangle$$

Here, our component functions are

$$\begin{aligned} x(t) =& \sqrt{9-t^2} \\ y(t) =& t \\ z(t) =& \sin^{-1}(3-t). \end{aligned}$$

The function $\vec{r}(t)$ is defined only if each one of the component functions is defined, so we need to determine the domain of each of these functions.

The domain of $x(t)$ is $[-3, 3]$; the domain of $y(t)$ is all real numbers, $\mathbb{R}$; and the domain of $z(t)$ is $(2, 4)$. The only place where each of these functions is defined is the interval $(2, 3]$. Thus the domain of $\vec{r}(t)$ is $(2, 3]$.

Curves in 3-space

Notice that every vector-valued function gives us a curve in 3-space. We imagine that for each vector $\vec{r}(t)$ we get when we plug in a value for t, the tail of that vector is at the origin. Then the tip of vector is some point in space. Imagine that a tiny pen or paint brush is attached to the tip of the vector. As the value of t changes, the tip of this vector changes, and the pen traces out a curve in space.

2.2 Limits, continuity, and tangent vectors

In the last section we defined vector-valued functions, and mentioned their relationship to parametric curves in 3-space. In this section we'll discuss how to take limits of these vector-valued functions, and what it means to say that such a function is continuous.

Limits

Recall that in the last section we said a vector-valued function was a function of the form

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

where each of $x(t), y(t)$, $z(t)$ was a scalar-valued functions. We're going to want to do calculus with these vector-valued functions, in the sense that we'll want to integrate and differentiate these functions. Since derivatives and integrals are defined in terms of limits, we first have to make sense of what the limit of a vector-valued function is.

Intuitively, $\lim_{t\to t_0} \vec{r}(t) = \vec{v}$ means that as t gets "really close" to t_0, the value of $\vec{r}(t)$ gets "really close" to $\vec{v}$. This is exactly what a limit of a function $f : \mathbb{R} \to \mathbb{R}$ means, and we're just extending that idea to vector-valued functions. The problem is that "really close" is a pretty vague term, so we'd like to make it precise.

Basically, we want $\vec{r}(t)$ to get arbitrarily close to $\vec{v}$. So if you give me an $\varepsilon > 0$, I want to be able to say that $\vec{r}(t)$ and $\vec{v}$ are less than ε-distance away from one another, provided t and t_0 are close enough. Now, how should we measure how far apart two vectors are? We'll define the distance between two vectors as the magnitude of their difference. I.e., the distance between $\vec{u}$ and $\vec{v}$ is defined to be $\|\vec{u} - \vec{v}\|$. (Notice this is the same as the distance between the tips of the arrows, if the tails are the origin.)

Caution: The next few points are a little bit technical and might seem strange if you're not used to these sorts of ideas. They're here if you're curious or ambitious enough to try to understand them, but you can skip over this stuff if you want and just go to the theorem below.

Now, to formally define the limit: We'll write $\lim_{t\to t_0} \vec{r}(t) = \vec{v}$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|\vec{r}(t) - \vec{v}\| < \varepsilon$ whenever $|t - t_0| < \delta$. If $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ and $\vec{v} = \langle x_0, y_0, z_0\rangle$, this means

$$\sqrt{(x(t) - x_0)^2 + (y(t) - y_0)^2 + (z(t) - z_0)^2} < \varepsilon.$$

Since $\sqrt{x}$ is an increasing function,

$$\begin{aligned}&\sqrt{(x(t)-x_0)^2+(y(t)-y_0)^2+(z(t)-z_0)^2}\\ <&\sqrt{(x(t)-x_0)^2}+\sqrt{(y(t)-y_0)^2}+\sqrt{(z(t)-z_0)^2}.\end{aligned}$$

We can rewrite this as

$$\begin{aligned}&\sqrt{(x(t)-x_0)^2+(y(t)-y_0)^2+(z(t)-z_0)^2}\\ <&|x(t)-x_0|+|y(t)-y_0|+|z(t)-z_0|\end{aligned}$$

So, as long as each of the following inequalities holds:

$$\begin{aligned}|x(t)-x_0| &< {}^{\varepsilon}/_{3}\\ |y(t)-y_0| &< {}^{\varepsilon}/_{3}\\ |z(t)-z_0| &< {}^{\varepsilon}/_{3}\end{aligned}$$

it's guaranteed that $\|\vec{r}(t)-\vec{v}\| < \varepsilon$. This means that if $x(t) \to x_0$, $y(t) \to y_0$, and $z(t) \to z_0$ as $t \to t_0$, then $\vec{r}(t) \to \vec{v}$.

In fact, you can go the other way to show that if the limit of $\vec{r}(t)$ exists, then the limits of the component functions must exist as well. If the limit exists, then for every $\varepsilon > 0$ there exists a $\delta > 0$ so that

$$\sqrt{(x(t)-x_0)^2+(y(t)-y_0)^2+(z(t)-z_0)^2} < \varepsilon$$

whenever $|t-t_0| < \delta$. This works for *every* $\varepsilon > 0$, so in particular it works for ε^2: We can guarantee that the expression under the radical above is less than ε^2 if we choose δ small enough. This implies that each of

$$\begin{aligned}(x(t)-x_0) &< \varepsilon^2\\ (y(t)-y_0) &< \varepsilon^2\\ (z(t)-z_0) &< \varepsilon^2.\end{aligned}$$

(If any of these quantities was greater than or equal to ε^2 then the expression under the radical would be greater than or equal to ε^2 as well.) Taking roots of both sides gives us the result.

The arguments above basically prove the following the theorem:

Theorem 2.1.
The limit of the vector-valued function $\lim_{t\to t_0} \vec{r}(t)$ exists and equals $\vec{v}$, if and only if the limits of the component functions exist and equal the components of $\vec{v}$.

Symbolically, the above theorem just states:

$$\lim_{t\to t_0} \vec{r}(t) = \left\langle \lim_{t\to t_0} x(t), \lim_{t\to t_0} y(t), \lim_{t\to t_0} z(t) \right\rangle.$$

So taking limits of vector-valued functions is about as easy as you could hope for.

Example 2.3.
Calculate the following limit:

$$\lim_{t\to 0} \left\langle \frac{\sin(t)}{t}, t^2+2, \frac{t^2-t}{t} \right\rangle$$

We simply take the limits of the component functions to get

$$\begin{aligned}&\lim_{t\to 0} \left\langle \frac{\sin(t)}{t}, t^2+2, \frac{t^2-t}{t} \right\rangle \\ &= \left\langle \lim_{t\to 0} \frac{\sin(t)}{t}, \lim_{t\to 0} \left(t^2+2\right), \lim_{t\to 0} \frac{t^2-t}{t} \right\rangle \\ &= \langle 1, 2, -1 \rangle\end{aligned}$$

Notice in the example above that the function is not defined at $t = 0$, but the limit still exists. (Compare this to the idea of a removable discontinuity from your first semester of calculus.)

If *any* of the component functions doesn't have a limit, then the vector-valued function doesn't have a limit either. This includes the case when one of the component functions has a limit of $\pm\infty$.

Example 2.4.
Does the following limit exist?

$$\lim_{t\to 2} \left\langle \frac{t^2-4}{t-2}, \frac{\cos(t-2)}{t}, \frac{1}{t-2} \right\rangle$$

What we need to see is if the limit of each of the component functions exists or not:

$$\begin{aligned}\lim_{t\to 2}\frac{t^2-4}{t-2} &= \lim_{t\to 2}\frac{(t+2)(t-2)}{t-2}\\ &= \lim_{t\to 2}(t+2)\\ &= 4\end{aligned}$$

$$\begin{aligned}\lim_{t\to 2}\frac{\cos(t-2)}{t} &= \frac{\cos 0}{2}\\ &= \frac{1}{2}\end{aligned}$$

$$\lim_{t\to 2}\frac{1}{t-2} \qquad \text{DNE}$$

The last limit does not exist because the left- and right-hand limits don't equal:

$$\lim_{t\to 2^+}\frac{1}{t-2} = \infty$$
$$\lim_{t\to 2^-}\frac{1}{t-2} = -\infty$$

Continuity

We say that the vector-valued function $\vec{r}(t)$ is ***continuous at*** t_0 if the following conditions are met:

(i) $\vec{r}(t_0)$ is defined

(ii) The limit $\lim_{t\to t_0}\vec{r}(t)$ exists

(iii) The two values above are the same: $\lim_{t\to t_0}\vec{r}(t) = \vec{r}(t_0)$.

This of course just amounts to saying that each of the component functions is continuous at t_0.

If a $\vec{r}(t)$ is continuous at every point in its domain, then we just say that $\vec{r}(t)$ is ***continuous***. This means that the component functions are continuous in their shared domain (in the domain of $\vec{r}(t)$).

Theorem 2.2.
The function $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ is continuous at t_0 if and only if each of the component functions is continuous at t_0. That is:

(*i*) $x(t_0)$, $y(t_0)$, *and* $z(t_0)$ *all exist*

(*ii*) *Each of the limits,* $\lim_{t\to t_0} x(t)$, $\lim_{t\to t_0} y(t)$, *and* $\lim_{t\to t_0} z(t)$ *exists.*

(*iii*) *The following equations hold:*

$$\begin{aligned} \lim_{t\to t_0} x(t) &= x(t_0) \\ \lim_{t\to t_0} y(t) &= y(t_0) \\ \lim_{t\to t_0} z(t) &= z(t_0) \end{aligned}$$

Example 2.5.
Is the function $\vec{r}(t) = \left\langle t^2, t+2, \frac{1}{t^2}\right\rangle$ continuous at the point $t = 0$?

No, this function is not continuous at $t = 0$ because of the last component function: the function isn't defined at $t = 0$.

Example 2.6.
Is the function $\vec{r}(t) = \left\langle t^2, t+2, \frac{1}{t^2}\right\rangle$ continuous?

Yes, this function is continuous. The domain of this function is $(-\infty, 0) \cup (0, \infty)$, and the function is continuous at every point in its domain.

The basic idea behind continuity (not just for vector-valued functions, but continuous functions in general) is that small changes in inputs produce small changes in outputs. This is really the content of all this technical ε-δ stuff above: if you change the input just a tiny little bit, the output can't change too drastically. For this reason continuous functions are really nice to work with, because they can't do anything too crazy. (Additionally, most of the real, physical quantities people care about are modelled by continuous functions. If you imagine that the temperature in a room is a function of the position in the room, the temperature can't change too quickly if you change the position just a little bit.)

Tangent vectors and tangent lines

In your first semester calculus class you learned that the derivative of a function $f(x)$ at a point x_0 gives the slope of the tangent line to $y = f(x)$ at $(x_0, f(x_0))$. We would like to do something similar for parametric curves and vector-valued functions. The slope of this line is difficult to compute, so we approximate it with something similar by simpler: the slope of a tangent line between two nearby points. In particular, we consider two points $(x_0, f(x_0))$ and $(x, f(x))$ (for x very near to x_0) on our graph and compute the slope of the line between these points to simply be

$$\frac{f(x) - f(x_0)}{x - x_0}.$$

This is only an approximation to the quantity we care about, though, and we get better approximations by moving x closer to x_0. Thus we are lead to the idea of taking the limit as x approaches x_0,

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

and this limit (if it exists) is the slope of the line tangent to our graph at that point.

We want to mimic this idea for parametric curves, however in three (and higher) dimensions the direction of a line isn't given simply by a single number, but rather is a vector. Thus the "slope of the tangent line" from first semester calculus is replaced by "a vector pointing along the tangent line" now, which we simply call the "tangent vector."

To be more precise, say we want to find the direction (tangent vector) for the line tangent to the curve parametrized by $\vec{r}(t)$ at the point $\vec{r}(t_0)$. This seems difficult to do, so we approximate with something simpler. We pick some nearby value t close to t_0 and consider the two points $\vec{r}(t)$

and $\vec{r}(t_0)$ on our curve. We can very easily get compute the displacement vector between these two points as

$$\vec{r}(t) - \vec{r}(t_0).$$

At first you may be tempted to go ahead and take the limit as t goes to t_0 of this quantity, but let's notice that if our function is continuous, then this limit will always go to zero. Let's fix this by dividing this expression by $t - t_0$. This will preserve the direction of our vector (since we're just doing scalar multiplication), but it the advantage of preventing our quantity from shrinking too quickly for the limit to be interesting (since we are dividing by very small values, which gives us bigger vectors, when t is very close to t_0). That is, we want to consider the limit

$$\lim_{t \to t_0} \frac{\vec{r}(t) - \vec{r}(t_0)}{t - t_0}.$$

This limit, if it exists, we will call the ***derivative*** of $\vec{r}(t)$ at t_0 and denote it by $\vec{r}\,'(t_0)$.

Let's go ahead and notice that if $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ then we can rewrite this quantity as follows:

$$\begin{aligned}
\vec{r}\,'(t_0) &= \lim_{t \to t_0} \frac{\vec{r}(t) - \vec{r}(t_0)}{t - t_0} \\
&= \lim_{t \to t_0} \frac{\langle x(t), y(t), z(t) \rangle - \langle x(t_0), y(t_0), z(t_0) \rangle}{t - t_0} \\
&= \lim_{t \to t_0} \frac{\langle x(t) - x(t_0), y(t) - y(t_0), z(t) - z(t_0) \rangle}{t - t_0} \\
&= \lim_{t \to t_0} \left\langle \frac{x(t) - x(t_0)}{t - t_0}, \frac{y(t) - y(t_0)}{t - t_0}, \frac{z(t) - z(t_0)}{t - t_0} \right\rangle \\
&= \left\langle \lim_{t \to t_0} \frac{x(t) - x(t_0)}{t - t_0}, \lim_{t \to t_0} \frac{y(t) - y(t_0)}{t - t_0}, \lim_{t \to t_0} \frac{z(t) - z(t_0)}{t - t_0} \right\rangle \\
&= \langle x'(t_0), y'(t_0), z'(t_0) \rangle .
\end{aligned}$$

That is, we can (very conveniently) compute the derivative of a vector-valued function by simply taking the derivatives of each of its components.

Example 2.7.

Find a vector tangent to the curve parametrized by

$$\vec{r}(t) = \left\langle 2\sin^2(t), \cos(t)\sin(t), \ln(t^2+1)\right\rangle$$

when $t = \pi$.

First we calculate the derivative vector,

$$\vec{r}\,'(t) = \left\langle 4\sin^2(t)\cos(t), \cos^2(t) - \sin^2(t), \frac{2t}{t^2+1}\right\rangle.$$

When $t = \pi$ this vector is

$$\vec{r}\,'(\pi) = \left\langle 0, 1, \frac{2\pi}{\pi^2+1}\right\rangle,$$

and by the above proposition, this vector is tangent to the curve when $t = \pi$.

Example 2.8.
Find a vector tangent to the curve parametrized by

$$\vec{r}(t) = \left\langle 3t^2 - t, 4t, e^{-\sqrt{t}}\right\rangle$$

at the point $(44, 16, e^{-2})$.

Notice that in this example we weren't given a value of t, but were instead given a point on the curve. For us to find a vector tangent to this curve, we first need to figure out what value of t corresponds to the point $(44, 16, e^{-2})$. To do this we solve the system

$$\begin{aligned} 3t^2 - t &= 44 \\ 4t &= 16 \\ e^{-\sqrt{t}} &= e^{-2}. \end{aligned}$$

These equations are satisfied only when $t = 4$. (In general, there may be several values of t that satisfy the equations $(x(t), y(t), z(t)) = (x_0, y_0, z_0)$. This corresponds to the curve passing through the same point several times. Each of these different values of t will give us different tangent vectors, and the one we choose to calculate may matter in certain problems!)

Now we calculate the derivative,

$$\vec{r}'(t) = \left\langle 6t - 1, 4, (2t)^{-1/2} e^{-\sqrt{t}} \right\rangle .$$

Plugging in $t = 4$ we have that the vector

$$\vec{r}'(4) = \left\langle 23, 4, \frac{e^{-2}}{\sqrt{8}} \right\rangle$$

is tangent to the curve at the point $(44, 16, e^{-2})$.

One subtle point to notice here is that the same curve can be parametrized in several different ways. For example, the helix we saw before (in Lecture 7) is traced out by each of the three vector-valued functions below:

$$\begin{aligned} \vec{r_1}(t) &= \langle \cos(t), \sin(t), t \rangle \\ \vec{r_2}(t) &= \langle \cos(2t), \sin(2t), 2t \rangle \\ \vec{r_3}(t) &= \langle \cos(-t), \sin(-t), -t \rangle . \end{aligned}$$

Though each of these vector-valued functions gives us the same curve, the points of $\vec{r_i}(t)$ move around the curve in different ways. In particular, $\vec{r_2}(t)$ moves through the curve twice as fast as $\vec{r_1}(t)$ does, and $\vec{r_3}(t)$ traverses the curve in reverse.

Differentiating each of the $\vec{r_i}(t)$ gives us different vectors that are tangent to the curve. For example, each of the three vectors below is tangent to the curve at the point $(1, 0, 0)$.

$$\begin{aligned} \vec{r_1}'(0) &= \langle 0, 1, 1 \rangle \\ \vec{r_2}'(0) &= \langle 0, 2, 2 \rangle \\ \vec{r_3}'(0) &= \langle 0, -1, -1 \rangle . \end{aligned}$$

(Note that for this particular point on the curve, the value $t = 0$ gives $(1, 0, 0)$ for each fo the $r_i(t)$. This is a special situation, and doesn't happen in general! For example, the point $(0, 1, 2\pi)$ is on the curve, but corresponds to $\vec{r_1}(2\pi)$, $\vec{r_2}(\pi)$ and $\vec{r_3}(-2\pi)$.)

For some problems and applications the choice of the parametrization (and thus tangent vector) matters, but for other problems it doesn't. For example, if $\vec{r}(t)$ is supposed to be position of a particle as it moves around in space, and so $\vec{r}\,'(t)$ is the velocity, changing the parametrization will (typically) change the velocity, and so the parametrization matters in such a situation. If however we were only interested in the geometry of a curve, and not really care about how quickly you move along the curve, then we wouldn't care about the parametrization or which tangent vector we used.

For this reason we'll often work with ***unit tangent vectors***. These are tangent vectors which are also unit vectors – that is, these vectors have length 1. In situations where the parametrization (and so choice of tangent vector) doesn't matter, it will be convenient to work with unit vectors. Given a parametrization $\vec{r}(t)$ of the curve, we can always calculate the unit tangent vector by making $\vec{r}\,'(t)$ have unit length. That is, for a curve parametrized by $\vec{r}(t)$, the unit tangent vector of the curve is

$$\vec{T}(t) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|}.$$

Example 2.9.
Find a unit tangent vector for the helix at the point $(\sqrt{2}/2, \sqrt{2}2/, \pi/4)$.

We know that the helix may be parametrized by

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

In this case the given point corresponds to $t = \pi/4$. Thus

$$\begin{aligned}
\vec{r}(t) &= \langle \cos(t), \sin(t), t \rangle \\
\implies \vec{r}\,'(t) &= \langle -\sin(t), \cos(t), 1 \rangle \\
\implies \vec{r}\,'(\pi/4) &= \langle -\sqrt{2}/2, \sqrt{2}/2, 1 \rangle \\
\implies \|\vec{r}\,'(\pi/4)\| &= \sqrt{1/2 + 1/2 + 1} = \sqrt{2} \\
\implies \vec{T}(\pi/4) &= \langle 1/2, 1/2, 1/\sqrt{2} \rangle .
\end{aligned}$$

Notice that if we had used another parametrization, such as $\vec{r}(t) = \langle \cos(2t), \sin(2t), 2t \rangle$, and so we'd use $t = \pi/8$, we'd calculate the same unit tangent vector.

$$\begin{aligned} \vec{r}(t) &= \langle \cos(2t), \sin(2t), 2t \rangle \\ \implies \vec{r}\,'(t) &= \langle -2\sin(2t), 2\cos(2t), 2 \rangle \\ \implies \vec{r}\,'\left(\pi/8\right) &= \left\langle -\sqrt{2}, \sqrt{2}, 2 \right\rangle \\ \implies \left\|\vec{r}\,'\left(\pi/8\right)\right\| &= \sqrt{2+2+4} = \sqrt{8} \\ \implies \vec{T}\left(\pi/4\right) &= \langle 1/2, 1/2, 1/\sqrt{2} \rangle . \end{aligned}$$

If we had used the third parametrization mentioned above, we'd have the negative of this vector.

This is the nice thing about using unit tangent vectors: we can (almost) forget the choice of parametrization.

Now that we know how to get a vector tangent to a curve, we can find the equation of a line tangent to the curve.

Example 2.10.
Find the equation of a line which is tangent to the curve $\vec{r}(t) = \langle t, t^2, t^3 \rangle$ at the point $(-2, 4, -8)$.

First notice that the point $(-2, 4, -8)$ corresponds to $t = -2$ with the above parametrization. The tangent vector is then

$$\begin{aligned} \vec{r}(t) &= \left\langle t, t^2, t^3 \right\rangle \\ \implies \vec{r}\,'(t) &= \left\langle 1, 2t, 3t^2 \right\rangle \\ \implies \vec{r}\,'(-2) &= \langle 1, -4, 12 \rangle \end{aligned}$$

We have a point on the line, $(-2, 4, -8)$, and a vector parallel to the line, $\langle 1, -4, 12 \rangle$. Hence the symmetric equations of the tangent line are

$$x + 2 = \frac{4 - y}{4} = \frac{x + 8}{12}.$$

2.3 Arclength

We know how to calculate the length of a line segment in 3-space by using the Pythagorean theorem. We can thus estimate the length of a curve by approximating the curve with line segments. Taking the limit as the line segments become arbitrarily short we obtain an integral which tells us the arclength of the curve.

Basic idea behind arclength

Recall that the line segment connecting the points (x_0, y_0, z_0) and (x_1, y_1, z_1) has length

$$\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2} = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}.$$

Suppose instead of measuring the length of a straight line segment, we instead wanted to measure the length of a curve. How could we use the tools we already know (lengths of line segments) to calculate the length of a curve?

What we'll do is approximate the curve with line segments, calculate the lengths of those line segments, and add up those lengths. This will give us an estimate of the length of the curve.

Example 2.11.
Suppose our curve is the helix parametrized by

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

where $0 \leq t \leq 2\pi$. To approximate the length of the curve, let's pick five points on the curve, and calculate the lengths of the line segments connecting those points. In order to pick points on the curve, we can just pick points in the interval $[0, 2\pi]$. Say we pick the points

$$0, {}^{\pi}/{}_{2}, \pi, {}^{3\pi}/{}_{2}, 2\pi.$$

Plugging these values of t into $\vec{r}(t)$ gives us the following points on the curve:

$$\begin{aligned} P_0 = \vec{r}(0) &= (1,0,0) \\ P_1 = \vec{r}\,(\pi/2) &= (0,1,\pi/2) \\ P_2 = \vec{r}\,(\pi) &= (-1,0,\pi) \\ P_3 = \vec{r}\,(3\pi/2) &= (0,-1,P_4 = 3\pi/2) \\ \vec{r}(2\pi) &= (1,0,2\pi) \end{aligned}$$

Now we calculate the distances between these points (aka the length of the line segment connecting the points):

$$\begin{aligned} |P_0P_1| &= \sqrt{1+1+\pi^2/4} = \frac{\sqrt{8+\pi^2}}{2\sqrt{2}} \\ &= |P_1P_2| \\ &= |P_2P_3| \\ &= |P_3P_4| \end{aligned}$$

(In this particular example it worked out that these line segments all had the same length). Summing these lengths, we estimate that the length of the helix is $\sqrt{16+2\pi^2}$.

If we wanted to make our estimate better, we'd approximate the curve with lots and lots of small line segments. Say our curve was parametrized by $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ where $a \le t \le b$. If we cut the interval $[a,b]$ into n different segments, say

$$\mathcal{P} = \{a = t_0, t_1, t_2, ..., t_{n-1}, t_n = b\}$$

where $t_{i-1} \le t_i$, then the length of the line segment connecting $\vec{r}(t_{i-1})$ and $\vec{r}(t_i)$ would be

$$\sqrt{(x(t_i)-x(t_{i-1}))^2 + (y(t_i)-y(t_{i-1}))^2 + (z(t_i)-z(t_{i-1}))^2} = \sqrt{\Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2}.$$

Our estimate to the length of the curve would then be

$$\sum_{i=1}^{\#\mathcal{P}} \sqrt{\Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2}.$$

Of course what we'd like to do now is take a limit as the points on the curve get arbitrarily close together: $|\mathcal{P}| \to 0$. Before we do this, let's notice that we can rewrite the summand above as

$$\begin{aligned}\sqrt{\Delta x_i^2 + \Delta y_i^2 + \Delta z_i^2} &= \sqrt{\left(\frac{\Delta x_i \Delta t_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i \Delta t_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta z_i \Delta t_i}{\Delta t_i}\right)^2} \\ &= \sqrt{\left(\frac{\Delta x_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta z_i}{\Delta t_i}\right)^2}\Delta t_i\end{aligned}$$

Now, if we assume that each of $x(t), y(t), z(t)$ is differentiable, then the mean value theorem tells us that there must exist some t_i^* in the interval $[t_{i-1}, t_i]$ such that

$$x'(t_i^*) = \frac{\Delta x_i}{\Delta t_i},$$

and similarly for $\frac{\Delta y_i}{\Delta t_i}$ and $\frac{\Delta z_i}{\Delta t_i}$. Hence we can write our estimate for the arclength of the curve as

$$\sum_{i=1}^{\#\mathcal{P}} \sqrt{x'(t_i^*)^2 + y'(t_i^*)^2 + z'(t_i^*)^2}\Delta t_i.$$

Taking the limit as $|\mathcal{P}| \to 0$, we have that the arclength of the curve is

$$\int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt$$

Example 2.12.
The arclength of the segment of the helix parametrized by $\vec{r}(t) = \langle \cos(t), \sin(t), t\rangle$ with $0 \le t \le 2\pi$ is

$$\begin{aligned}\int_0^{2\pi} \sqrt{\sin^2(t) + \cos^2(t) + 1}\, dt &= \int_0^{2\pi} \sqrt{1+1}\, dt \\ &= \sqrt{2}\int_0^{2\pi} dt \\ &= 2\pi\sqrt{2}.\end{aligned}$$

Example 2.13.
Notice that the arclength of a curve is independent of the parametrization used. For example, if we used the parametrization $\vec{r}(t) = \langle \cos(2t), \sin(2t), 2t\rangle$ for $0 \leq t \leq \pi$, we'd calculate the arclength to be

$$\begin{aligned}\int_0^\pi \sqrt{4\sin^2(2t) + 4\cos^2(2t) + 4}\,dt &= \int_0^\pi \sqrt{4+4}\,dt \\ &= \sqrt{8}\int_0^\pi dt \\ &= 2\sqrt{2}\int_0^\pi dt \\ &= 2\pi\sqrt{2}.\end{aligned}$$

Arclength parametrizations

Notice that the integrand,

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

that appears in calculating arclength could also be written as $\|\vec{r}\,'(t)\|$:

$$\int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}dt = \int_a^b \|\vec{r}\,'(t)\|dt.$$

In the special case that each tangent vector, $\vec{r}\,'(t)$, is a unit vector for every t, $\|\vec{r}\,'(t)\| = 1$, we say that $\vec{r}(t)$ is an *arclength parametrization* of the curve. The nice thing about an arclength parametrization is that it's very easy to calculate arclength:

$$\int_a^b \|\vec{r}\,'(t)\|dt = \int_a^b dt = b - a.$$

That is, if we walk along a curve using an arclength parametrization for t units of time, then we've walked t units of distance.

If we introduce an *arclength function* that measures the distance we've walked along a curve after walking for t units of time,

$$s(t) = \int_0^t \|\vec{r}\,'(u)\|du,$$

then an arclength parametrization is the special case that $s(t) = t$.

Example 2.14.
Find an arclength parametrization for the helix parametrized by

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

where $0 \leq t \leq 2\pi$.

When we calculate the arclength function we get

$$s(t) = \int_0^t \sqrt{\sin^2(u) + \cos^2(u) + 1}\, du = \sqrt{2}t$$

We want this to simply equal t, so let's solve the expression above for t, obtaining $t = {}^{s}/_{\sqrt{2}}$. The claim now is that $\vec{r}\,({}^{s}/_{\sqrt{2}})$. Let's double-check that this is indeed an arclength parametrization:

$$\begin{aligned}
\int_0^t \left\| \vec{r}\,' \left(\frac{u}{\sqrt{2}} \right) \right\| du &= \int_0^t \sqrt{\frac{1}{2}\sin^2\left(\frac{u}{\sqrt{2}}\right) + \frac{1}{2}\cos^2\left(\frac{u}{\sqrt{2}}\right) + \frac{1}{2}}\, du \\
&= \int_0^t \sqrt{\frac{1}{2} + \frac{1}{2}}\, du \\
&= t
\end{aligned}$$

So this is indeed an arclength parametrization.

2.4 Curvature

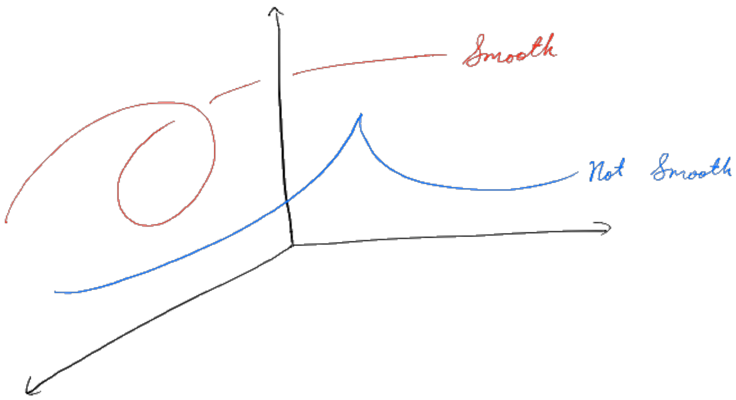
We now turn our attention to associating a number to each point on a curve that describes how "curvy" the curve is. This study of "curvature" is the beginning of a long and interesting discussion if taken to the extremes. For example, understanding the curvature of a curves can later be used to define a notion of curvature of surfaces, and extending the notion of curvature to higher-dimensional objects essentially leads to Einstein's general theory of relativity. (It is outside the scope of this class, but it's interesting to note that gravity is really just the curvature of the 4-dimensional spacetime we live in, and a serious study of this in physics basically boils down to studying four-dimensional geometry.)

Before we can define the "curvature" of a curve, we need some preliminaries. First, in order of our definition of curvature to make sense we need to be sure that our curve "curves" in a nice way without making any sharp kinks or corners. This is summed up by saying the curve is "smooth." To be precise, we will say that a parametrization $\vec{r}(t)$ of a curve C is ***smooth*** if the following thee conditions are satisfied:

(1) the tangent vector $\vec{r}\,'(t)$ exists for all t in the domain of $\vec{r}(t)$;

(2) the vector-valued function $\vec{r}\,'(t)$ is continuous; and

(3) the tangent vector $\vec{r}\,'(t)$ *is not* equal to the zero vector at any t.



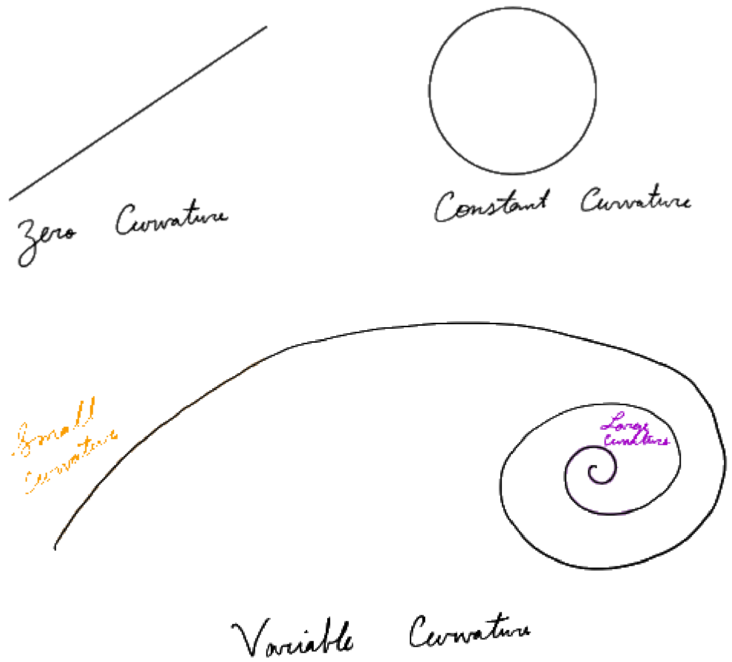
We will say that a curve C is ***smooth*** if it admits a smooth parametrization. Notice that a curve can be smooth even if a given parametrization is not smooth. For example, the following parametrization of the unit circle in the plane is not a smooth parametrization:

$$\vec{r}(t) = \begin{cases} \langle \cos(t), \sin(t) \rangle & \text{if } 0 \le t \le \pi \\ \langle -1, 0 \rangle & \text{if } \pi \le t \le 2\pi \\ \langle \cos(t-\pi), \sin(t-\pi) \rangle & \text{if } 2\pi \le t \le 3\pi. \end{cases}$$

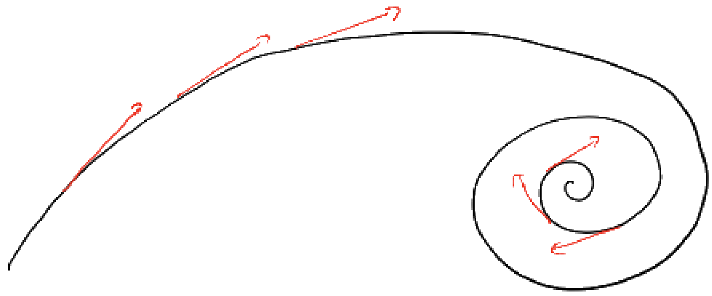
This parametrization walks around the top half of the curve, then just completely stops at the point $(-1, 0)$ for a long time before continuing. While stopped at $(-1, 0)$ (i.e., for $\pi < t < 2\pi$), the derivative will be zero and so the parametrization is not smooth. Of course, the circle itself is smooth because it does have other parametrizations which satisfy our definition of smooth above. Namely the "standard" parametrization $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle$ for $0 \le t \le 2\pi$ is a smooth parametrization as can easily be checked.

Before we give the proper definition of curvature, let's go ahead and mention a few properites we might expect a measure of "curviness" to have, just to help us build some intuition.

- A line should have zero curvature.
- A circle should have constant curvature (as the circle is "equally curvy" at every point).
- For curves which are "almost straight lines," the curvature should be small, but for curves which make quick turns, the curvature should be large.



To actually define curvature, we'll look at unit tangent vectors of the curve and see how quickly those tangent vectors change as we walk along the curve.



Since we're interested in how quickly these tangent vectors change, we want to define the curvature as the derivative of the unit tangent vectors. These vectors change as we walk along the curve, so we differentiate with respect to arclength. That is, we will define the ***curvature*** of the curve to be

$$\kappa = \left\| \frac{d\vec{T}}{ds} \right\|$$

. Notice here that we are differentiating with respect to s which is the arclength. That is, our definition of curvature should be independent of the parametrization we choose.

The definition of curvature we've stated makes intuitive sense (measure how the tangent vectors change as you walk along the curve), but it's actually a little difficult to calculate because of the fact we want to differentiate with respect to arclength. So, we'd like to change our definition a little bit so that it works for any parametrization $\vec{r}(t)$ of our curve.

To do this, let's recall that the arclength function $s(t)$ associated to a curve with parametrization $\vec{r}(t)$ with domain $[a, b]$ is given by

$$s(t) = \int_a^t \|\vec{r}\,'(u)\| du.$$

Differentiating this tells us

$$\frac{ds}{dt} = \|\vec{r}\,'(t)\|.$$

Now, notice that the unit tangent vector $\vec{T}$ appearing in our definition of curvature κ above is a function of s, but s is itself a function of t. That is, we really have a composition of functions, $\vec{T}(s(t))$. We can differentiate this composition with respect to t by applying the chain rule:

$$\begin{aligned}\frac{d\vec{T}}{dt} &= \frac{d}{dt}\vec{T}(s(t)) \\ &= \vec{T}\,'(s(t)) \cdot s'(t) \\ &= \frac{d\vec{T}}{dt} \cdot \frac{ds}{dt} \\ &= \frac{d\vec{T}}{dt} \cdot \|\frac{d\vec{r}}{dt}\|.\end{aligned}$$

Now let's notice that $\frac{d\vec{T}}{ds}$ is a vector and $\frac{d\vec{r}}{dt}$ is a scalar, so we're just doing scalar multiplication above. Taking the length of this vector we can factor out this scalar multiple and obtain

$$\|\frac{d\vec{T}}{dt}\| = \|\frac{d\vec{T}}{ds}\| \cdot \|\frac{d\vec{r}}{dt}\|.$$

Note, though, $\|\frac{d\vec{T}}{ds}\|$ is exactly our curvature κ and so $\|\frac{d\vec{T}}{dt}\| = \kappa \cdot \|\frac{d\vec{r}}{dt}\|$, which we can solve for the curvature to obtain

$$\kappa(t) = \frac{\|\vec{T}\,'(t)\|}{\|\vec{r}\,'(t)\|}.$$

Just to summarize, if we are given a parametrization $\vec{r}(t)$ of a curve C, then the curvature at point $\vec{r}(t)$ on the curve is given by

$$\kappa(t) = \frac{\|\vec{T}\,'(t)\|}{\|\vec{r}\,'(t)\|}$$

where

$$\vec{T}(t) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|}.$$

Example 2.15.
Find the curvature of the line

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

This is parametrized by

$$\vec{r}(t) = \langle x_0 + at, y_0 + bt, z_0 + ct \rangle,$$

and so we compute

$$\begin{aligned}
&\vec{r}\,'(t) = \langle a, b, c\rangle \\
\implies &\vec{T}(t) = \frac{\langle a, b, c\rangle}{\|\langle a, b, c\rangle\|} = \left\langle \frac{a}{\sqrt{a^2+b^2+c^2}}, \frac{b}{\sqrt{a^2+b^2+c^2}}, \frac{c}{\sqrt{a^2+b^2+c^2}} \right\rangle \\
\implies &\vec{T}\,'(t) = \langle 0, 0, 0\rangle \\
\implies &\kappa(t) = \frac{\|\vec{T}\,'(t)\|}{\|\vec{r}\,'(t)\|} = \frac{0}{\sqrt{a^2+b^2+c^2}} = 0,
\end{aligned}$$

and so the line has zero curvature.

Example 2.16.
Find the curvature of a circle of radius $\rho > 0$ in the plane.
We may parametrize the circle by

$$\vec{r}(t) = \langle \rho\cos(t), \rho\sin(t), 0\rangle$$

and so

$$\vec{r}\,'(t) = \langle -\rho\sin(t), \rho\cos(t), 0\rangle\,.$$

Now let's notice the magnitude $\|\vec{r}\,'(t)\|$ is easily computed by

$$\begin{aligned}
\|\vec{r}\,'(t)\| &= \sqrt{\rho^2\sin^2(t) + \rho^2\cos^2(t)} \\
&= \sqrt{\rho^2(\sin^2(t) + \cos^2(t))} \\
&= \sqrt{\rho^2} \\
&= |\rho| = \rho.
\end{aligned}$$

Now we can easily compute the unit tangent vectors as

$$\begin{aligned}
\vec{T}(t) &= \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|} \\
&= \frac{\langle -\rho\sin(t), \rho\cos(t), 0\rangle}{\rho} \\
&= \langle -\sin(t), \cos(t), 0\rangle\,,
\end{aligned}$$

which have derivatives

$$\vec{T}\,'(t) = \langle -\cos(t), \sin(t), 0\rangle \,.$$

Of course, then easily see

$$\|\vec{T}\,'(t)\| = \sqrt{\cos^2(t) + \sin^2(t)} = 1,$$

and using our formula for curvature above we have

$$\kappa(t) = \frac{\|\vec{T}\,'(t)\|}{\|\vec{r}\,'(t)\|} = \frac{1}{\rho}.$$

That is, a circle of radius ρ has curvature $\frac{1}{\rho}$.

Remark.
It's interesting to note that since curvature is one over the radius, the larger the radius, the smaller the curvature. This has the following interpretation: if you were to drive your car in a perfect circle, you'll drive in larger circles by rotating the wheel less (your rotation of the wheel is essentially the curvature of the curve you're driving along). For smaller, tighter circles you need to turn the wheel more. If you want to burn donuts in the parking lot, for example, you need to turn the wheel as much as you can.

Also, to a small insect living on your circle, the bigger the circle is, the flatter the circle appears (the curvature is closer to zero). This is something you essentially experience in your own life: the radius of the Earth is big, so its curvature is small, and so the Earth "feels" flat over small enough regions. (The Earth is a sphere and not a circle, but we can measure the curvature of surfaces and a similar type of statement will hold there as well.)

Since $\vec{T}(t) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|}$, it shouldn't be too surprising that we can rewrite the formula for curvature only in terms of $\vec{r}(t)$ and its derivatves. Notice, in particular, that our formula of $\kappa(t)$ involves the derivative of $\vec{T}\,'(t)$

which itself is defined in terms of the derivative of $\vec{r}\,'(t)$. That is, we should expect that the curvature involves a second derivative of $\vec{r}(t)$. This is basically what the following theorem tells us.

Theorem 2.3.
If C is a curve with smooth parametrization $\vec{r}(t)$, then its curvature function can be computed as

$$\kappa(t) = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3}$$

Example 2.17.
Find the curvature $\kappa(t)$ of the curve parametrized by $\vec{r}(t) = \langle t, t^2, e^t \rangle$.

First we need to compute all of the quantities required to use Theorem 2.3.

$$\begin{aligned}
\vec{r}\,'(t) &= \left\langle 1, 2t, e^t \right\rangle \\
\vec{r}\,''(t) &= \left\langle 0, 2, e^t \right\rangle \\
\|\vec{r}\,'(t)\| &= \sqrt{1 + 4t^2 + e^{2t}} \\
\vec{r}\,'(t) \times \vec{r}\,''(t) &= \det \begin{pmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 1 & 2t & e^t \\ 0 & 2 & e^t \end{pmatrix} \\
&= \left\langle 2e^t(t-1), -e^t, 2 \right\rangle \\
\|\vec{r}\,'(t) \times \vec{r}\,''(t)\| &= \sqrt{4e^{2t}(t^2 - 2t + 1) + e^{2t} + 4}.
\end{aligned}$$

Thus the curvature is given by

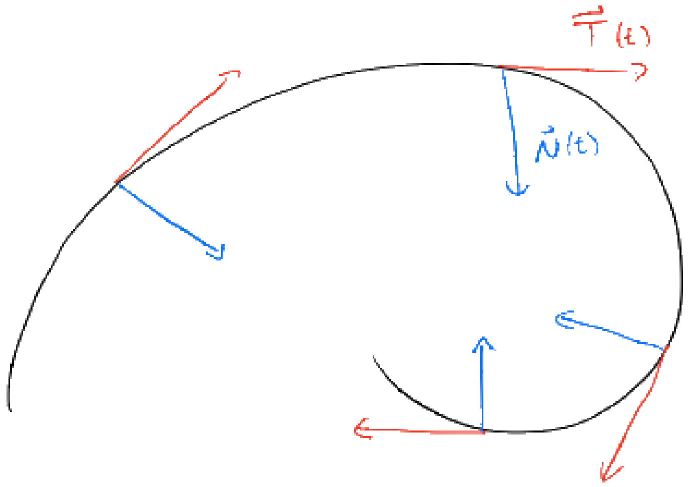
$$\kappa(t) = \sqrt{\frac{4e^{2t}t^2 - 8e^{2t}t + 4e^{7t} + e^{2t} + r}{1 + 4t^2 + e^{2t}}}.$$

At the point $\vec{r}(0) = (0, 0, 1)$ this becomes

$$\kappa(0) = \sqrt{\frac{0 - 0 + 4 + 1 + 4}{1 + 0 + 1}} = \frac{3}{\sqrt{2}}.$$

Given any point on a smooth curve we can define the ***unit normal vector*** of the curve at that point as the unit vector which points in the direction of curvature. As $\vec{T}'(t)$ points in the direction of curvature, we thus have that the unit normal vector (denoted $\vec{N}(t)$) is given by

$$\vec{N}(t) = \frac{\vec{T}'(t)}{\|\vec{T}'(t)\|}.$$



Example 2.18.
Calculate the unit normal of $\vec{r}(t) = \langle t^2, \frac{2}{3}t^3, t \rangle$ at $(1, \frac{2}{3}, 1)$.
We simply compute all the relevant quantities by differentiating

and normalizing:

$$\begin{aligned}
\vec{r}'(t) &= \left\langle 2t, 2t^2, 1 \right\rangle \\
\|\vec{r}'(t)\| &= \sqrt{4t^2+4t^4+1} = \sqrt{(2t^2+1)^2} = 2t^2+1 \\
\vec{T}(t) &= \left\langle \frac{2t}{2t^2+1}, \frac{2t^2}{2t^2+1}, \frac{1}{2t^2+1} \right\rangle \\
\vec{T}'(t) &= \left\langle \frac{4t^2+1-8t^2}{4t^2+4t^4+1}, \frac{8t^3+4t-8t^2}{4t^2+4t^4+1}, \frac{-4t}{4t^2+4t^4+1} \right\rangle .
\end{aligned}$$

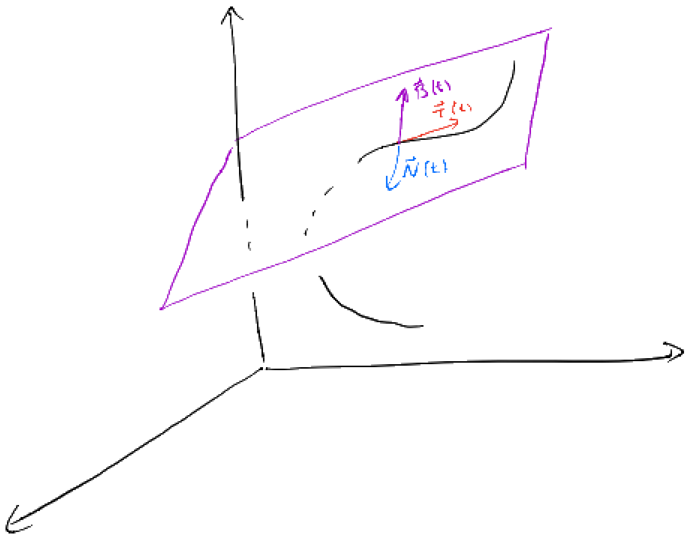
Notice now that our point corresponds to $t = 1$. We can thus compute

$$\begin{aligned}
\vec{T}'(1) &= \left\langle \frac{-3}{9}, \frac{4}{9}, \frac{-4}{9} \right\rangle = \frac{1}{9}\langle -3, 4, -4 \rangle \\
\|\vec{T}'(1)\| &= \sqrt{\frac{9}{81}+\frac{16}{81}+\frac{16}{81}} = \frac{\sqrt{41}}{9}.
\end{aligned}$$

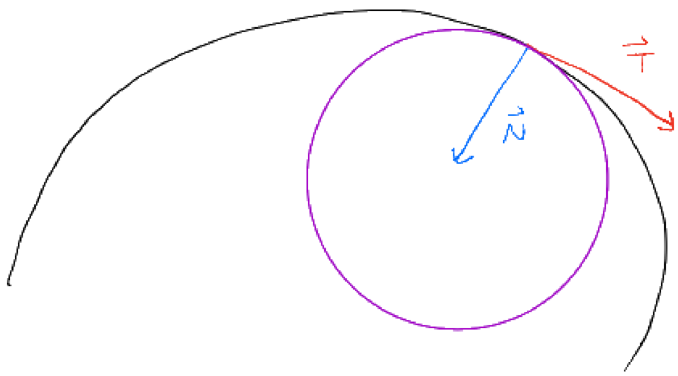
Now we can determine the unit normal vector as

$$N(1) = \frac{\vec{T}'(1)}{\|\vec{T}'(1)\|} = \frac{\langle -3, 4, -4 \rangle}{\sqrt{41}}.$$

At every point on our curve we can associate a plane by considering the plane containing our unit tangent vector $\vec{T}(t)$ and unit normal vector $\vec{N}(t)$ and which contains our point on the curve. Since the equation of a plane depends on a vector orthogonal go the plane, if we were to write down the equation of this plane we would need that orthogonal vector, which we can compute using the cross product $\vec{T}(t) \times \vec{N}(t)$. We typically call this vector the "normal vector" of the plane, but here we are already using the word "normal" to mean the unit normal vector $\vec{N}(t)$. Thus to avoid confusion (or, perhaps add to the confusion?) we call this orthogonal vector the ***binormal vector*** of the curve at the given point, denoted $\vec{B}(t) = \vec{T}(t) \times \vec{N}(t)$.



Inside of this plane we consider a circle called the ***osculating circle*** of the curve. This is a circle which is tangent to the curve (in the sense that the circle and the curve intersect at a given point and have the same (or opposite) unit normal vectors), has the same curvature as the curve, and the unit normal $\vec{N}(t)$ points towards the center of the circle.



Example 2.19.
Find the equation of the osculating circle to the graph $y = x^2$ at the

origin.

We first find a parametrization for our curve,

$$\vec{r}(t) = \left\langle t, t^2, 0 \right\rangle.$$

We can then compute the tangent vector,

$$\vec{r}\,'(t) = \langle 1, 2t, 0 \rangle,$$

which has magnitude

$$\|\vec{r}\,'(t)\| = \sqrt{1+4t^2}.$$

The unit tangent vector is then

$$\vec{T}(t) = \left\langle \frac{1}{\sqrt{1+4t^2}}, \frac{2t}{\sqrt{1+4t^2}}, 0 \right\rangle.$$

We can then compute the derivative of the unit tangent vectors to be

$$\vec{T}\,'(t) = \left\langle \frac{-8t}{2}\left(1+4t^2\right)^{-3/2}, 2(1+4t^2)^{-1/2} - 8t^2(1+4t^2)^{-3/2}, 0 \right\rangle$$

We can now compute the curvature as

$$\kappa(0) = \frac{\|\vec{T}\,'(0)\|}{\|\vec{r}\,'(0)\|} = 2.$$

Thus the osculating circle will have radius $^1/_2$.

The unit normal at the origin, corresponding to $t = 0$, is

$$\vec{N}(0) = \frac{\vec{T}\,'(0)}{\|\vec{T}\,'(0)\|} = \langle 0, 1, 0 \rangle.$$

and so the equation of the osculating circle is

$$x^2 + \left(y - \frac{1}{2}\right)^2 = \frac{1}{4}.$$

2.5 Motion in space and integrals of vector-valued functions

We end this chapter by quickly mentioning some applications of the material we have discussed to the motion of a particle in three-space. In doing so we will also introduce the idea of integrating a vector-valued function and see that this can be interpreted as computing displacement from velocity.

Motion of a particle in 3-space

We have thought about parametrizations and vector-valued functions as defining a curve in space, but if we interpret the variable t to be time, we can also think of a vector-valued function as telling us the position of a particle as it moves in space. That is, at some moment in time we begin to record the position of a small object as it flies around in 3-space by keeping track of its x-, y-, and z-coordinates. As these change over time, they determine functions $x(t)$, $y(t)$, and $z(t)$. That is, we have exactly the same information as in a parametrization of a curve.

There are several physical quantities we may associate to this particle as it moves around in space, such as its total displacement, its velocity, its acceleration, and the distance travelled. These are all closely related to quantities we've already discussed such as tangent vectors, normal vectors and curvature, and arclength.

To be more explicit, let's first consider determining the velocity of the particle. Velocity is a change in position (aka displacement) over a change in time. In particular, if the particle is at position $\vec{r}(t_0) = \langle x(t_0), y(t_0), z(t_0)\rangle$ at some moment in time t_0 and then moves to position

$$\vec{r}(t_0 + h) = \langle x(t_0 + h), y(t_0 + h), z(t_0 + h)\rangle$$

over the course of h units of time, then the average velocity of the particle is simply the displacement divided by the time it took for that displacement to occur, which is just

$$\frac{\vec{r}(t_0 + h) - \vec{r}(t_0)}{h} = \left\langle \frac{x(t_0 + h) - x(t_0)}{h}, \frac{y(t_0 + h) - y(t_0)}{h}, \frac{z(t_0 + h) - z(t_0)}{h} \right\rangle .$$

This quantity is the ***average velocity*** of the particle over the time interval from t_0 to $t_0 + h$. To compute the ***instantaneous velocity*** of the particle at time t_0, we simply take the limit of the average velocities as the amount of time shrinks to zero, and this gives us

$$\lim_{h\to 0} \frac{\vec{r}(t_0+h)-\vec{r}(t_0)}{h}$$
$$= \left\langle \lim_{h\to 0} \frac{x(t_0+h)-x(t_0)}{h}, \lim_{h\to 0} \frac{y(t_0+h)-y(t_0)}{h}, \lim_{h\to 0} \frac{z(t_0+h)-z(t_0)}{h} \right\rangle$$

Notice, though, this is exactly the derivative $\vec{r}\,'(t_0)$. That is, the velocity of a particle is simply the derivative of its position, just as in one-dimension. Notice that velocity is a vector-valued quantity: it tells us not simply how fast we are moving, but in which direction we are moving. If we only care about the ***speed*** of the particle, then we want to convert the velocity vector into a scalar, which is done simply by taking the magnitude. That is, the speed of a particle with position $\vec{r}(t)$ is just $\|\vec{r}\,'(t)\|$.

Example 2.20.
Suppose the position of a particle at time t is given by $(3\sin(\pi t),\ e^{t^2+3t},\ \sqrt{t})$. Find the instantaneous velocity of the particle at time $t=1$. Find the speed of the particle at time $t=1$ as well.

Letting $\vec{r}(t) = \left\langle 3\sin(\pi t),\ e^{t^2+3t},\ \sqrt{t}\right\rangle$, we simply calculate $\vec{r}\,'(1)$. Note first

$$\vec{r}\,'(t) = \left\langle 3\pi\cos(\pi t),\ e^{t^2+3t}(2t+3),\ \frac{1}{2\sqrt{t}} \right\rangle$$
$$\implies \vec{r}\,'(1) = \left\langle -3\pi, 5e^4, \frac{1}{2} \right\rangle$$

This is velocity, so the speed is the magnitude of this vector,

$$\|\vec{r}\,'(1)\| = \sqrt{9\pi^2 + 25e^8 + \frac{1}{4}} \approx 273.154.$$

The acceleration of a particle is, of course, simply the derivative of its velocity; the acceleration is the second derivative of position.

Example 2.21.
Find the instantaneous acceleration at time $t = 1$ of the particle

whose position at time t is $(3\sin(\pi t), e^{t^2+3t}, \sqrt{t})$.

First we calculate the second derivative,

$$\begin{aligned}\vec{r}(t) &= \left\langle 3\sin(\pi t),\ e^{t^2+3t},\ \sqrt{t}\right\rangle \\ \implies \vec{r}'(t) &= \left\langle 3\pi\cos(\pi t),\ e^{t^2+3t}(2t+3),\ \frac{1}{2\sqrt{t}}\right\rangle \\ \implies \vec{r}''(t) &= \left\langle -3\pi^2\sin(\pi t),\ (2t+3)^2e^{t^2+3t}+2e^{t^2+3t},\ \frac{-1}{4t^{3/2}}\right\rangle\end{aligned}$$

Thus the acceleration at $t=1$ is

$$\vec{r}''(1) = \left\langle 0, 27e^4, -1/4\right\rangle.$$

Some derivative rules

We have lots of ways of combining vectors (vector addition, scalar multiplication, dot products, cross products), and so do all of these operations on vector-valued functions as well. That is, if $\vec{r}(t)$ and $\vec{s}(t)$ are two vector-valued functions, it makes sense to talk about $\vec{r}(t)\cdot\vec{s}(t)$ or $\vec{r}(t)\times\vec{r}(t)$.

Example 2.22.
Let $\vec{r}(t)$ and $\vec{s}(t)$ be the following two vector-valued functions.

$$\begin{aligned}\vec{r}(t) &= \left\langle t, t^2, t^3\right\rangle \\ \vec{s}(t) &= \left\langle 2t^2, e^t, t\right\rangle.\end{aligned}$$

Calculate $\vec{r}(t)\cdot\vec{s}(t)$ and $\vec{r}(t)\times\vec{s}(t)$.

$$\begin{aligned}\vec{r}(t)\cdot\vec{s}(t) &= \left\langle t, t^2, t^3\right\rangle\cdot\left\langle 2t^2, e^t, t\right\rangle \\ &= 2t^3 + e^tt^2 + t^4.\end{aligned}$$

(Note this is a scalar-valued function!)

$$\begin{aligned}
\vec{r}(t) \times \vec{s}(t) &= \left\langle t, t^2, t^3 \right\rangle \times \left\langle 2t^2, e^t, t \right\rangle \\
&= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ t & t^2 & t^3 \\ 2t^2 & e^t & t \end{bmatrix} \\
&= \vec{\imath}(t^3 - e^t t^3) - \vec{\jmath}(t^2 - 2t^5) + \vec{k}(te^t - 2t^4) \\
&= \left\langle t^3(1 - e^t), 2t^5 - t^2, te^t - 2t^4 \right\rangle .
\end{aligned}$$

Given that we can combine vector-valued functions in various ways, we'd like to know how to take derivatives of vector-valued functions that are built out of other vector-valued functions. For example, we'd like to have something like a product rule for the cross product of two vector-valued functions.

Theorem 2.4.
Let $\vec{r}(t)$ and $\vec{s}(t)$ be vector-valued functions, let $f(t)$ be a scalar-valued function, and let $\lambda \in \mathbb{R}$ be a scalar. We then have the following:

(*i*) $\dfrac{d}{dt}\left(\vec{r}(t) + \vec{s}(t)\right) = \vec{r}\,'(t) + \vec{s}\,'(t)$

(*ii*) $\dfrac{d}{dt}\lambda\vec{r}(t) = \lambda\vec{r}\,'(t)$

(*iii*) $\dfrac{d}{dt}\vec{r}(t) \times \vec{s}(t) = \vec{r}\,'(t) \times \vec{s}(t) + \vec{r}(t) \times \vec{s}\,'(t)$

(*iv*) $\dfrac{d}{dt}\vec{r}(t) \cdot \vec{s}(t) = \vec{r}\,'(t) \cdot \vec{s}(t) + \vec{r}(t) \cdot \vec{s}\,'(t)$

(*v*) $\dfrac{d}{dt}f(t)\vec{r}(t) = f'(t)\vec{r}(t) + f(t)\vec{r}\,'(t)$

(*vi*) $\dfrac{d}{dt}\vec{r}(f(t)) = f'(t)\vec{r}\,'(f(t))$

Indefinite integrals

Recall that for a real-valued function $f : \mathbb{R} \to \mathbb{R}$, the *indefinite integral* of f was just the anti-derivative of f. That is, if $F(x) = \int f(x)\,dx$, then $F'(x) = f(x)$. We can define indefinite integrals for vector-valued functions as well. This means that if $\vec{r}(t)$ is a vector-valued function, then its *anti-derivative* is a vector-valued function $\vec{R}(t)$ such that $\vec{R}\,'(t) = \vec{r}(t)$. As in the case of real-valued functions, this anti-derivative is sometimes denoted

$$\vec{R}(t) = \int \vec{r}(t)\,dt.$$

Since we differentiate vector-valued functions by differentiating their components, we similarly integrate these functions by integrating their components.

Example 2.23.
Find the anti-derivative of $\vec{r}(t) = \langle \sin(t), t^3 + 3t^2, e^{5t} \rangle$.

$$\begin{aligned}
\vec{R}(t) &= \int \vec{r}(t)\,dt \\
&= \int \left\langle \sin(t), t^3 + 3t^2, e^{5t} \right\rangle\,dt \\
&= \left\langle \int \sin(t)\,dt, \int \left(t^3 + 3t^2\right)\,dt, \int e^{5t}\,dt \right\rangle \\
&= \left\langle -\cos(t) + C_1, \frac{t^4}{4} + t^3 + C_2, \frac{e^{5t}}{5} + C_3 \right\rangle \\
&= \left\langle -\cos(t), \frac{t^4}{4} + t^3, \frac{e^{5t}}{5} \right\rangle + \langle C_1, C_2, C_3 \rangle \\
&= \left\langle -\cos(t), \frac{t^4}{4} + t^3, \frac{e^{5t}}{5} \right\rangle + \vec{C}
\end{aligned}$$

Notice that just as we pick up a $+C$ when we integrate a real-valued function, we pick up a $+\vec{C}$ when we integrate a scalar-valued function, but now $\vec{C}$ is a vector (coming from the $+C$'s that appear when we integrate each component).

If we want to get rid of the $+\vec{C}$, we need some extra information about the antiderivative. This extra bit of information is called an *initial value*, and an *initial value problems* refers to an integration problem where we try to get rid of the $+\vec{C}$.

Example 2.24.
Suppose the velocity of a particle at time t is given by

$$\vec{v}(t) = \left\langle 3t^2 + 2t, \cos(\pi t), 3 \right\rangle$$

and that the position of the particle at time $t = 2$ is $(0, -1, 3)$. Find the position of the particle at time t.

First we integrate the velocity to get position:

$$\begin{aligned}\vec{r}(t) &= \int \vec{v}(t)\, dt \\ &= \left\langle \int \left(3t^2 + 2t\right)\, dt, \int \cos(\pi t)\, dt, \int 3\, dt \right\rangle \\ &= \left\langle t^3 + t^2, \frac{\sin(\pi t)}{\pi}, 3t \right\rangle + \langle C_1, C_2, C_3 \rangle .\end{aligned}$$

To solve for $\langle C_1, C_2, C_3 \rangle$ we use the fact that $\vec{r}(2) = \langle 0, -1, 3 \rangle$. Thus

$$\begin{aligned}\vec{r}(2) &= \left\langle 2^3 + 2^2, \frac{\sin(2\pi)}{\pi}, 6 \right\rangle + \langle C_1, C_2, C_3 \rangle \\ &= \langle 12, 0, 6 \rangle + \langle C_1, C_2, C_3 \rangle\end{aligned}$$

Since we're suppose to have

$$\langle 12, 0, 6 \rangle + \langle C_1, C_2, C_3 \rangle = \langle 0, -1, 3 \rangle ,$$

we must have that $C_1 = -12$, $C_2 = -1$, and $C_3 = -3$. Hence the position of the particle at time t is

$$\vec{r}(t) = \left\langle t^3 + t^2 - 12, \frac{\sin(\pi t)}{\pi} - 1, 3t - 3 \right\rangle .$$

Definite integrals

We can also calculate definite integrals of vector-valued functions. First recall how the definite integral of a function $f : \mathbb{R} \to \mathbb{R}$, over the interval $[a, b]$ is defined. We break $[a, b]$ into several pieces (called a *partition* of the interval),

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

This partition is just a set of points in the interval, call this $\mathcal{P}$ – so $\mathcal{P} = \{x_0, x_1, .., x_n\}$. We'll let $|\mathcal{P}|$ denote the *norm* of the partition – that is, $|\mathcal{P}|$ is just the largest distance between two points in the partition.

$$|\mathcal{P}| = \max_{1 \leq i \leq n} (x_i - x_{i-1}) .$$

Now we pick a point $x_i^* \in [x_{i-1}, x_i]$, and calculate a Riemann sum:

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

where $\Delta x_i = x_i - x_{i-1}$. The idea is essentially that we're trying to sum up *every* value of $f(x)$ for every $x \in [a, b]$. To do this we suppose that $f(x)$ is constant on the subintervals determined by the partition. This gives us an approximation to the integral, and to make the approximation better we use a smaller partition (that is, a partition where the points are close together). To get the "best" approximation, we take the limit as the points of the approximation get arbitrarily close together. This limit (if it exists) is called the *definite integral of f from $x = a$ to $x = b$* and is denoted

$$\int_a^b f(x)\, dx = \lim_{|\mathcal{P}| \to 0} \sum_{i=1}^{\#\mathcal{P}} f(x_i^*) \Delta x_i.$$

We can do the exact same thing for definite integrals of a vector valued function. We cut an interval $a < t < b$ into several pieces, assume our vector-valued function is constant on each piece, then add up these constants (which are vectors). We then take the limit as the points in our partition get arbitrarily close together.

$$\begin{aligned}\int_a^b \vec{r}(t)\,dt &= \lim_{|\mathcal{P}|\to 0}\sum_{i=1}^{\#\mathcal{P}} \vec{r}(t_i^*)\Delta t_i \\ &= \lim_{|\mathcal{P}|\to 0}\sum_{i=1}^{\#\mathcal{P}} \langle x(t_i^*)\Delta t_i, y(t_i^*)\Delta t_i, z(t_i^*)\Delta t_i\rangle \\ &= \left\langle \lim_{|\mathcal{P}|\to 0}\sum_{i=1}^{\#\mathcal{P}} x(t_i^*)\Delta t_i, \lim_{|\mathcal{P}|\to 0}\sum_{i=1}^{\#\mathcal{P}} y(t_i^*)\Delta t_i, \lim_{|\mathcal{P}|\to 0}\sum_{i=1}^{\#\mathcal{P}} z(t_i^*)\Delta t_i \right\rangle \\ &= \left\langle \int_a^b x(t)dt, \int_a^b y(t)dt, \int_a^b z(t)dt \right\rangle.\end{aligned}$$

Usually when you first learn about integration, definite integrals are motivated by trying to find the area under the graph of a function. In the case of vector-valued functions, we don't really have a nice explanation of areas under curves, but we can discuss other applications of the definite integral.

Example 2.25.
Suppose that the velocity of a particle at time t is

$$\vec{v}(t) = \left\langle e^t, t^4, t-1 \right\rangle.$$

Find the displacement of the particle from time $t = 3$ to time $t = 5$.

We know that displacement means change in position: that is, you subtract the starting point from the end point. We also know that position is the integral of velocity. So what we want to do is integrate velocity to get position, then subtract our start point from out end point to get displacement. This is exactly what it means for us to calculate the definite integral below.

$$\begin{aligned}
\text{Displacement} &= \int_3^5 \left\langle e^t, t^4, t-1 \right\rangle dt \\
&= \left\langle \int_3^5 e^t dt, \int_3^5 t^4 dt, \int_3^5 (t-1) dt \right\rangle \\
&= \left\langle e^5 - e^3, \frac{5^5 - 3^5}{5}, \left(\frac{5^2}{2} - 5\right) - \left(\frac{3^2}{2} - 3\right) \right\rangle \\
&= \left\langle e^3(e^2-1), \frac{2882}{5}, 6 \right\rangle .
\end{aligned}$$

3 Functions of Multiple Variables

The further a mathematical theory is developed, the more harmoniously and uniformly does its construction proceed, and unsuspected relations are disclosed between hitherto separated branches of the science.

David Hilbert

3.1 Basic Ideas

In this section we begin the study of functions of several variables, a topic which will be fundamental throughout the remainder of the course. For now we will only define the basic ideas, and will study more advanced notions, necessary to doing calculus with multivariable functions, in later lectures.

The basic idea

Recall that a *function* (also known as a *map*) from a set A to a set B, is a way of associating elements in B to elements in A. If we call our function f, we may write $f : A \to B$ to say that f takes elements in A to elements in B. We call A the *domain* of f, and we call B the *codomain* of f. The *range* of f is the set of elements in B which are actually obtained as outputs of f.

For example, we may write $f : \mathbb{R} \to \mathbb{R}$ to denote that f takes real numbers as its input and gives real numbers as its output. If our function was given by $f(x) = x^2$ (which might also be denoted as $x \mapsto x^2$), then the range of f is $[0, \infty)$, even though the codomain of f is all of $\mathbb{R}^2$. (This distinction between range and codomain may seem a bit subtle and strange at first.)

In previous courses you've primarily studied functions of the form $\mathbb{R} \to \mathbb{R}$, and during the last few weeks we've considered vector-valued functions which have the form $\mathbb{R} \to \mathbb{R}^2$ or $\mathbb{R} \to \mathbb{R}^3$. Now we want to consider functions which take two values as inputs and whose output is a real number. That is, we consider functions $\mathbb{R}^2 \to \mathbb{R}$.

When we write $f : \mathbb{R}^2 \to \mathbb{R}$ we're saying that the function f has two real numbers as inputs. Usually we'll call these two inputs x and y and consider the inputs as points in the xy-plane. For example, if we write $f(x, y) = xy + x^2$, we have a function of two variables whose output is a single real number. Since such functions have multiple inputs, we say these are *multivariable functions*. To evaluate a multivariable function, we simply plug in real numbers for x and y, then evaluate the expression defining the function.

In the case of the function $f(x, y) = xy + x^2$, to evaluate the function at a point $(x, y) = (-2, 3)$, we'd evaluate

$$f(-2, 3) = (-2) \cdot 3 + (-2)^2 = -6 + 4 = -2.$$

The domain of this function is all of $\mathbb{R}^2$ (we can plug in any values we'd like for x and y), and the range is all of $\mathbb{R}$. We can make $xy + x^2$ be any value we'd like by picking x and y appropriately. For example, to get an output of 9, we'd need to find an x and y such that $xy + x^2 = 9$. This will happen if we pick $x = 1$ and $y = 8$. In general, to get $f(x, y) = c$, we could set $x = 1$ and $y = c - 1$. (This is only one possible choice for picking an appropriate x and y – there are lots of other choices we could make!)

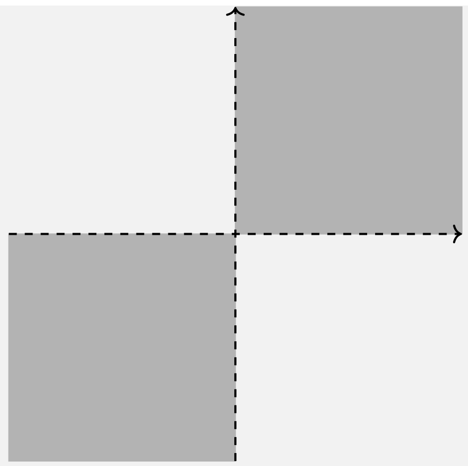
Example 3.1.
What's the domain and range of the function $f(x, y) = \ln(xy)$?

Here we have to worry about plugging a non-positive number into $\ln$. If either of x or y is zero, then we'll try to evaluate $\ln(0)$ which is undefined. So neither x nor y may be zero. Can x or y be negative? As long as both values are negative we'd be okay, since the product of two negative numbers is positive. Similarly, it'd be fine if x and y were both positive. However, if one of x or y is positive and the other is negative, then we'd try to evaluate $\ln$ of a negative number.

Putting all of this together, the domain of $f(x, y) = \ln(xy)$ is the set of all (x, y) pairs where $xy > 0$:

$$\left\{(x, y) \in \mathbb{R}^2 \,\middle|\, xy > 0\right\}.$$

This means we can plug in only (x, y)-pairs from the first and third quadrants.



The range of this function is all of $\mathbb{R}$. If we want $\ln(xy) = c$, then we can set $x = 1$ and $y = e^c$.

Notice that in image of the domain of $\ln(xy)$ in the example above, the edges of the domain of the function are dashed. It is a standard convention to dash the boundary of the domain if the boundary is not included, and to make the boundary solid if the boundary is part of the domain. This is illustrated in the next example.

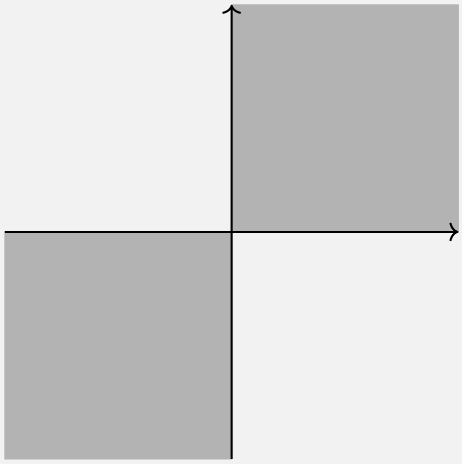
Example 3.2.
What's the domain and range of the function $f(x, y) = \sqrt{xy}$?

Here we have to worry about plugging a non-positive number into $\sqrt{xy}$. If either of x or y is zero, then we'll try to evaluate $\sqrt{0}$ which is undefined. So neither x nor y may be zero. Can x or y be negative? As long as both values are negative we'd be okay, since the product of two negative numbers is positive. Similarly, it'd be fine if x and y were both positive. However, if one of x or y is positive and the other is negative, then we'd try to evaluate the square root of a negative number.

Putting all of this together, the domain of $f(x, y) = \sqrt{xy}$ is the set of all (x, y) pairs where $xy \geq 0$:

$$\left\{(x, y) \in \mathbb{R}^2 \,\middle|\, xy \geq 0\right\}.$$

This means we can plug in only (x, y)-pairs from the first and third quadrants.



Unlike the last example, the range of this function is only the set of non-negative real numbers: there's no way to get a negative number out of $f(x, y) = \sqrt{xy}$.[a]

[a] This is true if we suppose the inputs to our function have to be real numbers. Of course we allowed complex numbers into our function, then we could get negative numbers out of the square root.

Note how in the example above the boundary of the domain is included as part of the domain, and so in our picture of the domain we made the boundary solid instead of dashed.

Sometimes to describe the output of a multivariable function we'll use a table where the top row and left-most column represent the values we plug in for x and y, respectively, and then the corresponding cell of the table indicates the value of the function at that point. For example, a portion of the possible outputs from the function $f(x, y)$ is indicated in the table below.

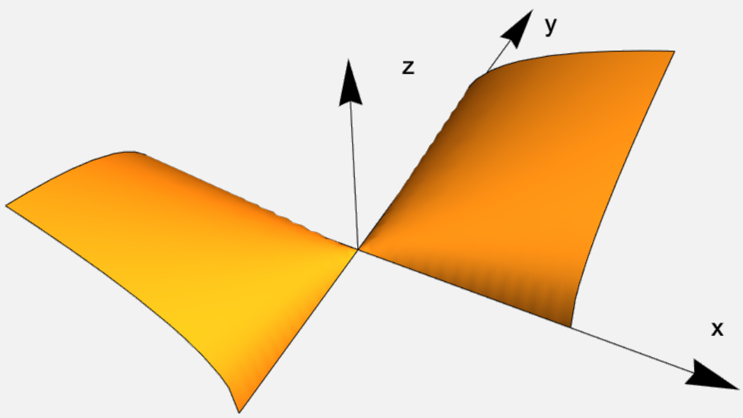
$x \backslash y$	**0**	**1**	**2**	**3**	**4**
0	0	0	0	0	0
1	0	1	$\sqrt{2}$	$\sqrt{3}$	$\sqrt{4}$
2	0	$\sqrt{2}$	2	$\sqrt{6}$	$\sqrt{8}$
3	0	$\sqrt{3}$	$\sqrt{6}$	3	$\sqrt{12}$
4	0	2	$\sqrt{8}$	$\sqrt{12}$	4

Graphs

In general, the graph of a function $f : A \to B$ is the set of pairs $(x, f(x)) \in A \times B$. If we're looking at functions $f : \mathbb{R}^2 \to \mathbb{R}$, this means that the graph will be the set of all triples $(x, y, f(x, y))$ in $\mathbb{R}^3$ where (x, y) is in the domain of f.

Example 3.3.
What is the graph of the function $f(x, y) = \sqrt{xy}$?

This is the set of all triples of the form $(x, y, \sqrt{xy})$ where (x, y) is in the domain of $\sqrt{xy}$; so (x, y) come from the first and third quadrants of the xy-plane, including the x- and y-axes. For example, the triples $(0, 0, 0)$, $(2, 3, \sqrt{6})$, $(-2, -3, \sqrt{6})$, $(4, 1, \sqrt{2})$, and so on are on the graph of this function. If we look at all possible triples we could get, this gives us set of points in $\mathbb{R}^3$ indicated below.



In general, the graph of a function $f(x, y)$ is called a *surface*. We will sometimes denote this surface by writing $z = f(x, y)$, in the same way that you'd write $y = f(x)$ to denote the curve in $\mathbb{R}^2$ given by graph of a function of one variable. Note that a surface defined in this way is a two-dimensional object that sits inside of three-dimensional space. (Why is this a two-dimensional object? Recall that we need two values to specify a point on a two-dimensional object. Here are two values are the x and y that we plug into $f(x, y)$.)

Example 3.4.
What is the graph of the function $f(x, y) = 3x + 2y + 4$?

We're looking for the graph of the function, so we want to find all of the pairs (x, y, z) where $z = 3x+2y+4$. If we do a little algebra, we see that this equation may be written as

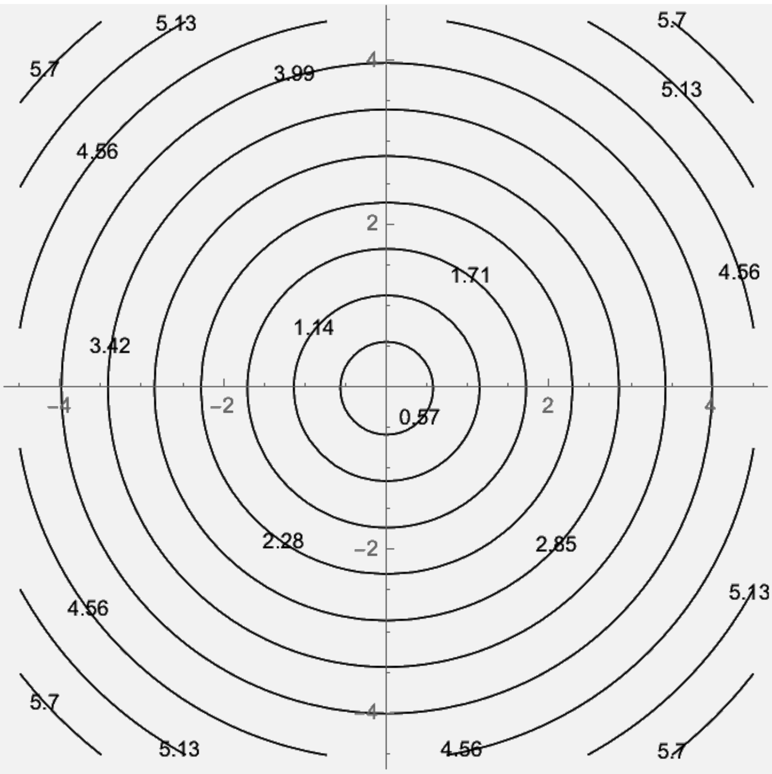
$$3x + 2y - z = -4.$$

Thus we see that the graph of our function is a plane. In fact, it's a plane passing through the point $(0, 0, 4)$ with normal vector $\langle 3, 2, -1\rangle$.

Level curves

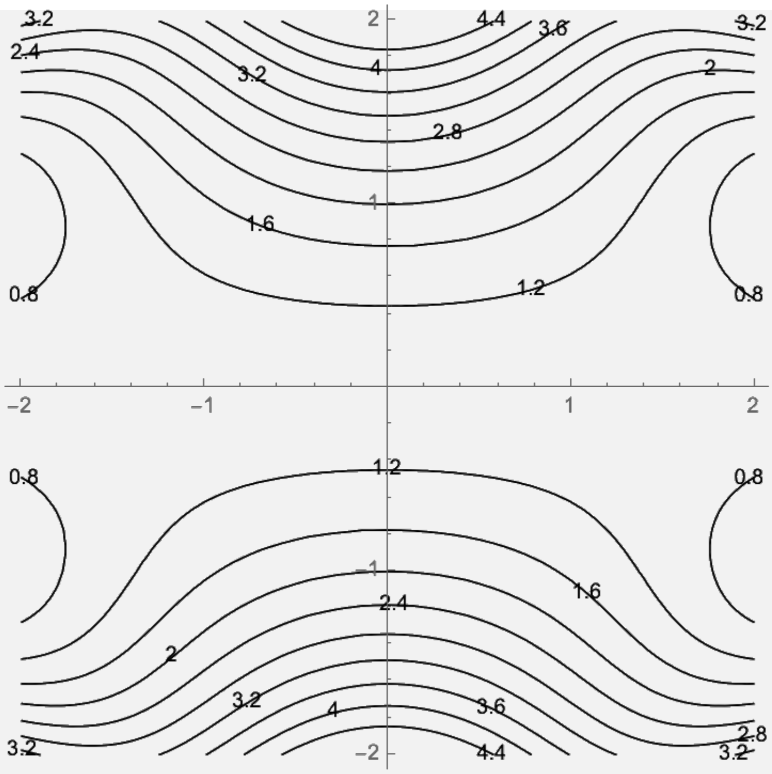
It can be difficult to figure out what the graph of a function looks like if you don't have a computer or graphing calculator handy. In particular, if you lived before the 1970's, you didn't have a very nice way of graphing these surfaces. One thing you could do to get a feel for what the surface is, however, is to look at a *contour map*. That is, instead of trying to visualize the graph $z = f(x, y)$ directly, you might instead look at curves of the form $c = f(x, y)$.

Example 3.5.
For example, let's try to figure out what the graph of $z = \sqrt{x^2 + y^2}$ looks like. To do this, let's consider what would happen if we intersect this surface with the plane $z = 1$. We'd then have the equation $1 = \sqrt{x^2 + y^2}$, which we could write as $x^2 + y^2 = 1$. This is a circle of radius 1 in the plane $z = 1$. If we instead looked at the intersection of the plane $z = 4$, then we'd have $4 = \sqrt{x^2 + y^2}$, or $x^2 + y^2 = 16$, which is a circle of radius 4 in the plane $z = 4$. Looking at lots and lots of these curves, projected down to the xy-plane, we see the contour map below.

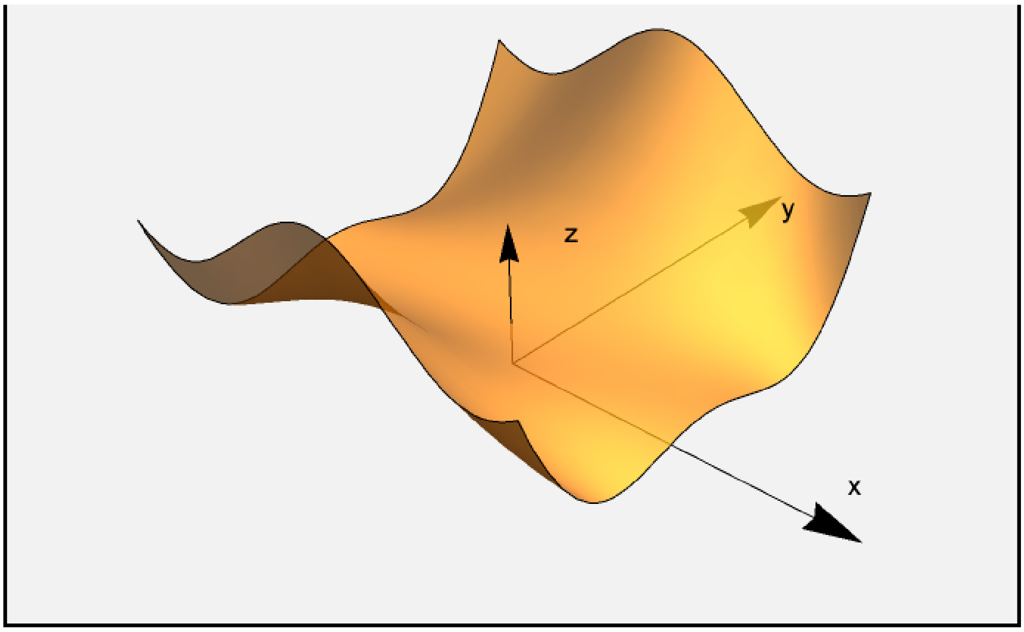


The contour plot in the example above helps us see that the surface $z = \sqrt{x^2 + y^2}$ is a cone that gets wider as z increases.

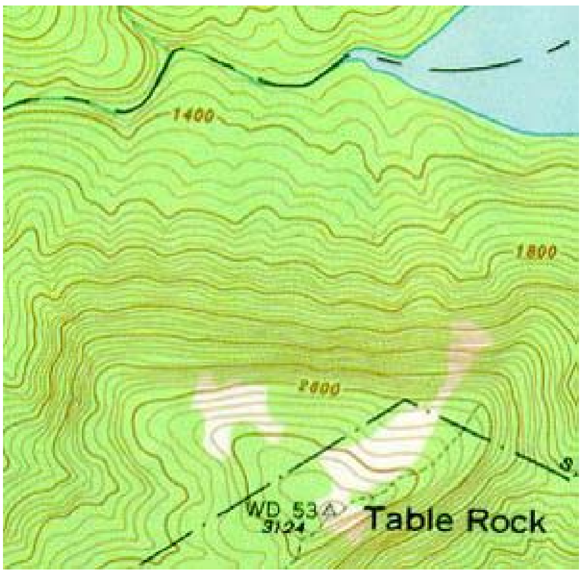
Example 3.6.
What does the contour plot of $z = \cos(xy) + y^2$ below tell us about the shape of the surface?



One thing we can immediately read off of this contour map is that if you move away from the x-axis, the surface starts to increase in elevation. As the contours are getting closer together, this means the surface is getting steeper. The actual surface is shown in the image below.



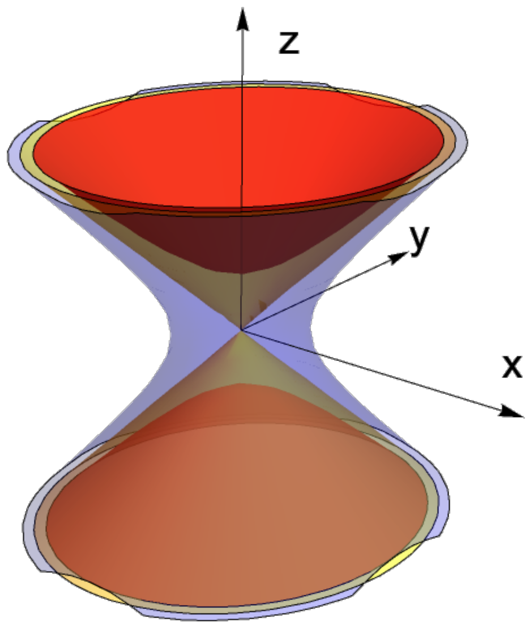
Since the contours, the individual curves given by $c = f(x, y)$ where c is a constant, represent a path we could walk on the surface to stay at the same elevation, these are sometimes called *level curves*. Note that this idea is really the same as a topographic map, a map that tells you elevations. See the topographic map for a portion of Table Rock State Park below.



Functions of more than two variables

We can use an idea similar to contour maps to try to wrap our heads around functions of three or more variables. If we have a function of three variables, say $f : \mathbb{R}^3 \to \mathbb{R}$, then the graph of this function lives inside of four-dimensional space. (The graph here is the set of points $(x, y, z, w) \in \mathbb{R}^4$ where $w = f(x, y, z)$.) This presents a little bit of a problem for us since we can't really visualize four-dimensional space. What we can do, however, is look at *level surfaces* of the function. That is, we look at surfaces of the form $c = f(x, y, z)$ where c is a constant. This means we look at the set of all points $(x, y, z) \in \mathbb{R}^3$ that satisfy the equation $c = f(x, y, z)$.

Consider for example the function $f(x, y, z) = x^2 + y^2 - z^2$. The level surfaces of this function are solutions to equations $c = x^2 + y^2 - z^2$. When $c = 0$ we have $0 = x^2+y^2-z^2$, or simply $z^2 = x^2+y^2$ which is a double cone. When $c = 1$ we have $x^2 + y^2 - z^2 = 1$ which is a *one-sheeted hyperboloid*. When $c = -1$ we have $x^2 + y^2 - z^2 = -1$ which is a *two-sheeted hyperboloid*. These level surfaces of $x^2 + y^2 - z^2$ are shown in the image below.



3.2 Limits, continuity, and the sandwich theorem

In the last section we introduced multivariable functions. In this lecture we pave the way for doing calculus with multivariable functions by introducing limits and continuity of such functions.

Limits

Informally, the notation $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$ means that as the inputs (x,y) gets "really close" to (x_0,y_0), the outputs $f(x,y)$ get "really close" to L. We won't spend the time to make this notion precise, but it comes down to an ε-δ definition of the limit, like we saw when we defined limits of vector-valued functions.

Recall that for functions of a single variable, we could talk about right-hand and left-hand limits. That is, if our inputs were from the real line, then we could approach a value from one of two directions. When our inputs live in the plane, there infinitely many different ways for the inputs (x,y) to approach (x_0,y_0). In order for the limit $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ to exist, we must get the same value for all possible ways of approaching (x_0,y_0). Put another way, if any two paths give different values, the limit does not exist.

Example 3.7.
Determine whether or not the limit $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$ exists.

If we approach $(0,0)$ from the x axis (so $y=0$), we have

$$\lim_{x\to 0} \frac{x^2}{x^2} = 1.$$

If we instead approach $(0,0)$ from the y axis (so $x=0$), we have

$$\lim_{y\to 0} \frac{-y^2}{y^2} = -1.$$

Since these two values disagree, the limit can not exist.

Example 3.8.
Determine whether or not the limit $\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^4+3y^4}$ exists.

Let's first approach from the x-axis to get

$$\lim_{x\to 0} \frac{0}{x^4+3y^4} = 0.$$

Approaching from the y-axis,

$$\lim_{y\to 0} \frac{0}{x^4+3y^4} = 0.$$

Finally, let's approach from the line $y = x$:

$$\lim_{x\to 0} \frac{x^4}{x^4+3x^4} = \lim_{x\to 0} \frac{x^4}{4x^4} = \frac{1}{4}.$$

These values don't all agree, so the limit can not exist.

Given that there are infinitely-many different paths we'd need to check to see if a limit exists, you may wonder how on earth we're supposed to check if limits exist. The answer is that we need to some tools to help us. The main tool we need is continuity of multivariable functions, since this will let us easily calculate limits.

Continuity

We say that a function $f(x,y)$ is *continuous at the point* (x_0,y_0) if the following three conditions are met:

(i) $f(x_0,y_0)$ is defined (i.e., (x_0,y_0) is in the domain of f).

(ii) $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ exists

(iii) $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$.

If a function is continuous at every point in its domain, then we simply say that the function is *continuous*.

This means that if we know a function is continuous, then its easy to take limits: we just evaluate the function. Now what we need is a repertoire of continuous functions.

Theorem 3.1.
The following types of multivariable functions are continuous:

(*i*) *Polynomials are always continuous on all of* $\mathbb{R}^2$.

(*ii*) *Rational functions (ratios of polynomials) are continuous where they're defined (i.e., where the denominator is not zero)*

(*iii*) *If* $g : \mathbb{R} \to \mathbb{R}$ *is continuous and* $f(x, y) : \mathbb{R}to\mathbb{R}$ *is continuous, then* $g(f(x, y))$ *is continuous.*

(*iv*) *Products and sums of continuous functions are always continuous.*

(*v*) *Quotients of continuous functions are continuous where they're defined.*

(*vi*) *A composition of continuous functions, in any number of variables, is continuous.*

Let's spend a little bit of time describing each of the types of functions described in the theorem above.

A *polynomial* in two variables is a sum where each term has the form cx^iy^j where c is a real number, and i and j are positive integers. So the following functions are all polynomials:

$$5x^3y^2 + 3x^2 - 2y^3 + 4xy + 5$$
$$-3x^{17} + y^3$$
$$32x^5y^4$$
$$13$$

Polynomials are very nice functions because they're built from the basic operations of arithmetic: addition and multiplication. Since the above theorem tells us that polynomials are continuous, it's very easy to take limits of polynomials.

Example 3.9.
Calculate the limit $\lim_{(x,y)\to(3,-1)} (3x^2y - 2y^2 + x)$.

$$\lim_{(x,y)\to(3,-1)} \left(3x^2y - 2y^2 + x\right) = 3{\cdot}3^2{\cdot}(-1){-}2{\cdot}(-1)^2{+}3 = -27{-}2{+}3 = 26$$

A *rational function* in two variables is just a ratio of two polynomials. So the functions below are examples of rational functions:

$$\frac{3x^2 + xy}{4y^2 - x}$$

$$\frac{3}{x + y}$$

$$\frac{x + y^2 - 2x}{2xy}$$

The theorem above tells us that rational functions are continuous everywhere they're defined. So taking limits of rational functions is also very easy, provided that we're taking the limit at a point that's in the domain of the function.

Example 3.10.
Calculate the following limit:

$$\lim_{(x,y)\to(1,2)} \frac{3xy - y^2}{4x + 3y}.$$

Notice that the denominator, $4x + 3y$ is not zero at the point $(x, y) = (1, 2)$, so this point is in the domain of the rational function, so to take the limit we just evaluate the function:

$$\lim_{(x,y)\to(1,2)} \frac{3xy - y^2}{4x + 3y} = \frac{3 \cdot 1 \cdot 2 - 2^2}{4 \cdot 1 + 3 \cdot 2} = \frac{2}{10} = \frac{1}{5}.$$

The third condition of the theorem above, that a composition of the form $g(f(x, y))$ is continuous when $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are continuous tells us that functions such as the following are continuous:

$$\cos(x+y)$$
$$\tan^{-1}(2x^3y)$$
$$e^{x-y}$$

Once nice property about such functions is the following:

Theorem 3.2.
If $g : \mathbb{R} \to \mathbb{R}$ is continuous and $f : \mathbb{R}^2 \to \mathbb{R}$ is any function, then

$$\lim_{(x,y)\to(a,b)} g(f(x,y)) = g\left(\lim_{(x,y)\to(a,b)} f(x,y)\right).$$

provided $\lim_{(x,y)\to(a,b)} f(x,y)$ is in the domain of g.

That is, we can move limits inside of continuous functions.

Example 3.11.
Calculate the following limit:

$$\lim_{(x,y)\to(-1,4)} e^{x+\sqrt{y}}.$$

$$\lim_{(x,y)\to(-1,4)} e^{x+\sqrt{y}} = e^{\lim_{(x,y)\to(-1,4)}\left(x+\sqrt{y}\right)} = e^{-1+\sqrt{4}} = e$$

Knowing that all of these functions are continuous is very helpful, but there are still times when they aren't able to help us take limits. For example, if we wanted to calculate the limit

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2},$$

continuity doesn't help us since $(0,0)$ isn't in the domain of $\frac{x^2y^2}{x^2+y^2}$. To evaluate limits like this we need one more tool: the sandwich theorem.

One of the interesting thing about limits of multivariable functions is that, if the limit exists, we can write it as a *double limit*:

$$\lim_{(x,y)\to(a,b)} f(x,y) = \lim_{x\to a}\lim_{y\to b} f(x,y) = \lim_{y\to b}\lim_{x\to a} f(x,y).$$

Again, this is contingent on the fact that we already know $\lim_{(x,y)\to(a,b)} f(x,y)$! This may seem like a minor observation, but it turns out to be a very useful fact, particularly for proving certain theorems.

The Sandwich Theorem

The sandwich theorem tells us that if we have a function that's "sandwiched" between two other functions, then the limit has to be sandwiched as well.

Theorem 3.3 (Sandwich theorem, (aka the squeeze theorem))**.** *Suppose that* $f, g, h : \mathbb{R}^2 \to \mathbb{R}$ *are three multivariable functions defined near the point* $(a, b) \in \mathbb{R}^2$ *and such that* $f(x, y) \leq g(x, y) \leq h(x, y)$ *for all* (x, y) *near* (a, b)*. If*

$$\lim_{(x,y)\to(a,b)} f(x,y) = L = \lim_{(x,y)\to(a,b)} h(x,y),$$

then we must also have that

$$\lim_{(x,y)\to(a,b)} g(x,y) = L$$

as well.

Example 3.12.
Evaluate the following limit:

$$\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2}.$$

Let's notice first that since $(0,0)$ isn't in the domain of this function, we can't use continuity to help us evaluate this limit. To use the sandwich theorem we need to find two functions which sandwich our $\frac{x^2y^2}{x^2+y^2}$ from above and below.

Let's first notice that $\frac{x^2y^2}{x^2+y^2}$ is never negative, and so we have

$$0 \leq \frac{x^2y^2}{x^2+y^2}.$$

Let's also notice that

$$x^2 \leq x^2 + y^2$$

since adding y^2 will always make x^2 larger (as $y^2 > 0$). This means

$$\frac{x^2}{x^2+y^2} \leq 1$$

If we multiply both sides by y^2 we have $\frac{x^2y^2}{x^2+y^2} \leq y^2$. Now we have our sandwich functions:

$$0 \leq \frac{x^2y^2}{x^2+y^2} \leq y^2.$$

Taking the limit as $(x, y) \to (0, 0)$ we have

$$\begin{aligned} 0 \leq & \quad \tfrac{x^2y^2}{x^2+y^2} & \leq y^2 \\ \implies \lim_{(x,y)\to(0,0)} 0 \leq & \quad \lim_{(x,y)\to(0,0)} \tfrac{x^2y^2}{x^2+y^2} & \leq \lim_{(x,y)\to(0,0)} y^2 \\ \implies 0 \leq & \quad \lim_{(x,y)\to(0,0)} \tfrac{x^2y^2}{x^2+y^2} & \leq 0 \\ \implies & \lim_{(x,y)\to(0,0)} \tfrac{x^2y^2}{x^2+y^2} = 0 \end{aligned}$$

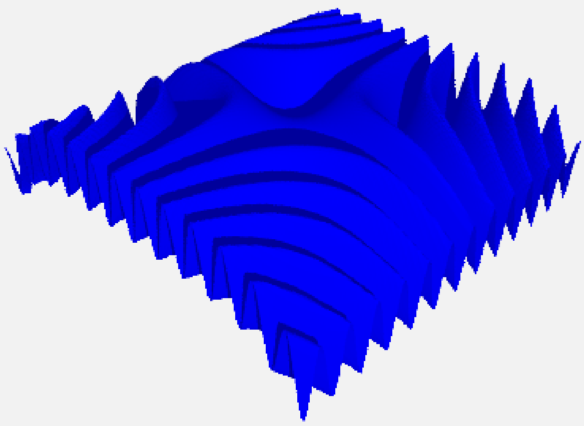
3.3 Partial derivatives

Given a function of two variables, $f(x, y)$, we may like to know how the function changes as the variables change. This knowledge of how the outputs change as the inputs change is important, for instance, when we maximize or minimize functions of multiple variables. However, in general this can be a complicated problem because the variables can change in so many different ways: one variable may increase exponentially as another decreases linearly; or one variable may decrease logarithmically while the other variable decreases quadratically; or many more complicated things could happen. In order to make our lives easier we will first suppose that only one variable changes at a time, since this is most directly related to the derivatives of single-variable functions that we understand. Later we will see how to deal with the case when both variables are changing simultaneously.

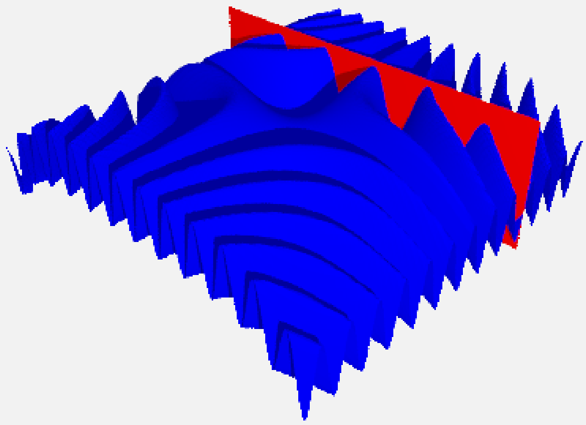
Example 3.13.
As a motivating example, consider the function

$$f(x, y) = \sin(xy) + \cos\left(\frac{x^3}{100}\right).$$

The surface $z = f(x, y)$ associated to this function consists of ripples and waves that change in complicated ways as we change x and y.

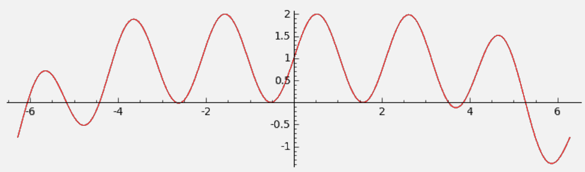


Let's suppose that y is some fixed constant, say $y = 3$, and only x-varies. Geometrically, we're intersecting the surface with the plane $y = 3$.



The intersection of our surface with the plane gives us a curve, which is naturally the graph of the function

$$F(x) = \sin(3x) + \cos\left(\frac{x^3}{100}\right).$$

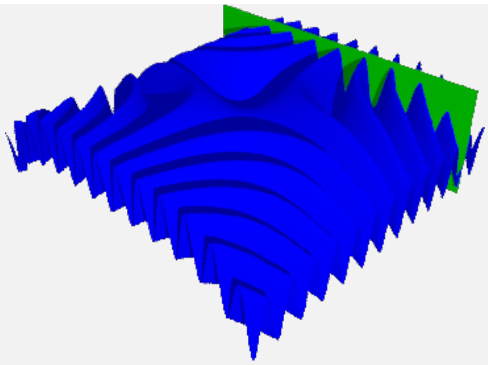


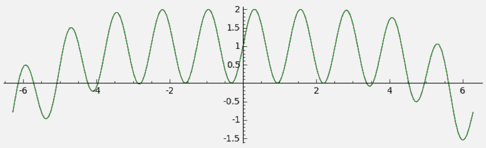
If we differentiate this function we have

$$F'(x) = 3\cos(3x) - \frac{3x^2}{100}\sin\left(\frac{x^3}{100}\right).$$

This derivative represents how the "elevation" of the surface $z = f(x, y)$ changes if we walk along the surface keeping y fixed at $y = 3$, but letting x change.

If we had instead picked $y = 5$ instead of $y = 3$, how would this change things? We intersect our surface with the plane $y = 5$ to get a different curve.





$$G(x) = \sin(5x) + \cos\left(\frac{x^3}{100}\right).$$

Now when we take the derivative we calculate,

$$G'(x) = 5\cos(5x) - \frac{3x^2}{100}\sin\left(\frac{x^3}{100}\right).$$

This is certainly a different derivative than what we calculated earlier, but it's very closely related to our previous derivative. In particular, all of the 3's that came from $y = 3$ before are now 5's, which isn't too surprising.

Partial Derivatives

In general, we could calculate the derivative of $f(x, y)$ after choosing y to be any constant we want. Instead of picking a new constant each time, though, let's suppose that we don't pick the constant we want to plug in for y yet, but instead just keep in mind that y represents a constant – whatever that constant may happen to be. When we then differentiate $f(x, y)$, treating y as a constant, we're calculating the *partial derivative of* $f(x, y)$ *with respect to* x. Notationally this partial derivative is usually denoted $\frac{\partial f}{\partial x}$ or f_x. Formally, it's defined as follows:

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \lim_{h\to 0}\frac{f(x_0+h,y_0)-f(x_0,y_0)}{h} = \lim_{x\to x_0}\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}$$

That is, we're calculate the derivative like normal, but we have this extra y thrown into our function. In the formula for the derivative above, though, notice that only x is changing (because of the $x+h$), and the y is not: the y is constant.

There's nothing special about x of course, we could repeat all of the above by letting y change and keeping x constant. This is the *partial derivative of $f(x,y)$ with respect to y*, which is denoted $\frac{\partial f}{\partial y}$ or f_y. In terms of limits,

$$\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = \lim_{h\to 0}\frac{f(x_0,y_0+h)-f(x_0,y_0)}{h} = \lim_{y\to y_0}\frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}$$

Notice that what we're doing is keeping one of the variable set to be a constant, and letting the other variable change. We can then apply all of our usual calculus rules to determine the partial derivative: just "pretend" one of the variables is constant.

Example 3.14.
Calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ of $f(x,y)=\sin(xy)+\cos\left(\frac{x^3}{100}\right)$.

To calculate $\frac{\partial f}{\partial x}$, just "pretend" that y is constant and apply your usual rules from calculus, differentiating with respect to x:

$$\frac{\partial f}{\partial x} = y\cos(xy) - \frac{3x^2}{100}\sin\left(\frac{x^3}{100}\right).$$

To calculate $\frac{partial f}{\partial y}$, repeat the process, but this time pretend x is the constant and y the variable:

$$\frac{\partial f}{\partial y} = x\cos(xy).$$

Example 3.15.
Calculate both partial derivatives of $z = x^y + 2^{xy}$.

$$\frac{\partial z}{\partial x} = yx^{y-1} + 2^{xy} y \ln(2)$$
$$\frac{\partial z}{\partial y} = x^y \ln(x) + 2^{xy} x \ln(2)$$

Higher-Order Partials

Notice that given a function of two variables, $f(x, y)$, both of our derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are again functions of x and y, so we can take the partial derivatives again. If $f(x, y) = x^3y^2 + \cos(xy)$, then we know

$$\frac{\partial f}{\partial x} = 3x^2y^2 - y\sin(xy)$$

We could then differentiate this derivative once again:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 6xy^2 - y^2\cos(xy).$$

This is called a *second-order partial derivative* of $f(x, y)$, and is denoted

$$\frac{\partial^2 f}{\partial x^2} \quad \text{or} \quad f_{xx}.$$

We didn't just have to take the partial with respect to x there: we could have taken the partial with respect to y of $\frac{\partial f}{\partial x}$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 6x^2y - xy\cos(xy)$$

This is another second-order derivative and is denoted

$$\frac{\partial^2 f}{\partial y\,\partial x} \quad \text{or} \quad f_{xy}$$

We could of course keep doing this, calculating partial derivatives of our partial derivatives, and the notation extends the way you would expect:

$$\frac{\partial^2 f}{\partial y^2} = f_{yy}$$
$$\frac{\partial^2 f}{\partial x\,\partial y} = f_{yx}$$
$$\frac{\partial^3 f}{\partial x^3} = f_{xxx}$$
$$\frac{\partial^3 f}{\partial x^2\,\partial y} = f_{yxx}$$
$$\frac{\partial^3 f}{\partial x\,\partial y\,\partial x} = f_{xyx}$$
$$\frac{\partial^3}{\partial x\,\partial y^2} = f_{yyx}$$
$$\vdots$$

When we take three partial derivatives, as in the case of $\frac{\partial f}{\partial x^2\,\partial y}$ where we first differentiate with respect to y and then differentiate with respect to x two more times, we've taken a *third-order* partial derivative. If we calculate four partial derivatives, we've taken a *fourth-order* partial derivative, and so on.

Any of these *higher-order* partial derivatives where we differentiate with respect to different variables is called a *mixed partial derivative*. The second-order mixed partials are

$$\frac{\partial}{\partial x\,\partial y} \quad \text{and} \quad \frac{\partial}{\partial y\,\partial x}.$$

Some of the third-order mixed partials are

$$\frac{\partial}{\partial x\,\partial y\,\partial x}, \quad \frac{\partial}{\partial x^2\,\partial y}, \quad \text{and} \quad \frac{\partial}{\partial y \partial x^2}.$$

A reasonable question to ask would be how these mixed partial derivatives are related: Is there any relationship between f_{xy} and f_{yx}? This is answered by *Clairaut's theorem*.

Theorem 3.4 (Clairaut's Theorem)**.**
If f is defined in a neighborhood of (x_0, y_0) and if f_{xy} and f_{yx} are both de-

fined and continuous in this neighborhood, then $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$.

Corollary 3.5.
Under all of the assumptions of Clairaut's theorem, we can move the order of differentiation around to our heart's content for any n*-th order partial derivative:*

$$\frac{\partial^4 f}{\partial x^2\, \partial y^2} = \frac{\partial^4 f}{\partial y \partial x^2 \partial y} = \frac{\partial^4 f}{\partial x \partial y \partial x \partial y} = \cdots .$$

The assumptions in Clairaut's theorem will be satisfied for most of the functions we care about, and so for practical purposes, in this class $f_{xy} = f_{yx}$. However, in general, it's possible to find functions that *do not* satisfy the assumptions of Clairaut's theorem, and in such situations there's no guarantee that the partial derivatives agree!

Functions of More Than Two Variables

This notion partially differentiating a function of two variables naturally extends to partial derivatives of functions of any number of variables. For a function of three variables, $f(x, y, z)$, for instance, we can define

$$\frac{\partial f}{\partial z} = \lim_{h \to 0} \frac{f(x, y, z + h) - f(x, y, z)}{h}.$$

To calculate $\frac{\partial f}{\partial z}$ we differentiate like normal, but now pretend that both x and y are constants.

Example 3.16.
Calculate $\frac{\partial f}{\partial z}$ of $f(x, y, z) = x^2 y^3 z^4$.

$$\frac{\partial f}{\partial z} = 4x^2 y^3 z^3.$$

This same idea works for any (finite!) number of variables. If we have with twenty-six variables, given by the twenty-six letters of the English alphabet, to differentiate with respect to any one, we treat the other twenty-five as constant:

$$\begin{aligned}\frac{\partial^2}{\partial \ell \, \partial u}\left(a^b c - \ell^5 m u - u^{\sin(x)}\right) =& \frac{\partial}{\partial \ell}\left(-\ell^5 m - \sin(x) u^{\sin(x)-1}\right) \\ =& -5\ell^4 m\end{aligned}$$

3.4 Directional derivatives

Directional derivatives can be thought of as generalizations of partial derivatives which describe how a multivariable function changes as you change the inputs to the function by walking along a line in the xy-plane. One important application of directional derivatives appears in an algorithm called the *simplex method* which is used to solve some special types of optimization problems.

Generalizing Partial Derivatives

When we discussed partial derivatives, we said that the partial derivative of a function $f(x, y)$ with respect to x (or y) represents the instantaneous change in the function as we change the x-value (y-value), leaving y (x) constant. Geometrically, we can think of this as intersecting the surface $z = f(x, y)$ with the plane $y = c$. This intersection is a curve, and the derivative of this curve corresponds to the partial derivative with respect to x.

There's nothing particularly special about the planes $x = c$ and $y = c$, however. We could take any plane $ax + by = c$ (this corresponds to a line in the xy-plane), intersect it with a surface $z = f(x, y)$, and obtain a curve which we can differentiate. The derivative of this curve is called a *directional derivative*, and it tells us how the function $f(x, y)$ as we change both the x- and y-inputs by moving in a given direction.

Motivating Example

Consider the surface $z = \cos(x)/(1 + y^2)$. If we choose a point in the xy-plane, say $(x_0, y_0) = (1, 1)$, then we get a point on the surface, $(1, 1, \cos(1)/2)$. Starting from this point, let's start changing the value (x, y) that we plug into the function by moving from this point in the direction $\langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$.

As we move from the original input $(1, 1)$ in direction $\langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$ (so we're plugging in points of the form $(1 + t/\sqrt{2}, 1 + t/\sqrt{2})$), we get a curve on the surface. This curve is the intersection of the surface $z = \cos(x)/(1+y^2)$ and the plane $y = x$. (The point $(1, 1)$ and vector $\langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$ in the xy-plane gives us the line $y = x$. In 3-space this is a plane.) The intersection of these surfaces, and corresponding curve, are given in Figure 3.2.

If we differentiate this curve, the derivative represents the change in the elevation of the surface as we walk along the curve.

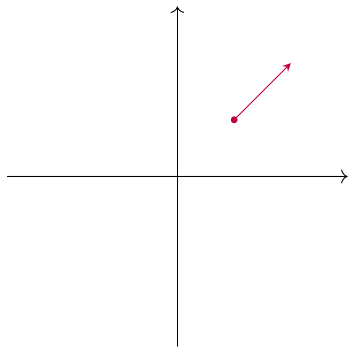


Figure 3.1: Changing inputs to the function.

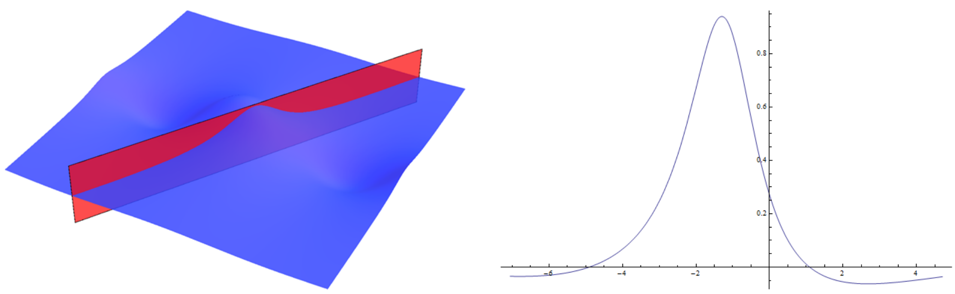


Figure 3.2: The intersection of our surface and a plane produces a curve.

Changing the line we're interested in – saying we walk in the direction $\langle \cos(15°), \sin(15°) \rangle$ from our point, we'd get a different line in the xy-plane, a different plane intersecting our surface, and a different curve representing the elevation of the surface along that curve. See Figure 3.3.

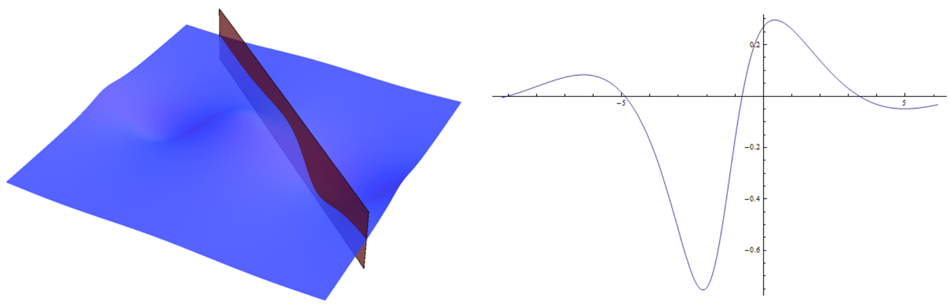


Figure 3.3: Changing directions gives a different curve.

We'd like to have a formula that tells us precisely how the surface's elevation changes as we walk in a given direction. This is given by directional derivatives.

Directional Derivatives

Let's suppose that with the function $f(x, y)$, we take the input (x_0, y_0) and change it by moving in direction $\vec{v} = \langle v_1, v_2 \rangle$. To make life easier, let's suppose that $\vec{v}$ is a unit vector. If we move h-units of distance away from the point (x_0, y_0) in direction $\langle v_1, v_2 \rangle$, then our new input point is $(x_0 + hv_1, y_0 + hv_2)$. The average change in our function is then

$$\frac{f(x_0 + hv_1, y_0 + hv_2) - f(x_0, y_0)}{h}.$$

To calculate the instantaneous change we take the limit as $h \to 0$. This is called the *direction derivative of* $f(x, y)$ *at* (x_0, y_0) *in the direction of* $\vec{v}$ and is denoted $D_{\vec{v}} f(x_0, y_0)$:

$$D_{\vec{v}} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + hv_1, y_0 + hv_2) - f(x_0, y_0)}{h}.$$

Notice that this value will depend on the (x_0, y_0) we choose. If we allow this original input to change, we have a function of x and y:

$$D_{\vec{v}} f(x, y) = \lim_{h \to 0} \frac{f(x + hv_1, y + hv_2) - f(x, y)}{h}.$$

Example 3.17.
Find the directional derivative of $f(x, y) = x \sin(y)$ in the direction of $\vec{v} = \langle 2, -1 \rangle$.

We want to write this as a limit, but let's first notice that $\langle 2, -1 \rangle$ *is not* a unit vector: it has length $\sqrt{5}$. For our formula above to work we need to divide by the distance between our points. We could modify the equation above to take this into account, but it's easier to just replace $\vec{v}$ with a unit vector that points in the same direction. Let's call this vector $\vec{u}$:

$$\vec{u} = \frac{1}{\|\vec{v}\|} \vec{v} = \frac{1}{\sqrt{5}} \langle 2, -1 \rangle = \left\langle 2/\sqrt{5}, -1/\sqrt{5} \right\rangle.$$

Now we can use our above formula for the directional derivative:

$$\begin{aligned}
D_{\vec{u}}f(x,y) &= \lim_{h\to 0}\frac{f(x+2h/\sqrt{5}, y-h/\sqrt{5}) - f(x,y)}{h}\\
&= \lim_{h\to 0}\frac{(x+2h/\sqrt{5})\sin(y-h/\sqrt{5}) - x\sin(y)}{h}\\
&= \lim_{h\to 0}\frac{x\sin(y-h/\sqrt{5}) + 2h\sin(y-h/\sqrt{5})/\sqrt{5} - x\sin(y)}{h}\\
&= \lim_{h\to 0}\left(\frac{x\sin(y-h/\sqrt{5}) - x\sin(y)}{h} + \frac{2h\sin(y-h/\sqrt{5})}{h\sqrt{5}}\right)\\
&= \lim_{h\to 0}\frac{x\sin(y-h/\sqrt{5}) - x\sin(y)}{h} + \lim_{h\to 0}\frac{2h\sin(y-h/\sqrt{5})}{h\sqrt{5}}\\
&= \lim_{h\to 0}\frac{x\sin(y-h/\sqrt{5}) - x\sin(y)}{h} + \lim_{h\to 0}\frac{2\sin(y-h/\sqrt{5})}{\sqrt{5}}\\
&\overset{\mathscr{L}}{=} \lim_{h\to 0}\frac{x\cos(y-h/\sqrt{5})\frac{-1}{5} - 0}{1} + \lim_{h\to 0}\frac{2\sin(y-h/\sqrt{5})}{\sqrt{5}}\\
&= \frac{2\sin(y) - x\cos(y)}{\sqrt{5}}
\end{aligned}$$

Since $\vec{v}$ and $\vec{u}$ point in the same direction, the directional derivatives are the same:

$$D_{\vec{v}}f(x,y) = D_{\vec{u}}f(x,y).$$

One thing to note about the above example is that we since our original vector wasn't a unit vector, we replaced it with a unit vector to do the calculations.

Using the Gradient

Even though the limit definition of the directional derivative should make intuitive sense, it's very difficult to compute as a limit. What w would like to do, then, is to relate the directional derivative to some easier-to-compute quantities, such as partial derivatives. To make sense of what we're about to do, let's first notice that any two-dimensional vector $\vec{v} = \langle v_1, v_2\rangle$ can be written as a sum of horizontal and vertical unit vectors:

$$\langle v_1, v_2\rangle = v_1\langle 1, 0\rangle + v_2\langle 0, 1\rangle.$$

Notice too that the directional derivatives of $f(x, y)$ in direction $\langle 1, 0\rangle$ and $\langle 0, 1\rangle$ are exactly the partial derivatives, f_x and f_y. Perhaps, then, we can write write the directional derivative of f in direction $\vec{v}$ in terms of these directional derivatives. It will turn out that we can compute directional derivatives using only information about the partial derivatives f_x and f_y together with the components of the vector pointing in the direction of our directional derivative, and in order to do this we need to introduce the "gradient" of a function.

Given a differentiable function $f(x, y)$, the *gradient* of $f(x, y)$ is the (multivariable) vector-valued function whose components are the partial derivatives of $f(x, y)$. This function is denoted $\nabla f(x, y)$ and is sometimes pronounced "del of $f(x, y)$":

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle .$$

Example 3.18.
Calculate the gradient of $f(x, y) = 3xy^2 - \sin(x) + y^x$.

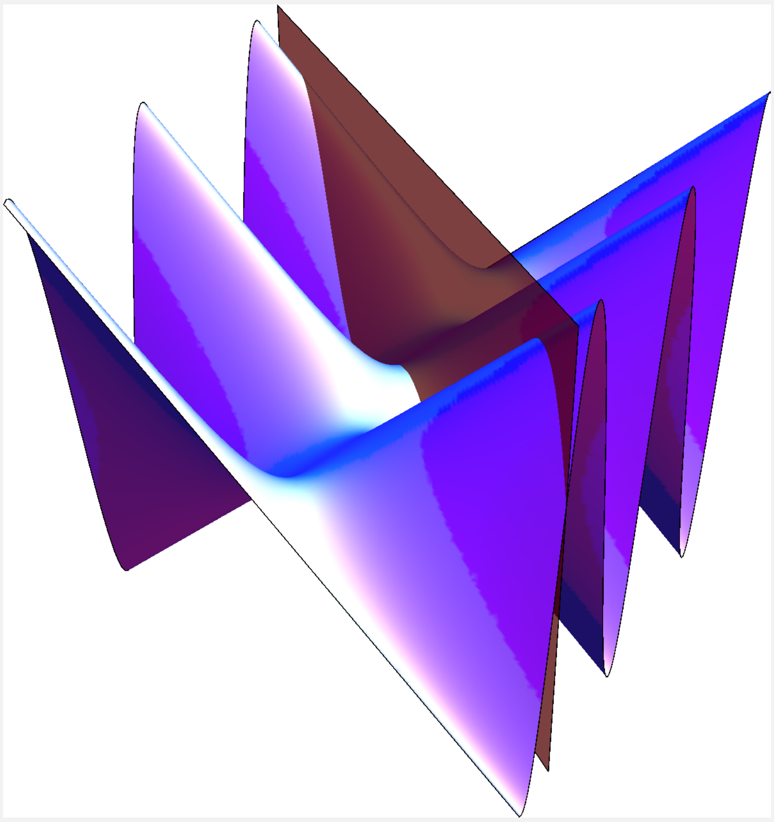
$$\begin{aligned}\nabla f(x, y) &= \left\langle \frac{\partial}{\partial x}\left(3xy^2 - \sin(x) + y^x\right), \frac{\partial}{\partial y}\left(3xy^2 - \sin(x) + y^x\right)\right\rangle \\ &= \left\langle 3y^2 - \cos(x) - y^x \ln(y), 6xy - xy^{x-1}\right\rangle .\end{aligned}$$

We'll see the gradient several times through the semester, particularly later when we discuss optimization problems. Right now we care about the gradient because of the following theorem which makes it easy for us compute directional derivatives.

Theorem 3.6.
If $\vec{u}$ is a unit vector and $f(x, y)$ is differentiable, then

$$D_{\vec{u}} f(x, y) = \nabla f(x, y) \cdot \vec{u}.$$

Example 3.19.
Find the directional derivative of $f(x,y) = x\sin(y)$ in the direction of $\vec{v} = \langle 2, -1\rangle$ using the above theorem.



To apply the theorem we need to replace $\vec{v}$ with a unit vector pointing in the same direction, $\vec{u} = \langle 2/\sqrt{5}, -1/\sqrt{5}\rangle$. We also need to calculate the gradient,

$$\nabla f(x,y) = \langle \sin(y), x\cos(y)\rangle .$$

Now the theorem tells us

$$\begin{aligned} D_{\vec{u}} f(x,y) =& \nabla f(x,y)\cdot\vec{u} \\ =& \langle \sin(y), x\cos(y)\rangle \cdot \left\langle 2/\sqrt{5}, -1/\sqrt{5}\right\rangle \\ =& \frac{2\sin(y) - x\cos(y)}{\sqrt{5}} \end{aligned}$$

Example 3.20.
Calculate $D_{\langle 3,-2\rangle}\left(x^2+\ln(\sec(xy))\right)$.

Note our vector does not have unit length, so we need to replace it with the vector

$$\vec{u}=\left\langle 3/\sqrt{10},-2/\sqrt{10}\right\rangle.$$

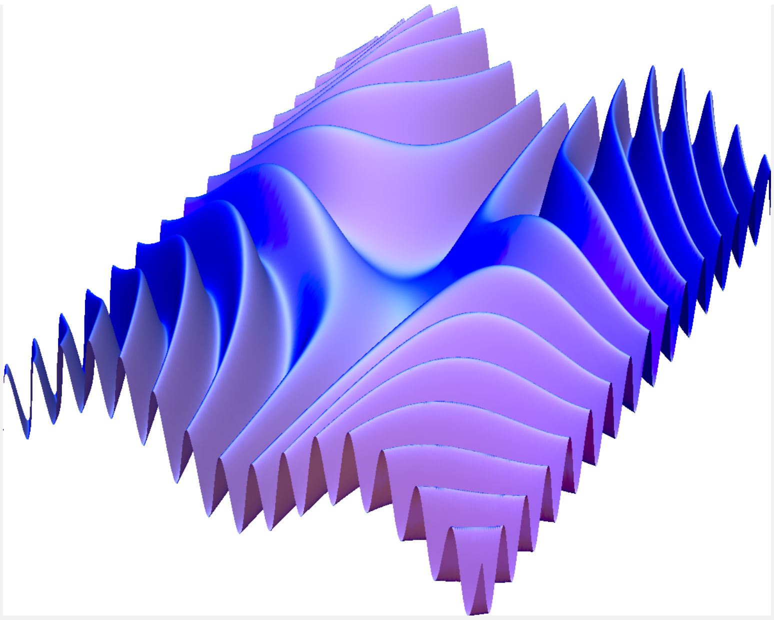
Now we need to calculate the gradient,

$$\begin{aligned}\nabla\left(x^2+\ln(\sec(xy))\right)&=\left\langle 2x-\frac{y\tan^2(xy)}{\sec(xy)},\frac{x\tan^2(xy)}{\sec(xy)}\right\rangle\\&=\langle 2x-y\sin(xy)\tan(xy),x\sin(xy)\tan(xy)\rangle\end{aligned}$$

Now we calculate

$$\begin{aligned}&D_{\langle 3,-2\rangle}\left(x^2+\ln(\sec(xy))\right)\\=&D_{\left\langle 3/\sqrt{10},-2/\sqrt{10}\right\rangle}\left(x^2+\ln(\sec(xy))\right)\\=&\langle 2x-y\sin(xy)\tan(xy),x\sin(xy)\tan(xy)\rangle\cdot\left\langle 3/\sqrt{10},-2/\sqrt{10}\right\rangle\\=&\frac{6x-3y\sin(xy)\tan(xy)-2x\sin(xy)\tan(xy)}{10}\\=&\frac{6x-\sin(xy)\tan(xy)(2x+3y)}{10}\end{aligned}$$

Example 3.21.
Suppose the elevation a mountain range at a point (x,y) is given by $\sin(xy)-\cos(x)$. What is the instantaneous rate of change in the elevation as you walk from the point $(\pi,1/2)$ in the North-East direction (this correspond to the direction $\left\langle 1/\sqrt{2},1/\sqrt{2}\right\rangle$)?



This is given by the directional derivative,

$$
\begin{aligned}
&D_{\langle 1/\sqrt{2}, 1/\sqrt{2}\rangle}\left(\sin(xy) - \cos(x)\right) \\
&= \langle y\cos(xy) - \sin(x), x\cos(xy)\rangle \cdot \left\langle 1/\sqrt{2}, 1/\sqrt{2}\right\rangle \\
&= \frac{\cos(xy)(y-x) - \sin(x)}{\sqrt{2}}
\end{aligned}
$$

Evaluating this at $(\pi, 1/2)$ tells us that the change in elevation is $\left(\cos(\pi/2)(\pi - 1/2) - \sin(\pi)\right)/\sqrt{2} = 0$.

The Simplex Method (Optional)

The simplex method is an algorithm that is used to solve linear optimization problems subject to linear constraints. The details of the algorithm require some linear algebra, so we can't give a precise version of the algorithm here, but we can give the basic idea which applies directional derivatives.

A *linear optimization problem* is a maximization or minimization problem of a linear function: a function of the form $a_1x_1 + a_2x_2 + \cdots a_nx_n$. In three variables, this means we'd like to maximize or minimize a function $ax + by + cz$. Saying the problem is subject to linear inequalities means that we have constraints of the form $d_1x_1 + \cdots + d_nx_n \leq k$. We could potentially have lots and lots of these constraints.

Since these constraints are linear inequalities, each constraint removes half of the possible points from consideration (half the plane in the case of two variables, half of 3-space in the case of three variables, etc.). The collection of all points satisfying all of the constraints forms a geometric object called a *simplex*. This simplex of possible points is called the *feasible region* of the problem.

Since the *objective function* we wish to minimize or maximize has is linear, different values of the objective function give parallel *hyperplanes*. Our goal is to find the hyperplane whose corresponding value is as big (or as small) as possible, and still intersects the feasible region. It's not difficult to see that such a maximum or minimum, if it exists, must occur at a corner of the simplex. This suggests an algorithm for determining the local of the extrema by traversing the corners of the feasible region.

Typically, if there are very many constraints, the feasible region will have very many corners. For computational reasons, we'd like to look at as few corners as possible. (The more corners you check, the more work you have to do, and this can take a very long time when you're checking millions of corners.) Starting from any corner, we will move to another corner by calculating the directional derivative of our objective function, in the direction of each adjacent corner of the feasible region.

If any of these directional derivatives is positive, then the function increases as we move in that direction. Of all the directions with positive directional derivative, pick the one which is largest (the one with the largest increase). Once we know which direction to move in, we move in that direction until we arrive at another corner of the feasible region, and then repeat the process until none of the directional derivatives is positive. When this happens, we have arrived at the corner giving the maximum value of the objective function.

3.5 The total derivative and the chain rule

In the previous sections we introduced the idea of partial and directional derivatives, we saw that computing directional derivatives was most easily accomplished by first computing a function's gradient. At the time the gradient was introduced simply as a convenient tool for doing these calculations. In this section we will see that the gradient is closely related to what is sometimes called the *total derivative* of a function.

Derivatives as linear transformations

In your first semester of calculus you learned that a function of one variable $f(x)$ was ***differentiable at a point*** a if the following limit existed,

$$f'(a) := \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

You also saw that this value, if it exists, can be interpreted as the slope of the tangent line to the graph $y = f(x)$ at the point a. Perhaps most importantly, though, is that $f'(a)$ is closely related to the idea of constructing a linear approximation to $f(x)$ near a. Writing out the equation of the tangent line to $y = f(x)$ at $(a, f(a))$ we have

$$y - f(a) = f'(a)(x - a)$$

which we may rewrite as

$$y = f'(a)(x - a) + f(a).$$

That is, $f'(a)$ tells us some information about how the output of our function should change as we make a small change to the inputs near $x \approx a$. Another way of saying this is that $f'(a)$ tells us the linear transformation that best approximates changes in $f(x)$ near a.

Recall that a linear transformation is a map $T : \mathbb{R}^n \to \mathbb{R}^m$ which is determined by an $m \times n$ matrix. In the special case $m = n = 1$ (so, we convert single real numbers into single real numbers), a linear transformation is given by a 1×1 matrix which is simply a single number. That is, all linear transformations $T : \mathbb{R} \to \mathbb{R}$ simply have the form $T(x) = mx$. As a consequence, notice that a linear transformation is required to map 0 to 0: $T(0) = 0$ regardless of what the value of m that appears in $T(x) = mx$ actually is.

What's *really* happening when you're calculating the derivative of a function is you're finding the best linear transformation that approximates the change in that function. That is, if $f'(a) = m$, then by the calc. 1 definition of a derivative this means

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = m.$$

We could rewrite this as

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - mh}{h} = 0.$$

This is something we can more easily generalize to higher dimensions. For reasons that will be clearer later (e.g., when we consider vector-valued functions of multiple variables), we will want to rewrite this as

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - mh}{|h|} = 0.$$

We should interpret the mh that appears in the numerator of the expression above as our linear transformation $T(x) = mx$ applied to h, so we could further rewrite this as

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - T(h)}{|h|} = 0.$$

Though this looks a little bit different than your limit definition of the derivative from first semester calculus, it is equivalent, and has the advantage of being easier to generalize to higher dimensions. To complete the generalization, though, we need to adopt one more convention.

For a function of multiple variables such as $f(x, y)$, we may sometimes wish to interpret the input as a vector. That is, if we have a two-dimensional vector $\vec{v} = \langle x, y\rangle$, then we will sometimes write $f(\vec{v})$ to mean plug the components of the vector into our function. Just to have a concrete example, if $f(x, y) = x^3y + y^2$ and $\vec{v} = \langle -1, 3\rangle$, then we would have $f(\vec{v}) = (-1)^3 \cdot 3 + 3^2 = 6$.

Thinking of the input to a multivariable function as a vector is convenient because we can add vectors together and so an expression like $f(\vec{v} + \vec{h})$ is meaningful.

In general for a function $f : \mathbb{R}^n \to \mathbb{R}^m$ we define the ***(total) derivative*** of f at a point $\vec{a}$ to be the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ satisfying the following equation

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - T(\vec{h})}{\|\vec{h}\|} = \vec{0},$$

providing such a linear transformation exists. If this linear transformation exists we say the function is ***differentiable*** at $\vec{a}$ and we write $Df(\vec{a})$ for that linear transformation.

Notice that we are taking the magnitude of $\vec{h}$ (which represents the change in input) in the denominator above, similarly to how we divided by $|h|$ in our rewritten calculus 1 derivative above. We need to do that here since $\vec{h}$ is a vector and we can't divide by a vector but *we can* divide by its magnitude since that's just a scalar.

This definition of a derivative may seem strange, but the idea is really that we're trying to generalize what you did in calculus one (which was really approximating a non-linear function with a linear one) to multi-variable functions. Let's notice a few things about this definition. First is that, unlike the case of partial derivatives and directional derivatives, this definition considers *all* directions in which the input to the function changes. This is simply because for a limit of a multivariable function to exist, we know that you must consider all directions. Let's also notice that this means that the our derivative is now a matrix instead of a single number.

Of course, the question now is how do we actually compute this matrix $Df(\vec{a})$? For the moment we will only consider the case when f gives a single real number as an output; i.e., we will only be concerned with functions $f : \mathbb{R}^n \to \mathbb{R}$ right now. Later we will consider more general vector-valued functions of multiple variables.

Let's notice that if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a}$, then the derivative $Df(\vec{a})$ will be an $n \times 1$ matrix. That is, it's just a list of n values written as a single "row" of numbers (sometimes such quantities are called ***row vectors***). The following theorems tie our notions of partial derivatives, directional derivatives, and the total derivative together.

Theorem 3.7.
If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a}$, then its derivative is equal to

$$Df(\vec{a}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{a}) & \frac{\partial f}{\partial x_2}(\vec{a}) & \cdots \frac{\partial f}{\partial x_n}(\vec{a}) \end{pmatrix}$$

That is, if f is differentiable at $\vec{a}$, then the total derivative simply consists of the component of the gradient of f evaluated at $\vec{a}$.

Example 3.22.
Assuming the function

$$f(x, y) = x^2y - 3x + y^3$$

is differentiable at $(1, 5)$, compute its derivative.

We simply compute the partial derivatives of the function,

$$\frac{\partial f}{\partial x} = 2xy - 3$$
$$\frac{\partial f}{\partial y} = x^2 + 3y^2$$

and evaluate these at our given point to obtain the values

$$\frac{\partial f}{\partial x}(1,5) = 7$$
$$\frac{\partial f}{\partial y}(1,5) = 16$$

Thus the derivative of f at $(1,5)$ is

$$Df(1,5) = \begin{pmatrix} 7 & 16 \end{pmatrix}.$$

Notice that for a functions $f : \mathbb{R}^n \to \mathbb{R}$, the derivative at $\vec{a}$, $Df(\vec{a})$, evaluated at a unit vector $\vec{u}$ is precisely the same thing as the directional derivative of f in the direction of $\vec{u}$ evaluated at $\vec{a}$.

Example 3.23.
Compute the total derivative of $f(x,y) = x^2y - 3x + y^3$ at $(1,5)$ and evaluate this at the vector $\vec{u} = \langle \sqrt{3}/2, 1/2 \rangle$. Compare this to the directional derivative $D_{\vec{u}}f(1,5)$.

For the first part we simply evaluate the linear transformation described in Example 3.22 at the given vector $\vec{u}$, which means we just do matrix multiplication. For this matrix multiplication to be defined, though, notice we need to write $\vec{u}$ as a column vector:

$$\begin{pmatrix} 7 & 16 \end{pmatrix} \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} = 7\frac{\sqrt{3}}{2} + 16\frac{1}{2} = \frac{7\sqrt{3} + 16}{2}.$$

For the second part we compute the gradient,

$$\nabla f(x,y) = \left\langle 2xy - 3,\ x^2 + 3y^2 \right\rangle,$$

evaluate this at $(1,5)$,

$$\nabla f(1,5) = \langle 7, 16 \rangle ,$$

and take the dot product with the given vector,

$$\langle 7, 16 \rangle \cdot \langle \sqrt{3}/2, 1/2 \rangle = 7\frac{\sqrt{3}}{2} + 16\frac{1}{2} = \frac{7\sqrt{3}+16}{2}.$$

As the above example shows, directional derivatives and total derivatives are very closely related, *but they're not quite the same thing*. In particular, our definition of directional derivative required that we convert our vectors to unit vectors before doing any calculations, but that's not required for the total derivative. In particular, the total derivative will give us multiples of the directional derivative in the corresponding direction, the multiple being the same as the multiple of the unit vector pointing in the same direction.

Example 3.24.
Compute the total derivative of $f(x,y) = x^2y - 3x + y^3$ at $(1,5)$ and evaluate this at the vector $\vec{v} = \left\langle \sqrt{3}, 1 \right\rangle$.

We simply compute

$$\begin{pmatrix} 7 & 16 \end{pmatrix} \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} = 7\sqrt{3} + 16$$

Notice this quantity is twice the value computed in Example 3.23 and the vector given in this problem is twice the vector that was given in that example.

Any time we are talking about limits, we have to worry about whether the limit exists or not. In particular, it is possible that no linear transformation T satisfies the equation involving limits in our definition of the derivative. If this were to happen then our function would not be differentiable.

Example 3.25.
The following function is not differentiable at $(0,0)$:

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

To see that this function can not be differentiable at $(0,0)$, we need to show that there is no linear transformation $T : \mathbb{R}^2 \to \mathbb{R}$ satisfying

$$\lim_{\vec{h} \to \vec{0}} \frac{f(\vec{0}+\vec{h}) - f(\vec{0}) - T(\vec{h})}{\|\vec{h}\|} = 0.$$

To see no such T exists, let's suppose that we had an arbitrary linear transformation determined by the matrix $\begin{pmatrix} a & b \end{pmatrix}$. The limit above would become

$$\lim_{(h_x,h_y)\to(0,0)} \frac{\frac{h_x h_y}{h_x^2+h_y^2} - 0 - ah_x - bh_y}{\sqrt{h_x^2+h_y^2}}.$$

Consider taking this limit along the x axis by setting h_y equal to zero. We would then have

$$\lim_{h_x \to 0} \frac{0 - 0 - ah_x - 0}{\sqrt{h_x^2}} = \lim_{h_x \to 0} -a\frac{h_x}{|h_x|}$$

and this limit does not exist (taking the limit as h_x approaches 0 from the right would give us $-a$, and taking the limit as h_x approaches 0 from the left would give a).

Notice that even though the function appearing in Example 3.25 is not differentiable, the partial derivatives of the function exist: they are are

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{-x^2y+y^3}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{x^3-xy^2}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

That is *existence of partial derivatives does not imply a function is differentiable!* Let's notice, however, that even though these partial derivatives exist, they are discontinuous. For example, to compute the limit as $(x, y) \to (0, 0)$ of $\frac{\partial f}{\partial x}$ we may consider the limit along the y-axis, setting $x = 0$ to obtain

$$\lim_{y \to 0} \frac{y^3}{y^4} = \lim_{y \to 0} \frac{1}{y}.$$

This limit does not exist, however, as the limit from the left-hand side $(y \to 0^-)$ gives $-\infty$ and the limit from the right-hand side $(y \to 0^+)$ gives ∞. It turns out that this is the only thing that can go wrong when comparing differentiability to the existence of partial derivatives.

Theorem 3.8.
A function $f : \mathbb{R}^2 \to \mathbb{R}$ will be differentiable at (x_0, y_0) if and only if the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at (x_0, y_0).

Chain rule for composition with a single-variable function

Before we dive into the chain rule for multivariable functions, let's quickly recall the chain rule for functions of a single variable. If $f, g : \mathbb{R} \to \mathbb{R}$ are two differentiable functions, the chain rule for functions of a single variable tells us

$$\frac{d}{dt} f(g(t)) = f'(g(t))\, g'(t).$$

If we write $y = f(x)$ and $x = g(t)$, then we have $y = f(x) = f(g(t))$, and so we can differentiate y with respect to t. The chain rule can then be written as

$$\frac{dy}{dt} = \frac{df}{dx}\frac{dx}{dt}.$$

If this looks a little bit strange, just realize that $\frac{df}{dx} = f'(x)$ and since $x = g(t)$, we have $\frac{df}{dx} = f'(x) = f'(g(t))$, and similarly $\frac{dx}{dt} = g'(t)$ since $x = g(t)$. So this really is the "usual" chain rule, just written in a way that might look a little bit foreign.

Now let's change things a little bit, let's suppose that f is a function of two variables, $f : \mathbb{R}^2 \to R$. But let's also suppose that instead of plugging in (x, y) values into f, we plug in two different functions. Let's say that we take $x = g(t)$ and $y = h(t)$ and plug that into our function. This gives us a function of a single variable,

$$z(t) = f(x(t), y(t)).$$

Since this is a function of a single variable, we ought to be able to calculate its derivative, $\frac{dz}{dt}$. However, we have this multivariable function $f(x, y)$ involved. How are the partial derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, related to the derivative $\frac{dz}{dt}$? The answer to this question is provided by the multivariable version of the chain rule.

Theorem 3.9 (Multivariable chain rule, version 1)**.**
Suppose that $f(x, y)$ is a multivariable function whose partial derivatives exist and are continuous. Suppose also that $x = g(t), y = h(t)$ are two single variable functions whose derivatives exist and are continuous as well. The single variable function

$$z(t) = f(g(t), y(t))$$

has the following derivative:

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Example 3.26.
Suppose that $z = f(x, y) = x^2 + y^2 + xy$ where $x = \sin(t)$ and $y = e^t$. Calculate $\frac{dz}{dt}$.

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= (2x + y)\cos(t) + (2y + x)e^t \\ &= (2\sin(t) + e^t)\cos(t) + (2e^t + \sin(t))e^t.\end{aligned}$$

Note that we could have first also replaced x and y with $\sin(t)$ and e^t, respectively, and then calculated the derivative:

$$\frac{dz}{dt} = \frac{d}{dt}\left(\sin^2(t) + e^{2t} + \sin(t)e^t\right).$$

This would give us the same derivative, but the advantage of the chain it lets us easily calculate the derivative, possibly shortcutting any necessary algebraic simplifications. In particular, if we repeated the above example keeping $f(x, y) = x^2 + y^2 + xy$, but changing x and y to some other functions, most of the work is already done. All we'd have to do is change our final answer.

Example 3.27.
Suppose that $z = f(x, y) = x^2 + y^2 + xy$ where $x = \sqrt{t}$ and $y = \tan(t)$. Calculate $\frac{dz}{dt}$.

$$\begin{aligned}\frac{dz}{dt} &= (2x + y)\frac{1}{2\sqrt{t}} + (2y + x)\sec^2(t) \\ &= (2\sqrt{t} + \tan(t))\frac{1}{2\sqrt{t}} + (2\tan(t) + \sqrt{t})\sec^2(t).\end{aligned}$$

Chain rule for composition with a multivariable function

Now suppose that we have $z = f(x, y)$, but instead of replacing x and y with functions of a single variable, we replace them with functions of two variables. Say $x = g(s, t)$ and $y = h(s, t)$. How z is a function of two variables,

$$z(s, t) = f(g(s, t), h(s, t)).$$

Since z is a function of two variables, we can discuss the partial derivatives of z with respect to either s or t. What does the chain rule look like in this situation?

If we wanted to calculate $\frac{\partial z}{\partial t}$, we'd differentiate like normal, treating s as a constant. This means that our chain rule above still applies, but we'd replace $\frac{dx}{dt}$ and $\frac{dy}{dt}$ with $\frac{\partial x}{\partial t}$ and $\frac{\partial y}{\partial t}$, respectively. We'd perform similar replacements if we wanted to calculate $\frac{\partial z}{\partial s}$, giving us the following version of the chain rule.

Theorem 3.10 (Multivariable chain rule, version 2).
Suppose that $f(x, y)$ is a multivariable function whose partial derivatives exist and are continuous. Suppose also that $x = g(s, t), y = h(s, t)$ are two multivariable functions whose partial derivatives exist and are continuous as well. The multivariable function

$$z(s, t) = f(g(s, t), y(s, t))$$

has the following partial derivatives:

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

$$\frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

Example 3.28.
Suppose $z = f(\theta, \varphi) = \sin(\theta)\cos(\varphi)$ where we replace $\theta = st^2$ and $\varphi = t$. Calculate both of the first-order partial derivatives of z with respect to s and t.

$$\begin{aligned}\frac{\partial z}{\partial s} &= \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial s} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial s} \\ &= (\cos(\theta)\cos(\varphi)) \cdot t^2 - (\sin(\theta)\sin(\varphi)) \cdot 0 \\ &= \cos\left(st^2\right)\cos(t)t^2\end{aligned}$$

$$\begin{aligned}\frac{\partial z}{\partial t} &= \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial t} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial t} \\ &= (\cos(\theta)\cos(\varphi)) \cdot 2st - (\sin(\theta)\sin(\varphi)) \cdot 1 \\ &= 2st\cos\left(st^2\right)\cos(t) - \sin\left(st^2\right)\sin(t)\end{aligned}$$

Implicit Differentiation

We can use these ideas to reinterpret implicit differentiation for functions of a single variable.

Suppose that we have a curve in $\mathbb{R}^2$ given by the equation $y^2 = x^3 - x + 1$. See Figure 3.4.

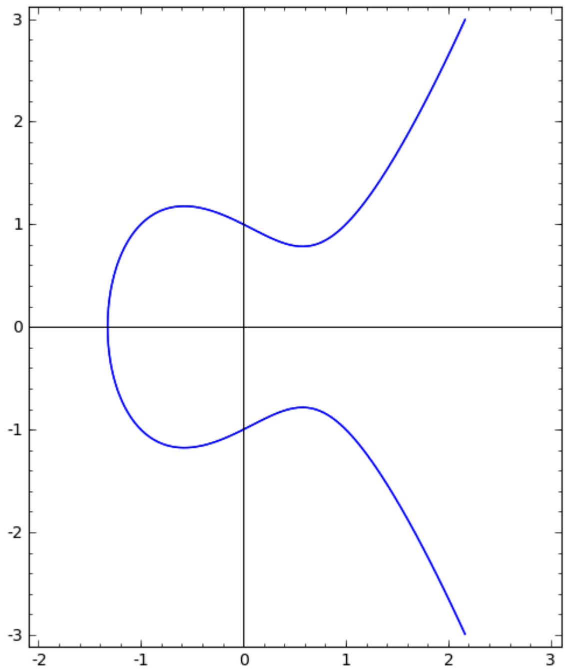


Figure 3.4: The curve determined by $y^2 = x^3 - x + 1$

Notice that this curve isn't the graph of a function of x; you can't solve the above equation for y (you'd get two possibilities for each y, a positive and a negative). However, it still makes sense to talk about slopes of lines tangent to this curve: we can still make sense of $\frac{dy}{dx}$, even though y isn't a function of x!

You of course already know how to do this with implicit differentiation from your first-semester calculus courses, but we can also view this as a multivariable problem. To do this, let's first rewrite our equation

$$y^2 = x^3 - x + 1$$

by putting everything on one side of the equation:

$$y^2 - x^3 + x - 1 = 0.$$

Now consider the function $F(x,y) = y^2 - x^3 + x - 1$. Notice that our curve is the set of (x,y) points satisfying $F(x,y) = 0$.

Now, since we want to calculate $\frac{dy}{dx}$, let's suppose that y is a function of x – just call this $y(x)$. Then

$$\begin{aligned}
& y^2 - x^3 + x - 1 \\
\implies & F(x,y) = 0 \\
\implies & \frac{d}{dx}F(x,y) = \frac{d}{dx}0 = 0 \\
\implies & \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0 \\
\implies & F_x + F_y\frac{dy}{dx} = 0 \implies \qquad \frac{dy}{dx} = -\frac{F_x}{F_y}.
\end{aligned}$$

In our particular example, $F_x = -3x^2 + 1$ and $F_y = 2y$, hence

$$\frac{dy}{dx} = \frac{3x^2 - 1}{2y}.$$

This means the slope of the tangent line to $y^2 = x^3 - x + 1$ at the point $(2, \sqrt{7})$ on the curve is $\frac{5}{2\sqrt{7}}$.

3.6 Tangent planes

In your first semester calculus class you learned how to use derivatives to determine the equation of a line tangent to a curve $y = f(x)$. In this lecture we'll see how to use partial derivatives to determine the equation of a plane tangent to a surface $z = f(x,y)$. In the next lecture we'll apply these ideas to determine linear approximations of multivariable functions. Before we discuss tangent planes, we need to review some prerequisites that will be needed just to make sense of the idea of a tangent plane.

Linear combinations

Suppose that $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_n$ are n different vectors all of the same dimension (i.e., each $\vec{v}_i$ is a vector with m components for some fixed m). Then we define a *linear combination* of these vectors as any expression of the form

$$\sum_{i=1}^{n} \lambda_i \vec{v}_i = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n.$$

where each $\lambda_i \in \mathbb{R}$ is a scalar. Notice that a linear combination is itself a vector.

Example 3.29.
The vector $\vec{u} = \langle 6, -2, 5\rangle$ in 3-space is a linear combination of the vectors $\vec{v}_1 = \langle -2, 1, 0\rangle$ and $\vec{v}_2 = \langle 0, 4, 20\rangle$ since we may write $\vec{u} = -3\vec{v}_1 + \frac{1}{4}\vec{v}_2$:

$$\begin{aligned} -3\vec{v}_1 + \frac{1}{4}\vec{v}_2 &= -3\,\langle -2, 1, 0\rangle + {}^1\!/\!_4\,\langle 0, 4, 20\rangle \\ &= \langle 6, -3, 0\rangle + \langle 0, 1, 5\rangle \\ &= \langle 6, -2, 5\rangle\,. \end{aligned}$$

So linear combinations allow us to express one vector by breaking it down into vectors that point in a fixed direction; we can decompose the vector into different components.

Example 3.30.
Every vector in 2-space can be decomposed into a linear combination of horizontal and vertical vectors. That is, any given vector $\langle x, y\rangle$ may be written as

$$\langle x, y\rangle = x\,\langle 1, 0\rangle + y\,\langle 0, 1\rangle = x\vec{\imath} + y\vec{\jmath}.$$

This easy observation is sometimes very useful in simple physical problems. E.g., calculating work done by a force vector that points in a direction different from the direction of motion.

Example 3.31.
Recalling that work may be defined as force times distance, how much work is done in moving a cart four meters horizontally, if the cart is pulled by a rope which makes a $45°$ angle from the horizontal with $\sqrt{18}$ Newtons of tension in the rope?

We know the rope makes a $45°$-angle, so its horizontal and vertical components must be the same. Since these components have to give us a vector of magnitude $\sqrt{18}$, the components must both be 3: the force vector is $\langle 3, 3\rangle$. Since there's 3N of force pulling horizontally, over a distance of 4m, the work done is 12 Nm (aka 12 joules).

This idea of breaking a 2-dimensional vector into its horizontal and vertical components extends to 3-dimensions. Here we have a horizontal component (the x-direction), a vertical component (the y-direction), and also a "depth component" (the z-direction). This means that any vector in 3-space can be written as a linear combination of the vectors $\vec{\imath} = \langle 1, 0, 0\rangle$, $\vec{\jmath} = \langle 0, 1, 0\rangle$, and $\vec{k} = \langle 0, 0, 1\rangle$:

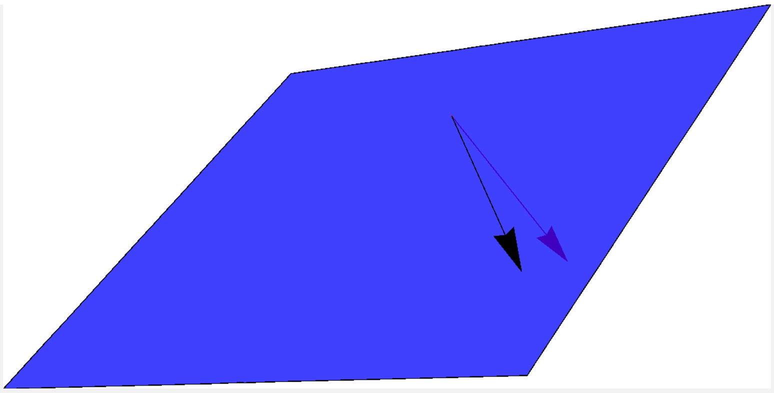
$$\langle x, y, z\rangle = x\vec{\imath} + y\vec{\jmath} + z\vec{k}.$$

Planes and linear combinations

What we'd like to do is use our knowledge about linear combinations to help us determine a plane. To do this, we first need to talk about vectors that live in the plane.

Recall that a vector $\vec{v}$ in 3-space doesn't have a distinguished footpoint: we can slide the vector around as much as we'd like, and as long as we don't stretch or rotate the vector, we have the same vector. We'll say that a vector $\vec{v}$ "lives" in a given plane, say $ax + by + cz = d$, if we can slide the vector around so that its tail is a point in the plane and its tip is also a point in the plane. Calling the tip of the vector the point P, and the tail of the vector the point Q, this just means that the coordinates of P and Q both satisfy the equation $ax + by + cz + d = 0$.

Example 3.32.
Do the vectors $\vec{v}_1 = \langle 3, -4, 1\rangle$ and $\vec{v}_2\ \langle 2, 0, -2\rangle$ live in the plane $5x + y - 11z = 2$?



Let's say that we slide the vectors so that they both have their tails at some given point in the plane. This means we first need to find a point in the plane: we need an (x, y, z) that satisfies $5x + y - 11z = 2$. One easy way to do this would be to set $(x, y, z) = (0, 2, 0)$. So let's take $P = (0, 2, 0)$ as the tail of both of our vectors.

Now we need the tips of the vectors $\vec{v}_1$ and $\vec{v}_2$ to also be points in the plane. Let's call the tip of $\vec{v}_1$ the point Q_1, and the tip of $\vec{v}_2$ the point Q_2. To determine the coordinates of Q_1 and Q_2, we add the components of the vectors $\vec{v}_1$ and $\vec{v}_2$ to the coordinates of P.

Adding $\vec{v}_1 = \langle 3, -4, 1\rangle$ to $(0, 2, 0)$ we find that $Q_1 = (3, -2, 1)$. Adding $\vec{v}_2 = \langle 2, 0, -2\rangle$ to $(0, 2, 0)$, we find that $Q_2 = (2, 2, -2)$. Now we need to determine if Q_1 and Q_2 are points in our plane. To do this we see if the coordinates of Q_1 and Q_2 satisfy the equation of our plane, $5x + y - 11z = 2$.

Plugging in the coordinates of Q_1:

$$5 \cdot 3 + 1 \cdot (-2) - 11 \cdot 1 = 15 - 2 - 11 = 2.$$

So Q_1 is a point in the plane, and since the tail and tip of $\vec{v}_1$ are points in the plane, we have that $\vec{v}_1$ lives in the plane.

Plugging in the coordinates of Q_2:

$$5 \cdot 2 + 1 \cdot 2 - 11 \cdot (-2) = 10 + 2 - 22 = -10 \neq 2$$

and so Q_2 is not a point in the plane. Since the tail of $\vec{v}_2$ is in the plane, but the tip is not, the vector $\vec{v}_2$ does not live in the plane.

An observation we'll need to construct the tangent plane of a surface is that if $\vec{v}_1$ and $\vec{v}_2$ are any two non-parallel vectors living in a given plane, then *all* the vectors living in the plane can be expressed as linear combinations of $\vec{v}_1$ and $\vec{v}_2$. Before showing this, let's first note the following:

Proposition 3.11.
If $\vec{v}_1$ and $\vec{v}_2$ are two vectors in the plane $ax + by + cz = d$, then all linear combinations of $\vec{v}_1$ and $\vec{v}_2$ also live in the plane.

Proof.
Let's write the coordinates of $\vec{v}_1$ and $\vec{v}_2$ as

$$\begin{aligned} \vec{v}_1 &= \langle x_1, y_1, z_1 \rangle \,, \text{ and} \\ \vec{v}_2 &= \langle x_2, y_2, z_2 \rangle \,. \end{aligned}$$

Note that these components satisfy the equations

$$ax_i + by_i + cz_i = 0$$

where $i = 1$ or $i = 2$. To see this, note that if $P = \langle \alpha, \beta, \gamma \rangle$ is a point in the plane that we use as a tail of $\vec{v}_i$, then the tip $Q = (\alpha + x_i, \beta + y_i, \gamma + z_i)$ is also a point in the plane. This means

$$\begin{aligned} &a\alpha + b\beta + c\gamma = d, \text{ and} \\ &a(\alpha + x_i) + b(\beta + y_i) + c(\gamma + z_i) = d. \end{aligned}$$

Subtracting the first equation from the second, we have

$$ax_i + by_i + cz_i = 0.$$

Now, suppose that we pick some point (A, B, C) in the plane (so $aA + bB + cC = d$), and we add the vector $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2$ to (A, B, C) to get the point

$$(A + \lambda_1 x_1 + \lambda_2 x_2, B + \lambda_1 y_1 + \lambda_2 y_2, C + \lambda_1 z_1 + \lambda_2 z_2).$$

We need to show this point also lives in the plane, but this is obvious if we just rearrange the terms above:

$$\begin{aligned}
&a(A+\lambda_1 x_1+\lambda_2 x_2)+b(B+\lambda_1 y_1+\lambda_2 y_2)+c(C+\lambda_1 z_1+\lambda_2 z_2)\\
&=aA+bB+cC+\lambda_1(x_1+y_1+z_1)+\lambda_2(x_2+y_2+z_2)\\
&=aA+bB+cC+0+0\\
&=d.
\end{aligned}$$

□

So any linear combination of two vectors living in the plane gives us another vector in the plane. Now the question is can we go in the other direction: if we fix two vectors in the plane, call them $\vec{v}_1$ and $\vec{v}_2$, and then choose some third vector in the plane, call it $\vec{u}$, can we write $\vec{u}$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

To prove this, let's first consider a simpler situation: let's suppose our plane is just the xy-plane.

Proposition 3.12.
If $\vec{v}_1$ and $\vec{v}_2$ live in the xy-plane and aren't parallel, then any other vector, $\vec{u}$, in the xy-plane is linear combination of $\vec{v}_1$ and $\vec{v}_2$.

Proof.
Let's first suppose the coordinates of our vectors as follows:

$$\begin{aligned}
\vec{v}_1 &= \langle x_1, y_1\rangle\\
\vec{v}_2 &= \langle x_2, y_2\rangle\\
\vec{u} &= \langle x_u, y_u\rangle .
\end{aligned}$$

We need to find a λ_1, λ_2 such that

$$\vec{u} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2.$$

In terms of components we have

$$\langle x_u, y_u \rangle = \lambda_1 \langle x_1, y_1 \rangle + \lambda_2 \langle x_2, y_2 \rangle .$$

Equating components we have a system of equations,

$$\begin{aligned} x_u &= \lambda_1 x_1 + \lambda_2 x_2 \\ y_u &= \lambda_1 y_1 + \lambda_2 y_2 \end{aligned}$$

You may have learned in high-school that a system of equations such as this can be converted into an equation with matrices that is easily solved. (Even if you didn't learn this before, you will learn how to solve such a system when you take linear algebra next semester.)

This system is solved by setting

$$\lambda_1 = \frac{x_u y_2 - y_u x_2}{x_1 y_2 - x_2 y_1}$$

$$\lambda_2 = \frac{x_1 y_u - y_1 x_u}{x_1 y_2 - x_2 y_1}.$$

Even if you don't see how to come up with this solution, it's easy to check that this is in fact the solution:

$$\begin{aligned} \lambda_1 x_1 + \lambda_2 x_2 &= \left(\frac{x_u y_2 - y_u x_2}{x_1 y_2 - x_2 y_1} \right) x_1 + \left(\frac{x_1 y_u - y_1 x_u}{x_1 y_2 - x_2 y_1} \right) x_2 \\ &= \frac{x_u x_1 y_2 - y_u x_1 x_2 + y_u x_1 x_2 - x_u y_1 x_2}{x_1 y_2 - x_2 y_1} \\ &= \frac{x_u (x_1 y_2 - y_1 x_2)}{x_1 y_2 - x_2 y_1} \\ &= x_u \\ \lambda_1 y_1 + \lambda_2 y_2 &= \left(\frac{x_u y_2 - y_u x_2}{x_1 y_2 - x_2 y_1} \right) y_1 + \left(\frac{x_1 y_u - y_1 x_u}{x_1 y_2 - x_2 y_1} \right) y_2 \\ &= \frac{x_u y_1 y_2 - y_u y_1 x_2 + y_u x_1 y_2 - x_u y_1 y_2}{x_1 y_2 - x_2 y_1} \\ &= \frac{y_u (-y_1 x_2 + x_1 y_2)}{x_1 y_2 - x_2 y_1} \\ &= y_u \end{aligned}$$

Notice that the denominator, $x_1 y_2 - x_2 y_1$ can never be zero if our vectors are not parallel. If this did equal zero, we could rearrange the terms to have $x_1/x_2 = y_1/y_2$. Calling this value μ, it's not hard to see that $\langle x_1, y_1 \rangle = \mu \langle x_2, y_2 \rangle$, but we know this can't happen because our vectors aren't parallel. □

Proposition 3.13.
Suppose that $\vec{v}_1 = \langle x_1, y_1, z_1 \rangle$ and $\vec{v}_2 = \langle x_2, y_2, z_2 \rangle$ are two non-parallel vectors living in the plane $ax + by + cz = d$. If $\vec{u} = \langle x_u, y_u, z_u \rangle$ is any other vector that also lives in the plane, then $\vec{u}$ may be written as a linear combination of $\vec{v}_1$ and $\vec{v}_2$.

Proof.
We need to find a λ_1 and λ_2 so that $\vec{u} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2$. To see that such λ_1, λ_2 exist, we'll use our proposition above. Simply project the vectors $\vec{v}_1$ and $\vec{v}_2$ to each of the coordinate planes (the xy-plane, the yz-plane, and the xz-plane). If $\vec{v}_1$ and $\vec{v}_2$ are not parallel, then in at least one of the coordinate planes, the projections are also not parallel. The above lemma then tells us how to find the corresponding λ_1 and λ_2. □

Curves on surfaces

Suppose that we have a surface, $z = f(x, y)$. A *curve on the surface* is a parametric curve $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ so that for each value of t, the point $(x(t), y(t), z(t))$ lies on the surface. That is, $z(t) = f(x(t), y(t))$. It is extremely easy to take parametric curves in the xy-plane and lift them up to the surface $z = f(x, y)$.

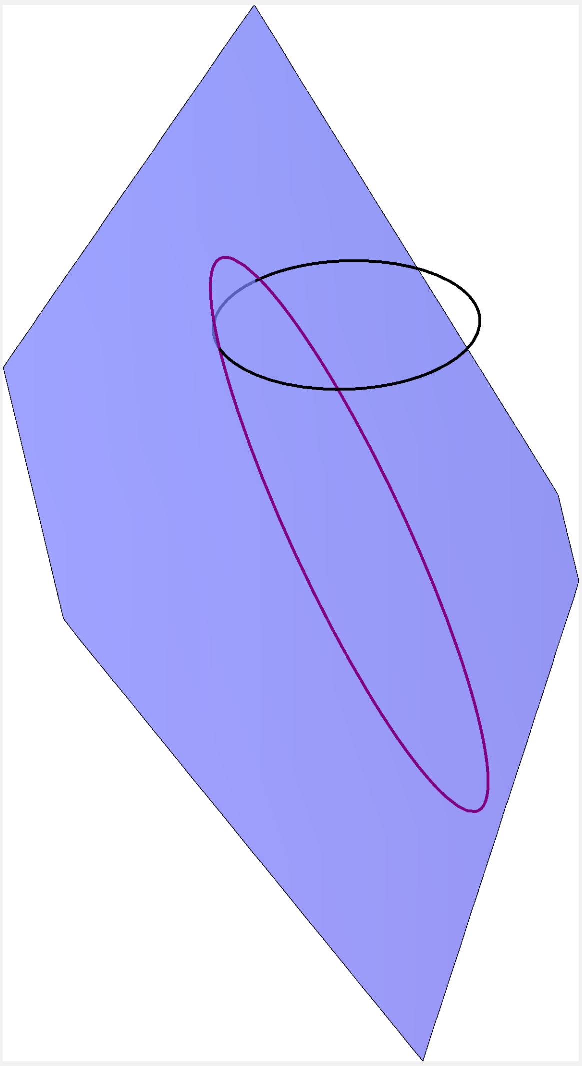
Suppose that $\vec{\rho}(t) = \langle x(t), y(t) \rangle$ is a parametric curve in the xy-plane. We can turn this into a curve on the surface $z = f(x, y)$ by setting $\vec{r}(t) = \langle x(t), y(t), f(x(t), y(t)) \rangle$.

Example 3.33.
Lift the circle $\vec{\rho}(t) = \langle \cos(t), \sin(t) \rangle$ to a curve on the surface $z = 3x - 2y - 3$.

We simply set

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

This picks the circle up from the xy-plane, and puts it into the surface.



To calculate the tangent plane, we'll need to be able to calculate tangent vectors of the curves on the surface, and to do this we need to use the chain rule for multivariable functions.

If we lift the curve $\vec{\rho}(t) = \langle x(t), y(t) \rangle$ to the surface $z = f(x, y)$ by setting $\vec{r}(t) = \langle x(t), y(t), f(x(t), y(t)) \rangle$, then the tangent vectors of the curve are given by

$$\frac{d}{dt}\vec{r}(t) = \left\langle \frac{d}{dt}x(t), \frac{d}{dt}y(t), \frac{d}{dt}f(x(t), y(t)) \right\rangle.$$

The first two components are of course just $x'(t)$ and $y'(t)$. By the multivariable chain rule we discussed last time, we know the derivative of the third component is

$$\begin{aligned} \frac{d}{dt}f(x(t), y(t)) &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t). \end{aligned}$$

Example 3.34.
Lift the curves $\vec{\rho}_1(t) = \langle t^2, t \rangle$ and $\vec{\rho}_2(t) = \left\langle e^{-t} - 1, e^{\sin(t)} - 2 \right\rangle$ to the surface $z = xy$ and calculate the tangent vectors of the lifted curves at time $t = 0$.

Our curves are of course just

$$\begin{aligned} \vec{r}_1(t) &= \left\langle t^2, t, t^3 \right\rangle \\ \vec{r}_2(t) &= \left\langle e^{-t} - 2, e^{\sin(t)} - 1, (e^t - 1) \cdot (e^{\sin(t)} - 2) \right\rangle. \end{aligned}$$

The tangent vector of a lifted curve has the form

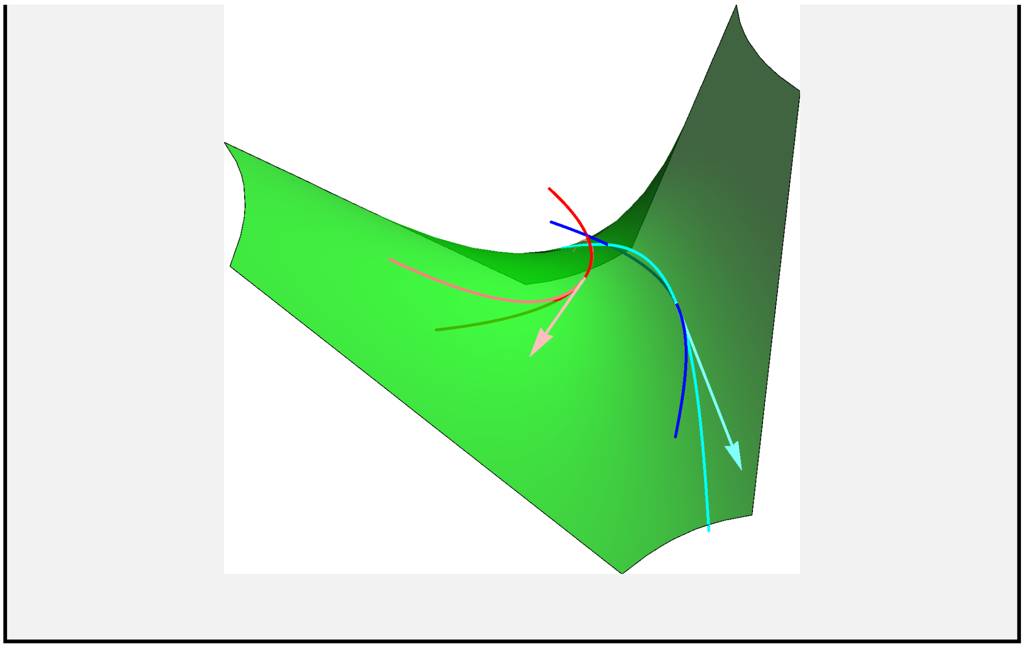
$$\langle x'(t), y'(t), y(t)x'(t) + x(t)y'(t) \rangle,$$

and so our two lifted curves have derivatives

$$\begin{aligned} \vec{r}_1'(t) &= \left\langle 2t, 1, t^2 + 2t^2 \right\rangle = \left\langle 2t, 1, 3t^2 \right\rangle \\ \vec{r}_2'(t) &= \left\langle -e^{-t}, \cos(t)e^{\sin(t)}, -e^{-t}(e^{\sin(t)} - 1) + \cos(t)e^{\sin(t)}(e^{-t} - 2) \right\rangle. \end{aligned}$$

At time $t = 1$ our tangent vectors as

$$\begin{aligned} \vec{r}_1'(1) &= \langle 0, 1, 0 \rangle \\ \vec{r}_2'(1) &= \langle -1, 1, -1 \rangle. \end{aligned}$$



Tangent Planes

We now have enough information to determine a tangent plane. Geometrically a "tangent plane" is exactly what you think it should be: a plane which is "tangent" to a surface. To make this idea precise, we'd like to relate it to things we already know, such as tangent vectors of curves. We'll use this as our definition of tangent plane because it's easy to work with.

The *tangent plane* of the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$ on the surface is the collection of all vectors tangent to curves on the surface which pass through the point $(x_0, y_0, f(x_0, y_0))$.

This may sound like a strange definition, but it's actually fairly straightforward. If you have any curve on the surface (and we know how to take curves in the xy-plane and lift them up to curves on the surface), then a tangent vector to that curve (which we know how to calculate), should intuitively live in the tangent plane. So we simply define the tangent plane to be the collection of all possible tangent vectors we could get.

Now the question is how to do we go about actually finding the equation of this tangent plane. Here we'll need to combine several things we've learned. We know that the equation of a plane through the point (x_0, y_0, z_0) with normal vector $\vec{n} = \langle a, b, c \rangle$ is given by

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

We also know that since the normal vector of a plane is orthogonal to all of the vectors that live in that plane, we can find a normal vector by taking the cross product of two vectors in the plane.

To find the tangent plane of the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$, we need two vectors that live in this plane. We'll then take the cross product of these two vectors to get our normal. To find two such vectors we need to find two curves on the surface that pass through the point $(x_0, y_0, f(x_0, y_0))$. We'll then differentiate these curves to get our tangent vectors.

To get two curves on the surface, it suffices to take two curves in the xy-plane, and then lift them up to the surface. The only thing we have to be careful of is that our curves on the surface need to pass through $(x_0, y_0, f(x_0, y_0))$. This just means that the corresponding curves in the xy-plane need to both pass through the point (x_0, y_0).

There are of course lots and lots of curves in the xy-plane that pass through the point (x_0, y_0), but to make our lives a little bit easier, let's just pick the two simplest possible curves: a horizontal line and a vertical line.

The horizontal line in the xy-plane that passes through the point (x_0, y_0) may be parametrized as

$$\vec{\rho}_h(t) = \langle x_0 + t, y_0 \rangle .$$

The vertical line in the xy-plane that passes through the point (x_0, y_0) may be parametrized as

$$\vec{\rho}_v(t) = \langle x_0, y_0 + t \rangle .$$

(There are lots of other ways we could parametrize this line, but we'll choose this particular parametrization since it goes through the point (x_0, y_0) at time $t = 0$.)

Lifting these curves up to our surface $z = f(x, y)$, we have the curves

$$\begin{aligned}\vec{r}_h(t) &= \langle x_0 + t, y_0, f(x_0 + t, y_0) \rangle , \\ \vec{r}_v(t) &= \langle x_0, y_0 + t, f(x_0, y_0 + t) \rangle .\end{aligned}$$

These curves go through the point $(x_0, y_0, f(x_0, y_0))$ at time $t = 0$. Hence the corresponding tangent vectors are $\vec{r}_h'(0)$ and $\vec{r}_v'(0)$:

$$\begin{aligned}\vec{r}_h'(0) &= \langle 1, 0, f_x(x_0, y_0) \rangle \\ \vec{r}_v'(0) &= \langle 0, 1, f_y(x_0, y_0) \rangle .\end{aligned}$$

Taking the cross product of these two vectors, the normal vector to our tangent plane is

$$\begin{aligned}\vec{r}_v'(0) \times \vec{r}_h'(0) &= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ 0 & 1 & f_y(x_0, y_0) \\ 1 & 0 & f_x(x_0, y_0) \end{bmatrix} \\ &= \langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle\end{aligned}$$

Combining all of the above gives us the following.

Theorem 3.14.
The tangent plane of the surface $z = f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0)$ is given by the equation

$$f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) - (z - f(x_0, y_0)) = 0.$$

Example 3.35.
Find the equation of the plane tangent to the cone $z = \sqrt{x^2 + y^2}$ at the point $(1, 1, \sqrt{2})$.

To use the above formula for the tangent plane, we first need to find the partial derivatives:

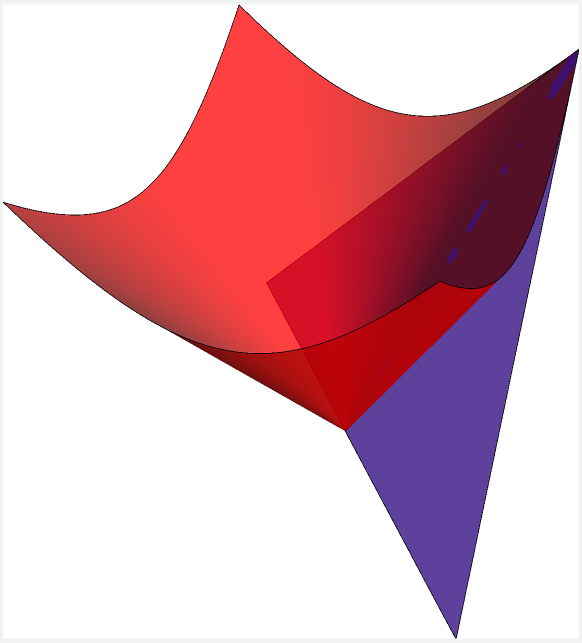
$$\begin{aligned}f_x(x, y) &= \frac{2x}{2\sqrt{x^2 + y^2}} \\ f_y(x, y) &= \frac{2y}{2\sqrt{x^2 + y^2}}\end{aligned}$$

At the point $(x, y) = (1, 1)$, these give us the values

$$\begin{aligned}f_x(1, 1) &= \frac{1}{\sqrt{2}} \\ f_y(1, 1) &= \frac{1}{\sqrt{2}}.\end{aligned}$$

Hence the equation of our tangent plane is

$$\frac{x-1}{\sqrt{2}} + \frac{y-1}{\sqrt{2}} - (z - \sqrt{2}) = 0.$$



Tangent planes and implicitly defined functions

Most of the surfaces we have considered in this class have been given as graphs of functions of two variables, $z = f(x, y)$. Just as there are curves in the plane which do not appear as graphs of functions $y = f(x)$, there are surfaces in 3-space which are not the graphs of functions of two variables. One simple example is the unit sphere,

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

This can not be the graph of a function because it also fails the vertical line test: a vertical line (i.e., parallel to the z-axis) through a given (x, y) point in the plane (i.e., $(x, y, 0)$ in 3-space) may intersect the sphere in two points. However, a function can only have one output for any given input. That is, there is at most one point on the surface $z = f(x, y)$ for each point (x, y) in the plane.

Though not all surfaces are graphs of functions, we may still be interested in their tangent planes. (This has applications to computer graphics, for example.) So, how can we compute the tangent plane of an implicitly defined surface? What we really need to find is a vector that is orthogonal to the surface at a given point, and we can find this by reinterpreting our implicitly defined surface as a level surface of a function of three variables and using the gradient.

That is, if we have an expression of three variables that defines our surface, such as the sphere above, we could move all the variables to one side of the equation and consider the function of three-variables defined by that expression. In the case of the sphere this would be $f(x, y, z) = x^2 + y^2 + z^2$. The gradient of this function, $\nabla f(x, y, z)$ will *always* be perpendicular to the level surfaces of the function. More precisely, we have the following theorem:

Theorem 3.15.
For any scalar-valued function of n-variables, $f : \mathbb{R}^n \to \mathbb{R}$ and any point $(a_1, ..., a_n)$ on the level set $f(x_1, ..., x_n) = c$, the gradient at that point $\nabla f(a_1, ..., a_n)$ will be perpendicular to every curve on the level set at $(a_1, ..., a_n)$.

This theorem tells us, in particular, that $\nabla f(x_0, y_0, z_0)$ is orthogonal to the tangent vector of every curve on the level surface that passes through (x_0, y_0, z_0), which implies it is orthogonal to the tangent plane. That is, the gradient evaluated at the point,

$$\nabla f(x_0, y_0, z_0) = \langle f_x(x_0, y_0, z_0),\ f_y(x_0, y_0, z_0),\ f_z(x_0, y_0, z_0) \rangle$$

is a normal vector for the tangent plane. Since we have a normal vector and a point on the plane, we can now write down the equation of the plane as

$$f_x(x_0, y_0, z_0) \cdot (x - x_0) + f_y(x_0, y_0, z_0) \cdot (y - y_0) + f_z(x_0, y_0, z_0) \cdot (z - z_0) = 0.$$

Example 3.36.
Find the equation of the tangent plane to the unit sphere $x^2+y^2+z^2 = 1$ at the point $(1/\sqrt{14}, 2/\sqrt{14}, 3/\sqrt{14})$.

Thinking of the unit sphere as a level curve of the function $f(x, y, z) = x^2 + y^2 + z^2$, we consider the gradient

$$\nabla f = \langle 2x, 2y, 2z \rangle$$

and evaluate it at the given point to obtain our normal vector,

$$\left\langle \frac{2}{\sqrt{14}}, \frac{4}{\sqrt{14}}, \frac{6}{\sqrt{14}} \right\rangle.$$

The equation of the tangent plane is thus

$$\frac{2}{\sqrt{14}}\left(x - \frac{1}{\sqrt{14}}\right) + \frac{4}{\sqrt{14}}\left(y - \frac{2}{\sqrt{14}}\right) + \frac{6}{\sqrt{14}}\left(z - \frac{3}{\sqrt{14}}\right) = 0$$

It's interesting to notice that we could fit our discussion of tangent planes of graphs $z = f(x, y)$ into the framework above by thinking of $z = f(x, y)$ as the level surface $f(x, y) - z = 0$. That is, we would consider the function $g(x, y, z) = f(x, y) - z$. The gradient of this function is

$$\nabla g(x, y, z) = \langle f_x(x, y), f_y(x, y), -1 \rangle$$

which gives us the normal vectors we had considered above.

3.7 Linearization

In this lecture we'll apply tangent planes, the topic of the previous lecture, to show how to obtain a linear approximation of a multivariable function.

Motivation

Suppose that we're given a function of a single variable, $f(x)$. If this function is differentiable, then we can use the tangent line of the graph of the function to get an approximation of the function. That is, we know the equation of the tangent line of the surface at some point $(x_0, f(x_0))$ is

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Moving the $f(x_0)$ to the right-hand side, we have a line

$$y = f'(x_0)(x - x_0) + f(x_0)$$

which is the graph of the function

$$L(x) = f'(x_0)(x - x_0) + f(x_0).$$

What's nice about this function $L(x)$ is that it's something we can actually evaluate. If you think about the mathematical, numerical operations you can actually perform – the things you can in principle sit down and work out with a pencil and paper – you realize that there are basically only four operations: addition, subtraction, multiplication, and division. (These are the four *arithmetic operations*.) There are of course some more complicated things we know how to do (for example, cubing a number), but these are really built out of combinations of addition, subtraction, multiplication, and division (e.g., $x^3 = x \cdot x \cdot x$).

Using a computer, by the way, doesn't really let you do any more operations than what you can do with pencil and paper. Ultimately, computers also can only do arithmetic: they aren't magically able to perform things that people in principle can not. In fact, in some ways computers are worse at these arithmetic operations than people. A computer has to represent numbers using a finite number of *bits*: values that can only be 1 or 0. Since any computer only has a finite number of these bits – even if it's a very large number! – "most" numbers can't be represented exactly on a computer. It turns out that "most" numbers would require an infinite number of bits to represent exactly, so the computer has to use approximations. You might be surprised to learn that a number as simple as $1/10$ can't be represented exactly with a finite number of bits (at least not the way computers usually represent numbers)! A really simple way to demonstrate this is to ask the computer to add $0.1 + 0.2$. If you did this in Python, for example, you'll get

```
>>> 0.1 + 0.2
0.30000000000000004.
```

Usually the computer truncates – cuts off – the values before printing them, but internally represents the number to more decimal places. When you do lots of calculations with these values, these little errors start to add up!

Anyway, the four arithmetic operations are basically the only tools we have to do numerical computations. However, certain types of functions are defined in terms of these arithmetic operations. The trig functions, $\cos\theta$ and $\sin\theta$, for example, are defined geometrically: they're the x- and y-coordinates of points on the circle. Yet somehow your computer is able to spit out a value if you enter $\cos(0.3245)$. If the computer can only do arithmetic, how is it able to determine this value? The answer is that the computer uses calculus (or, rather, someone who knew calculus programmed the computer) to determine approximations of $\cos\theta$ that can actually be evaluated by using only arithmetic operations. This is of course the Taylor polynomials you learned about in your second-semester calculus class.

The techniques of Taylor polynomials you learned before are very powerful, and are the basis for all of the fancy technology we have today: breakthroughs in science and medicine are possible today because people are able to use computers to do calculations and analyze very large amounts of data, and they're able to do this because we can use Taylor polynomials (and related ideas) to convert complicated calculations into arithmetic.

However, the material you've learned before is only applicable to functions of a single variable. Our goal in this lecture is to start studying the comparable ideas for functions of several variables. To do this, we'll use tangent planes to get a *linear approximation* of a multivariable function.

Linearization

Recall from the last lecture that we said tangent plane of the surface $z = f(x, y)$ at the point (x_0, y_0) is given by

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - f(x_0, y_0)) = 0.$$

Solving this for z we have

$$z = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + f(x_0, y_0).$$

Notice that this is a function of x and y, and this is a plane which is a graph of the function

$$L(x, y) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + f(x_0, y_0),$$

which we call the *linearization* of the function $f(x, y)$ at the point (x_0, y_0).

The reason we care about linearizations is that they're functions we (or a computer) can actually compute: we can get numerical answers using a linearization. We can thus use linearizations to approximate multivariable functions.

Example 3.37.
Calculate the linearization of $f(x, y) = xe^y$ at the point $(2, 0)$, and use the linearization to approximate $f(2.01, -0.1)$.

To calculate the linearization, we first need to calculate the partial derivatives.

$$f_x(x, y) = e^y$$
$$f_y(x, y) = xe^y$$

Now we must evaluate these partial derivatives, and our original function, at the point $(x_0, y_0) = (2, 0)$.

$$f(2, 0) = 2$$
$$f_x(2, 0) = 1$$
$$f_y(2, 0) = 2$$

We use these values to now determine our linearization,

$$L(x, y) = (x - 2) + 2y + 2 = x + 2y$$

We can now use this approximation to estimate the value $f(2.01, -0.1)$:

$$\begin{aligned} f(2.01, -0.1) &\approx L(2.01, -0.1) \\ &= 2.01 + 2 \cdot (-0.1) \\ &= 2.01 - 0.2 \\ &= 1.81 \end{aligned}$$

Let's take a moment to think about what we've just done in the example above. We used linearization, which is essentially just the equation of a tangent plane, to estimate the value $2.01 \cdot e^{-0.1}$ and got an actual numerical value. This is an extremely important idea: we can complicated functions and estimate them with things we actually know how to calculate – this we can sit down and really do with a pencil and paper. That is, we were able to say

$$2.01 \cdot (2.718281824...)^{-0.1} \approx 1.81.$$

By the way, if you plug $2.01e^{-0.1}$ in a calculator or computer, it will probably spit back the answer

$$2.01e^{-0.1} \approx 1.81872321025.$$

This means two things: 1) our approximation above, which was super easy to actually calculate by hand, is a decent approximation; and 2) the computer uses a different type of approximation than what we used. The computer is using a Taylor polynomial, probably to some high degree, to get its approximation. The idea of Taylor polynomials in one variable, you may recall, is really just an extension of the idea of linearization (using tangent lines as the approximation). We can also do Taylor polynomials in several variables, but won't work on that right now. For this lecture we'll focus on linearization, and may come back to multivariable Taylor polynomials at the end of the semester if we have extra time.

Example 3.38.
Calculate the linearization of $f(x,y) = x^3y + y^2x$ at the point $(-1,3)$, and use the linearization to approximate $f(-0.93, 2.976)$.

First we calculate the partial derivatives,

$$\begin{aligned} f_x(x,y) &= 3x^2y + y^2 \\ f_y(x,y) &= x^3 + 2xy \end{aligned}$$

Evaluating these partials, and the original function, at $(-1,3)$ we have

$$\begin{aligned} f(-1,3) &= (-1)^3 \cdot 3 + 3^2 \cdot (-1) = -3 - 9 = -12 \\ f_x(-1,3) &= 3 \cdot (-1)^2 \cdot 3 + 3^2 = 9 + 9 = 18 \\ f_y(-1,3) &= (-1)^3 + 2 \cdot (-1) \cdot 3 = -1 - 6 = -7 \end{aligned}$$

Hence the linearization is

$$L(x,y) = 18(x+1) - 7(y-3) - 12.$$

We now use this to get an approximation,

$$\begin{aligned} f(-0.93, 2.976) &\approx L(-0.93, 2.976) \\ &= 18 \cdot (1 - 0.93) - 7(2.976 - 3) - 12 \\ &= 18 \cdot (0.07) - 7 \cdot (-0.024) - 12 \\ &= 1.26 + 0.168 - 12 \\ &= 1.428 - 12 \\ &= -10.576 \end{aligned}$$

Differentials

Notice in both of the examples above we had to pick values (x_0, y_0), the "center" of our approximation, where we could actually calculate the true value of the function. This is always the case for these linearizations: we have to find a place to "anchor" our approximation; we need somewhere where we know definitively what the function equals. Let's say we know the true value $z_0 = f(x_0, y_0)$. Calling $L(x, y) = z$, our linearization has the form:

$$z = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0) + z_0.$$

Moving the z_0 to the other side we have

$$z - z_0 = f_x(x_0, y_0) \cdot (x - x_0) + f_y(x_0, y_0) \cdot (y - y_0).$$

Notice that $z - z_0$, $x - x_0$, and $y - y_0$ are just the change in the values of z, x, and y when change our inputs from x_0 to x; from y_0 to y; and then the output changes from z_0 to z. That is, each of $x - x_0$, $y - y_0$, and $z - z_0$ represents the change in x, y, and z. Let's write these changes as dx, dy, and dz. (We use the letter 'd' for "difference.") Our equation above then becomes

$$dz = f_x(x_0, y_0)dx + f_y(x_0, y_0)dy.$$

Thinking of dx and dy as variables (just saying how much we vary the original inputs x and y), we have that dz is a function of two variables. This function is called the *differential* of $z = f(x, y)$.

The idea here is that differentials measure the change in our approximation. For example, in our example above, we have

$$dz = 18dx - 7dy.$$

This means we can determine the change in approximation, dz, by just plugging in the changes dx and dy in our variables x and y. In the example above we changed x from -1 to -0.93. This is a change of

$$dx = -0.93 - (-1) = 0.07.$$

We changed y from 3 to 2.976. This is a change of

$$dy = 2.976 - 3 = -0.024$$

So the change in z from $f(-1, 3) = 12$ to our approximation $L(-0.93, 2.976)$ is

$$dz = 18 \cdot (0.07) - 7 \cdot (-0.024) = 1.428.$$

This means our function $f(x, y)$ changes by approximately 1.428 when we move the inputs of the function from $(x_0, y_0) = (-1, 3)$ to $(-0.93, 2.976)$; so the approximation is $-12 + dz = -12 + 1.428 = -10.576$, as we saw above.

Differentials and linearizations are two sides of the same coin: they're basically the same thing, just represented different ways. More precisely, a differential is just a change in linearization. This means that

$$f(x, y) \approx L(x, y) = f(x_0, y_0) + dz.$$

(Since we'll usually write $z = f(x, y)$, we may sometimes write df for dz, and call this value the *differential* of f instead of the differential of z. These are the same thing, just different words.)

Example 3.39.
Calculate the differential dz of $z = \sin(x + 3y^2)$:

$$\begin{aligned} dz &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \\ \implies dz &= \frac{\partial}{\partial x}\sin(x + 3y^2)dx + \frac{\partial}{\partial y}\sin(x + 3y^2)dy \\ \implies dz &= \cos(x + 3y^2)dx + 6y\cos(x + 3y^2)dy \end{aligned}$$

Example 3.40.
Calculate the differential dz of $z = f(x, y) = (x^3 - 2) \cdot \tan^{-1}(y)$, then use the differential to approximate $f(-2.1, 0.22)$ and $f(-1.99, 0.18)$.

$$dz = 3x^2 \tan^{-1}(y)dx + \frac{x^3 - 2}{1 + y^2}dy$$

Use differentials (or linearizations), we need to find a point (x_0, y_0) to use as the "center" of our approximation; some value near the values we're trying to approximate, where we can exactly calculate the true value of the function. Let's use $(x_0, y_0) = (-2, 0)$. Then our differential becomes

$$dz = 12dx - 10dy$$

and the true value of the function is $f(-2, 0) = -10$.

For $(-2.1, 0.22)$, we have $dx = -0.1$ and $dy = 0.22$, so

$$dz = 12 \cdot (-0.1) - 10 \cdot (0.22) = -1.2 - 2.2 = -3.4$$

so our approximation for the function is

$$f(-2.1, 0.22) \approx f(-2, 0) + dz = -10 - 3.4 = -13.4.$$

For $(-1.99, 0.18)$, $dx = 0.01$ and $dy = 0.18$, so

$$dz = 12 \cdot 0.01 - 10 \cdot 0.18 = 0.12 - 1.8 = -1.68$$

so

$$f(-1.99, 0.18) \approx -10 - 1.68 = -11.68.$$

We can describe differentials in any number of variables, by the way. If $z = f(x_1, x_2, ..., x_n)$ is a function of n variables, the differential of z is defined to be

$$dz = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n.$$

3.8 Optimization

In this lecture we'll describe extend the optimization techniques you learned in your first semester calculus class to optimize functions of multiple variables. In particular, we'll see how the tangent planes of a function's graph will allow us to us to search for local maxima and minima. We'll also describe a second-derivative test for functions of two variables, and discuss some conditions on the domain of a multivariable function that guarantee the existence of extrema.

Maxima & Minima

Given a function of two variables, $f(x, y)$, we say that the point (x_0, y_0) is a *local maximum* of f if there exists some $\varepsilon > 0$ such that $f(x_0, y_0) \geq f(x, y)$ for all (x, y) that are within ε-distance of (x_0, y_0). (For this lecture we will assume all functions are differentiable.) Graphically this means that the point $(x_0, y_0, f(x_0, y_0))$ occurs at the top of a "hill" on the surface $z = f(x, y)$ (or the surface could be flat near this point).

A *local minimum* is defined the same way, except we require $f(x_0, y_0) \leq f(x, y)$ for all (x, y) within ε-distance of (x_0, y_0). So instead of being at the top of a hill, $(x_0, y_0, f(x_0, y_0))$ is at the bottom of a valley.

If $f(x_0, y_0) \geq f(x, y)$ for all (x, y) in the domain of f, then we say that (x_0, y_0) is a *global maximum*. If $f(x_0, y_0) \leq f(x, y)$ for all (x, y) in the doamin of f, then we say that (x_0, y_0) is a *global minimum*. Sometimes the words *absolute maximum* and *absolute minimum* are used instead. Graphically, the global maxima (minima) occur at the peaks (valleys) of the largest (deepest) hills (valleys).

Notice that local and global extrema are not unique in general, and that global extrema are also local extrema.

The obsveration that extrema occur at hills and valleys suggest a nice way of looking for extrema: find all of the hills and valleys. Geometrically, this means we want to find all of the places where the tangent plane of the function is "flat" (parallel to the xy-plane). This is the same thing as saying that the normal vector of this tangent plane points straight up, or straight down. Since we know from the previous lecture on tangent planes that the normal vector of the tangent plane of $z = f(x, y)$ at the point (x_0, y_0) is given by

$$\langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle ,$$

this is equivalent to asking that $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$.

Theorem 3.16.
If $f(x, y)$ is differentiable and (x_0, y_0) is a local extremum, then $f_x(x_0, y_0) =$

$f_y(x_0, y_0) = 0$.

It turns out that in order to guarantee a point is a global extremum, and not just a local one, we'll need a few more conditions on the function. For this reason we'll only search for local maxima and minima for the time being.

If a function isn't differentiable at a point in its domain (which effectively means the function's graph has a sharp point), it could be that this non-differentiable point is an extremum of the function. As a simple example, consider the cone $z = \sqrt{x^2 + y^2}$. The function isn't differentiable at the point $(0, 0)$, but this is clearly a local (in fact global) minimum of the function.

We will call a point (x_0, y_0) a *critical point* of $f(x, y)$ if any of the following conditions are met:

1. $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$, or
2. $f_x(x_0, y_0)$ does not exist, or
3. $f_y(x_0, y_0)$ does not exist.

Example 3.41.
Find all of the critical points of $f(x, y) = x^2 + xy + y^2 + y$.

First we calculate the partial derivatives,

$$f_x(x, y) = 2x + y$$
$$f_y(x, y) = x + 2y + 1$$

These partial derivatives obviously exist, and are continuous, for all (x, y) since they're polynomials. So to find the critical points, we now need to look for the solutions to the system

$$2x + y = 0$$
$$x + 2y + 1 = 0.$$

Moving the 1 in the second equation to the right-hand side, this is equivalent to solving the system

$$\begin{aligned} 2x + y &= 0 \\ x + 2y &= -1. \end{aligned}$$

Let's multiply the second equation by 2 and subtract it from the first:

$$2x + y - 2(x + 2y) = -3y.$$

Since the first equation equals 0 and the first equation equals -1, we must have

$$-3y = 0 - 2(-1) = 2.$$

Hence $y = {}^2/_3$. Plugging $y = {}^2/_3$ back into the first equation, we find

$$\begin{aligned} 2x + {}^2/_3 &= 0 \\ \implies 2x &= -{}^2/_3 \\ \implies x &= -{}^1/_3. \end{aligned}$$

Thus $(-{}^1/_3, {}^2/_3)$ is the only critical point of the function.

Example 3.42.
Find all of the critical points of the function $f(x, y) = \sin(x)\sin(y)$.
Taking partial derivatives we have

$$\begin{aligned} f_x(x, y) &= \cos(x)\sin(y) \\ f_y(x, y) &= \sin(x)\cos(y) \end{aligned}$$

We need the x's and y's which simultaneously make both of these quantities zero. Notice

$$\begin{aligned} \cos\theta &= 0 \text{ if } \theta = (2k+1)\frac{\pi}{2} \text{ for } k \in \mathbb{Z} \\ \sin\theta &= 0 \text{ if } \theta = k\pi \text{ for } k \in \mathbb{Z}. \end{aligned}$$

So the points $f_x(x, y) = f_y(x, y) = 0$ are

$$\begin{aligned}&\{((2k+1)\pi/2, (2j+1)\pi/2)\} \cup \{(k\pi, j\pi)\} \\ &= \left\{ (m\pi/2, n\pi/2) \,\middle|\, m, n \in \mathbb{Z}, \text{ with } m \text{ and } n \text{ both odd, or both even} \right\}.\end{aligned}$$

Once we know where all of the critical points of f are, we need some way of determining if the points maxima, minima, or neither. If the function $f(x, y)$ has continuous second-order partial derivatives, then we can use the multivariable version of the second derivative test.

Theorem 3.17 (Second Derivative Test)**.**
If $f(x, y)$ has continuous second-order partial derivatives and if (x_0, y_0) is a critical point of f, then define

$$\begin{aligned} D(x_0, y_0) &= \det \begin{pmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{pmatrix} \\ &= f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0) \cdot f_{yx}(x_0, y_0) \\ &= f_{xx}(x_0, y_0) \cdot f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 \end{aligned}$$

Then

1. *If $D(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) > 0$, then $f(x_0, y_0)$ is a local minimum.*
2. *If $D(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) < 0$, then $f(x_0, y_0)$ is a local maximum.*
3. *If $D(x_0, y_0) < 0$, then (x_0, y_0) is neither a maximum nor a minimum. (In this case we say that (x_0, y_0) is a* saddle point*.)*
4. *If $D(x_0, y_0) = 0$, then the test is inconclusive.*

Remark.
We won't prove the second derivative test, but will mention the un-

derlying idea. Recall that the second derivative of a function of a single variable measures the function's concavity at a point. The three-dimensional analogue of concavity is called *curvature*. Curvature at a point on a surface is determined by looking at the curvatures of curves on the surface which pass through the point. The largest and smallest curvatures of these curves are then multiplied together, and this is the curvature of the surface. A surface has *positive curvature* if these two largest and smallest curvatures of curves on the surface (called the *principal curvatures*) have the same sign; *negative curvature* means the principal curvatures have opposite signs; and *zero curvature* if one of the principal curvatures is zero – in this case the surface is "flat" near the point. Intuitively, surfaces of positive curvature look like a bowl; surfaces of negative curvature look like a Pringles chip; and surfaces of zero curvature contain a straight line. The D that appears in the second-derivative test is (essentially) the curvature of the surface. So when the curvature is positive ($D > 0$) we have a bowl, and it's only a question of whether the bowl opens up or down ($f_{xx}(x_0, y_0)$ is negative or positive).

Example 3.43.
Find the maxima and minima of the function $f(x, y) = y^3 + 3x^2y - 6x^2 - 6y^2 + 2$.

First we have to find the critical points, which means calculating the partial derivatives.

$$f_x(x, y) = 6x - 12x = 6x(y - 2)$$
$$f_y(x, y) = 3y^2 + 3x^2 - 12y = 3(y^2 - 4y + x^2).$$

Notice these functions are continuous everywhere, so we need to look for solutions to the system $f_x(x, y) = f_y(x, y) = 0$. If $f_x(x, y) = 0$, then either $x = 0$ or $y = 2$. We'll plug each of these into $y^2 - 4y + x^2$.

If $x = 0$, $y^2 - 4y + x^2 = 0$ becomes $y^2 - 4y = y(y - 4) = 0$, and so $y = 0$ or $y = 4$. Thus $(0, 0)$ and $(0, 4)$ are two of our critical points.

If $y = 2$, $y^2 - 4y + x^2 = 0$ becomes $-4 + x^2 = 0$ which means $x^2 = 4$, so $x = \pm 2$. Hence $(2, 2)$ and $(-2, 2)$ are our other critical points.

We have four critical points: $(0,0)$, $(0,4)$, $(2,2)$, and $(-2,2)$. To which are maxima and which are minima, we apply our second derivative test, which means we need to calculate $D(x,y)$ for each value.

In our situation,

$$\begin{aligned} D(x,y) &= \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \\ &= \det \begin{pmatrix} 6y-12 & 6x \\ 6x & 6y-12 \end{pmatrix} \\ &= (6y-12)^2 - 36x^2 \end{aligned}$$

Now we calculate this value for each of our critical points.

For $(0,0)$ we have $D(0,0) = (-12)^2 - 0^2 = 144$. So $D > 0$ and we do have a local extremum. To see if it's a local max or a local min, we need to calculate $f_{xx}(0,0) = -12 < 0$, so $(0,0)$ is a local max.

For $(0,4)$ we have $D(0,4) = (24-12)^2 - 0 = 144$, and $f_{xx}(0,4) = 12 > 0$, so $(0,4)$ is a local minimum.

For $(2,2)$ we have $D(2,2) = (12-12)^2 - 36 \cdot 4 = -144$, and so $(2,2)$ is a saddle point.

For $(-2,2)$ we have $D(-2,2) = (12-12)^2 - 36 \cdot 4 = -144$ and so $(-2,2)$ is also a saddle point.

Finding Global Extrema

In general a local extremum doesn't have to be a global extremum. You might like to say that the largest local maximum is a global maximum, and the smallest local minimum is a global minimum, but this isn't necessarily the case. [Insert picture here.]

What we'd like is some condition that will guarantee that a function does in fact have global extrema, and furthermore say that the largest local maximum is a global maximum. Again, this doesn't have to happen in general, so we're looking for special cases where this does happen.

What prevents us from making such a statement is that the domain of a function could be infinite, and the outputs of the function could become arbitrarily large. So what we'd like to do is to consider functions whose domain is "small enough" that this sort of thing can't happen. Unfortunately this is a little bit more involved than you might think, so first we need to make a few definitions.

We'll say that a subset $S \subseteq \mathbb{R}^2$ of the plane is *bounded* if points of S can't become arbitrariliy far away from the origin. Equivalently, all of S can be fit into a circle of finite radius.

A point P is called a *boundary point* of a set $S \subseteq \mathbb{R}^2$ if every circle centered at P contains points that are both inside of and outside of S. The collection of all boundary points of S is called the *boundary* of S and denoted ∂S.

Finally we'll say that a set S which contains all of its boundary points is *closed*.

For example, the set $S = \{(x, y) | \sqrt{x^2 + y^2} \leq 1\}$ is closed, but the set $S = \{(x, y) | \sqrt{x^2 + y^2} < 1\}$ is not closed.

Remark.
If the complement of S is closed (that is, if the collection of all points that are in $\mathbb{R}^2$ but not in S is closed), then we say that S is *open*. Equivalently, for every point in S, we can put a small disc around that point, and the disc lies entirely in S.

Notice that there are sets that are neither open nor closed, as well as sets that are both open and closed.

If a set is both closed and bounded, we say the set is *compact*. Intuitively, compact sets are "small" and particularly small enough to ensure that any function whose domain is compact has global extrema.

Theorem 3.18 (Extreme Value Theorem)**.**
If $f(x, y)$ is a continuous function, and if the domain of f is compact, then f has both a global maximum and a global minimum.

3.9 Lagrange multipliers

In this lecture we'll describe a way of solving certain optimization problems subject to constraints. This method, known as *Lagrange multipliers*,

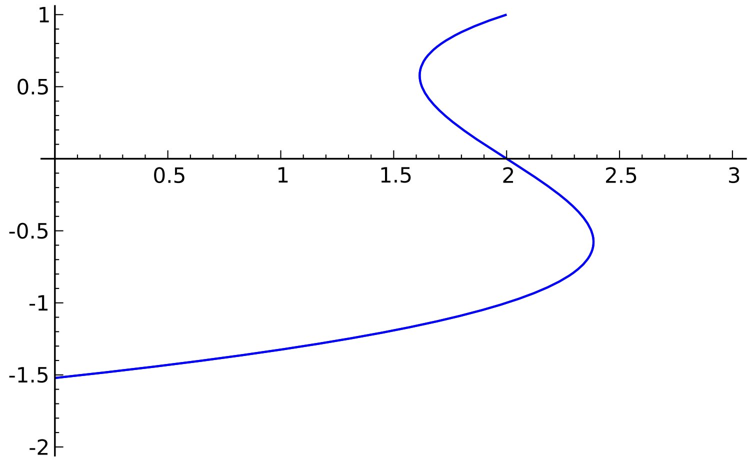


Figure 3.5: The constraint $x - y^3 + y = 2$.

gives us a way to algebraically solve such optimization problems and, unlike the previously described simplex algorithm, doesn't require that our function or constraints be linear.

Geometric Motivation

Suppose that we want to solve the following problem:

$$\begin{array}{ll} \text{minimize} & xy \\ \text{subject to} & x - y^3 + y = 2 \end{array}$$

That is, we want to find the (x, y) that makes xy as small as possible, but we only are conscerned with the points that satisfy $x - y^3 + y = 2$. To get an idea for how to do this, suppose that we first graph the points satisfying $x - y^3 + y = 2$. See Figure 3.5.

Our goal is to find the point on this curve, $x - y^3 + y = 2$, which makes xy as small as possible. To do this we'll consider curves of the form $xy = c$, where c is some constant, and try to make this value c (which is what we're trying to minimize) as small as possible *and still intersect the curve* $x - y^3 + y = 2$. Let's begin by setting $c = -1$; so we plot the curve $xy = -1$. This gives us the red curve in Figure 3.6.

Notice that the red curve touches the blue curve in exactly two spots. The coordinates of these intersection points tell us which (x, y) we can plug in for xy to get the value -1 and still satisfy the constraint $x - y^3 + y = 2$.

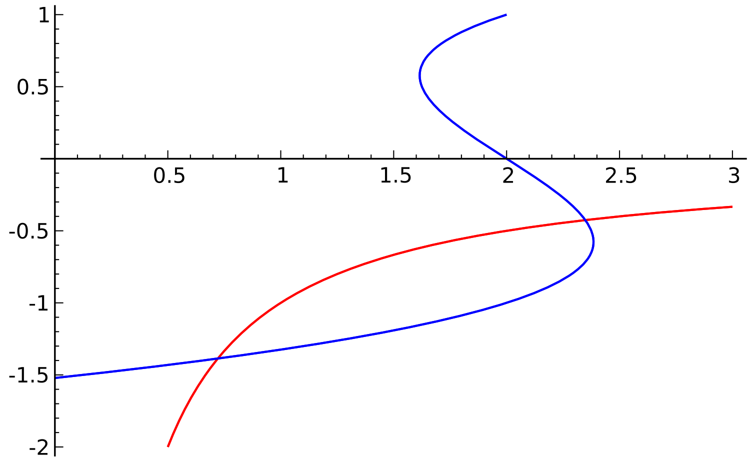


Figure 3.6: The constraint curve together with the curve $xy = -1$.

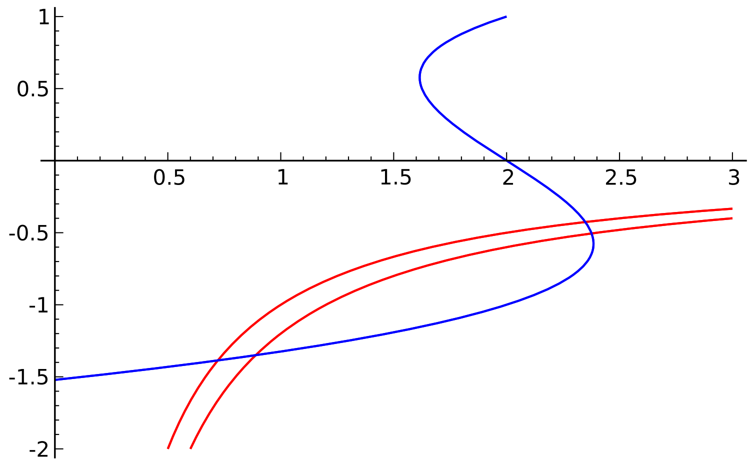


Figure 3.7: The constraint curve together with the curves $xy = -1$ and $xy = -1.2$.

Since we want to make c as small as possible and still intersect the blue curve, let's repeat the above process with the value $c = -1.2$. So we plot $xy = -1.2$. To compare with the previous plot, for $c = -1$, both curves are shown together in Figure 3.7.

Continuing this process, let's plot the curve $xy = c$ for lots of values of c, and then try to find which of these c's is as small as possible with the curve $xy = c$ still intersecting our constraint curve $x - y^3 + y = 2$. This is shown in Figure 3.8.

From the picture it appears that $c = -2$ is the smallest we can make

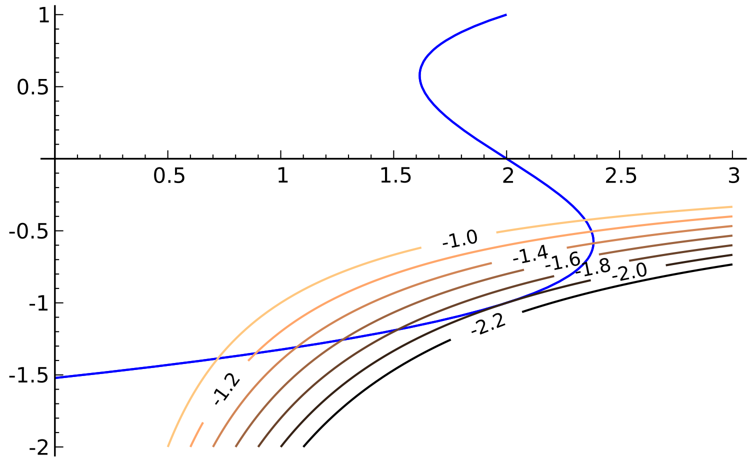


Figure 3.8: The curves $xy = c$ plotted for several values of c

$xy = c$ and still intersect the curve $x - y^3 + y = 2$. To find the (x, y)-coordinates of the point where this happens we'd want to solve the system $xy = 2$ and $x - y^3 + y = 2$.

Let's make a few simple observations about the pictures we've constructed above. The minimal value occured when the curves $xy = c$ and $x - y^3 + y = 2$ were tangent. This isn't a coincidence: if our value of c was "almost" optimal, then we'd have two very close points of intersection. Pushing c a little bit closer to the optimal value makes the points move closer together, and when we get to the optimal value the two points collide and we have a point of tangency. Notice that this is *not* the same as saying there is exactly one point where the optimal value occurs, nor does it guarantee that a point of tangency is a global max/min.

Even though we were solving a minimization problem above, we could try to solve the corresponding maximization problem in the exact same way: keep increasing c until we find the largest c so that the curves $x - y^3 + y = 2$ and $xy = c$ are tangent.

Gradients

In order to solve these problems in a precise way, we need to make use of the gradient of a multivariable function. Given a differentiable function $f(x, y)$, the *gradient* of $f(x, y)$ is the (multivariable) vector-valued function whose components are the partial derivatives of $f(x, y)$. This function is denoted $\nabla f(x, y)$ and is sometimes pronounced "del of $f(x, y)$":

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle .$$

Example 3.44.
Calculate the gradient of $f(x, y) = 3xy^2 - \sin(x) + y^x$.

$$\begin{aligned}\nabla f(x, y) &= \left\langle \frac{\partial}{\partial x}\left(3xy^2 - \sin(x) + y^x\right), \frac{\partial}{\partial y}\left(3xy^2 - \sin(x) + y^x\right)\right\rangle \\ &= \left\langle 3y^2 - \cos(x) - y^x \ln(y), 6xy - xy^{x-1}\right\rangle.\end{aligned}$$

We'll see the gradient several times through the semester, particularly in the next lecture when we talk about directional derivatives. Right now we care about the gradient because of the following theorem:

Theorem 3.19.
If $f(x, y)$ is differentiable, then any vector tangent to the curve $f(x, y) = c$ at (x_0, y_0) is orthogonal to $\nabla f(x_0, y_0)$.

Lagrange Multipliers

Now we need some way of taking the intuitive ideas above and putting them in a more formal framework. So let's suppose we have a more general optimization problem:

$$\begin{array}{ll}\text{maximize} & f(x, y) \\ \text{subject to} & g(x, y) = k\end{array}$$

So of all the (x, y) pairs that satisfy $g(x, y) = k$, we want to pick the pair that makes $f(x, y)$ as big as possible. In the motivating example above we considered curves $f(x, y) = c$ and tried to make c as large as possible while still intersecting $g(x, y) = k$. When this happened, the curves $g(x, y) = k$ and $f(x, y) = c$ were tangent. So in general what we want to do is find the (x, y) so that the tangent vectors of $g(x, y) = k$ and $f(x, y) = c$ are parallel.

Thus if $f(x, y)$ and $g(x, y)$ are both differentiable, then at any point where the curves $f(x, y) = c$ and $g(x, y) = k$ are tangent, the gradient vectors $\nabla f(x, y)$ and $\nabla g(x, y)$ have to be parallel. By the above theorem, to find these points of tangency, and hence to find the maximum (or minimum) of $f(x, y)$ subject to $g(x, y) = k$, we need to find the (x, y) pairs that simultaneously solve $\nabla f(x, y) = \lambda \nabla g(x, y)$ (because these vectors are parallel; they are scalar multiples) and $g(x, y) = k$ (because the points have to satisfy our constraint). This turns an optimzation problem into a problem of algebra: solving a system of equations. This is known as the *method of Lagrange multipliers*. (The "Lagrange multiplier" is the value λ above.)

$$\begin{array}{ll} \text{maximize} & f(x,y) \\ \text{subject to} & g(x,y) = k \end{array} \quad \leftrightarrow \quad \begin{array}{l} \text{Solve the system:} \\ \nabla f(x,y) = \lambda \nabla g(x,y) \\ g(x,y) = k \end{array}$$

Notice that $\nabla f(x, y) = \lambda \nabla g(x, y)$, when written out in components, really is just an algebraic system of equations.

$$\begin{array}{ll} \text{maximize} & f(x,y) \\ \text{subject to} & g(x,y) = k \end{array} \quad \leftrightarrow \quad \begin{array}{l} \text{Solve the system:} \\ f_x(x,y) = \lambda g_x(x,y) \\ f_y(x,y) = \lambda g_y(x,y) \\ g(x,y) = k \end{array}$$

Example 3.45.
Use the method of Lagrange multipliers to solve the following optimization problem:

$$\begin{array}{ll} \text{maximize} & 5x - 3y \\ \text{subject to} & x^2 + y^2 = 136 \end{array}$$

Here $f(x, y) = 5x - 3y$, $g(x, y) = x^2 + y^2$, and $k = 136$. First we find our gradients:

$$\nabla f(x,y) = \langle 5, -3 \rangle$$
$$\nabla g(x,y) = \langle 2x, 2y \rangle$$

And so we want to solve the following system of equations:

$$\begin{aligned} 5 =& \lambda 2x \\ -3 =& \lambda 2y \\ x^2 + y^2 =& 136 \end{aligned}$$

Solving the first two equations for x and y, we have

$$\begin{aligned} x =& \frac{5}{2\lambda} \\ y =& \frac{-3}{2\lambda}. \end{aligned}$$

Notice that there's no possible way for λ to equal zero in our problem because of the equation $5 = \lambda 2x$; if we did have $\lambda = 0$, this equation would give us $5 = 0$, which is certainly not true.

Plugging these into the constraint $x^2 + y^2 = 136$, we have

$$\begin{aligned} \frac{25}{4\lambda^2} + \frac{9}{4\lambda^2} &= 136 \\ \implies \frac{34}{4\lambda^2} &= 136 \\ \implies \lambda^2 = \frac{34}{544} &= \frac{1}{16} \end{aligned}$$

Thus $\lambda = \pm 1/4$.

If $\lambda = 1/4$, the first two equations become $5 = \frac{x}{2}$ and $-3 = \frac{y}{2}$, or $x = 10$ and $y = -6$. If $\lambda = -1/4$, then we'd have $x = -10$ and $y = 6$.

So there are two solutions to our system of equations: $(10, -6)$ and $(-10, 6)$. One of these is the maximum, and one is the minimum. To see which is which we have to plug them into our original function. Plugging in $(10, -6)$ gives $50 + 18 = 68$; plugging in $(-10, 6)$ gives $-50 - 18 = -68$. Thus $(10, -6)$ is the max, and $(-10, 6)$ is the min.

(The "symmetry" in the solutions here is not typical of these sorts of problems, but arises from the geometry of our optimization problem. Our constraint is a circle centered at the origin, while the objective function we're trying to optimize is linear. This means the coordinates of the max and min will be opposite points on the circle.)

As usual, there's nothing really special about the fact that we're using functions two variables above. We could just as easily use Lagrange multipliers to solve optimization problems with several variables.

maximize	$f(x_1, x_2, ..., x_n)$	$\leftrightarrow$	Solve the system:
subject to	$g(x_1, x_2, ..., x_n) = k$		$\nabla f(x_1, x_2, ..., x_n) = \lambda \nabla g(x_1, x_2, ..., x_n)$
			$g(x_1, x_2, ..., x_n) = k$

Example 3.46.
Find the dimensions of the largest box that can be made with 64 square feet of cardboard.

Suppose the dimensions of our box are x, y, and z. Then we want to maximize $V(x, y, z) = xyz$ subject to the constraint that $g(x, y, z) = 2xy + 2yz + 2xz = 64$. The gradient of our objective function is

$$\nabla V(x, y, z) = \langle yz, xz, xy \rangle .$$

The gradient of our constrant function is

$$\nabla g(x, y, z) = \langle 2y + 2z, 2x + 2z, 2y + 2x \rangle .$$

Thus our system of equations is

$$\begin{aligned} yz =& \lambda(2y + 2z) \\ xz =& \lambda(2x + 2z) \\ xy =& \lambda(2y + 2x) \\ 2xy + 2yz + 2xz =& 64. \end{aligned}$$

To solv this equation notice that if we multiply the first equation by x, the second equation by y, and the third equation by z we have:

$$\begin{aligned} xyz =& \lambda x(2y + 2z) \\ xyz =& \lambda y(2x + 2z) \\ xyz =& \lambda z(2y + 2x) \end{aligned}$$

Each of the right hand sides equals xyz, so the right hand sides are all equal:

$$\lambda x(2y + 2z) = \lambda y(2x + 2z) = \lambda z(2y + 2x).$$

Now notice that for our problem λ can never be zero. If λ was zero, then $\nabla V(x, y, z) = 0$ at the maximum point. However, the only way $\nabla V(x, y, z) = \langle yz, xz, xy \rangle$ is zero is if at least two of our coordinates are zero. Since we're talking about the volume of a box, our coordinates are all strictly positive and this can't happen.

Since $\lambda \neq$ we can divide out the λ's, and also divide out the 2's to get

$$xy + xz = xy + yz = yz + xz.$$

For the first equation, $xy + xz = xy + yz$, subtracting xy from each side gives $xz = yz$. Now dividing by z (again, $z \neq 0$ because the coordinates are all positive) gives $x = y$. Similarly, the second equation, $xy + yz = yz + xz$ gives $y = z$. So $x = y = z$.

Plugging $y = x$ and $z = x$ into our constraint equation gives

$$\begin{aligned} &2x^2 + 2x^2 + 2x^2 = 64 \\ \implies &6x^2 = 64 \\ \implies &x^2 = \frac{64}{6} = \frac{32}{3} \\ \implies &x = \pm\sqrt{\frac{32}{3}} \end{aligned}$$

Again, all of our coordinates are positive because of the physical interpretation of our problem, so $x = y = z = \sqrt{32/3}$ are the dimensions that maximize the volume of a cube that can be made with 64 square feet of material.

Example 3.47.
Write down the system of equations, but do not solve (it's difficult!), for the following optimization problem:

$$\begin{array}{ll} \text{maximize} & 1 + x^4 + y^4 - 4xy \\ \text{subject to} & x^2 + y^3 = 2. \end{array}$$

Here $f(x, y) = 4xy - x^4 - y^4 - 1$, $g(x, y) = x^2 + y^3$ and $k = 2$. Thus our gradient vectors are

$$\nabla f(x, y) = \left\langle 4y - 4x^3, 4x - 4y^3 \right\rangle$$
$$\nabla g(x, y) = \left\langle 2x, 3y^2 \right\rangle$$

So the system of equations we wish to solve is

$$4y - 4x^3 = \lambda 2x$$
$$4x - 4y^3 = \lambda 3y^2$$
$$x^2 + y^3 = 2$$

4 Integration

If one looks at the different problems of the integral calculus which arise naturally when one wishes to go deep into the different parts of physics, it is impossible not to be struck by the analogies existing.

HENRI POINCARÉ

4.1 Integration in two variables

We now begin the study of integration in several variables. Motivated by the geometric problem of finding the volume between a surface $z = f(x, y)$ and the xy-plane, the ideas and geometric intuition of integration in one variable easily extend to two variables. However, as we will see, integration is about much more than simply calculating areas and volumes, and has applications throughout mathematics and physics.

Review of Integration in One Variable

Recall that if $f : [a, b] \to \mathbb{R}$ is a continuous function, the (*definite*) *integral* of f is defined as a limit of Riemann sums. In particular, we choose a partition $\mathcal{P}$ of $[a, b]$:

$$\mathcal{P} = \{x_0, x_1, x_2, ..., x_n\},$$
$$\text{where } a = x_0 < x_1 < x_2 < ... < x_n = b.$$

A *Riemann sum* of f with respect to the partition $\mathcal{P}$ is the quantity

$$\sum_{i=1}^{n} f(x_i^*)\Delta x_i,$$

where $\Delta x_i = x_i - x_{i-1}$ (this is the length of the i-th subinterval in the partition) and x_i^* is any point in $[x_{i-1}, x_i]$. Obviously the value of this sum depends on the choice of $\mathcal{P}$ and the choice of each x_i^*. Incredibly, if we take the limit as the pieces of the partition get arbitrarily small, we always get the same value, regardless of the $\mathcal{P}$ and x_i^*'s we choose in calculating each of the Riemann sums.

Writing $|\mathcal{P}| = \max_i \Delta x_i$ (so $|\mathcal{P}|$ is the length of the widest subinterval determined by $\mathcal{P}$) the integral of f over $[a, b]$ is defined as

$$\int_a^b f(x)\,dx = \lim_{|\mathcal{P}|\to 0} \sum_{i=1}^{n_{\mathcal{P}}} f(x_i^*)\Delta x_i.$$

The number of terms in the sum depends on the partition $\mathcal{P}$ we choose. Here we're letting $n_{\mathcal{P}}$ denote the number of subintervals into which $[a, b]$ is partitioned into by $\mathcal{P}$.

Since we're taking a limit, we always have to ask ourselves if this limit exists or not. It is a theorem (that we won't try to prove) that this limit will always exist if f is continuous and $[a, b]$ is a closed, bounded interval.

When you first learn about integration, you build the integral up for a singular purpose: to find the area between some curve $y = f(x)$ and the x-axis. The idea is to approximate the area under the curve with things that are much simpler to work with: rectangles. In the Riemann sum, the value $f(x_i^*)$ acts as the height of the i-th rectangle, and the Δx_i is the width. So we calculate the area of each rectangle and add them all up.

Of course, you realize very quickly that integrals can do much more than simply calculate areas. Integration is ubiquitous in mathematics: from geometry to statistics to physics, integrals are everywhere. The reason that integrals are such a useful tool is that they can be thought of as a very special type of infinite summation. The integral $\int_a^b f(x)\,dx$ is, in some sense, the result of summing up the values of $f(x)$ for every single x in $[a, b]$; it's just that we weight the values in the sum by a very small number (this is basically what the dx is) to keep this "sum" from blowing up to infinity.

Calculating Volumes

To motivate integration in several variables, consider the following problem. Suppose that $f : D \to \mathbb{R}$ is a continuous function and that D, the domain of f, is a rectangle in $\mathbb{R}^2$. Suppose also that $f(x, y) \geq 0$ for all $(x, y) \in D$. We now construct a three-dimensional object whose top is $z = f(x, y)$, whose bottom is D, and we fill in all of the space in-between. See Figure 4.1 on the next page.

Now we want to determine what the volume of this solid is. To do this we do the same sort of thing we did to calculate the area under a curve: approximate the volume with simpler objects. The simpler objects we'll use are rectangular prisms. If a prism has height h, length ℓ, and width w, then we know its volume is $h\ell w$.

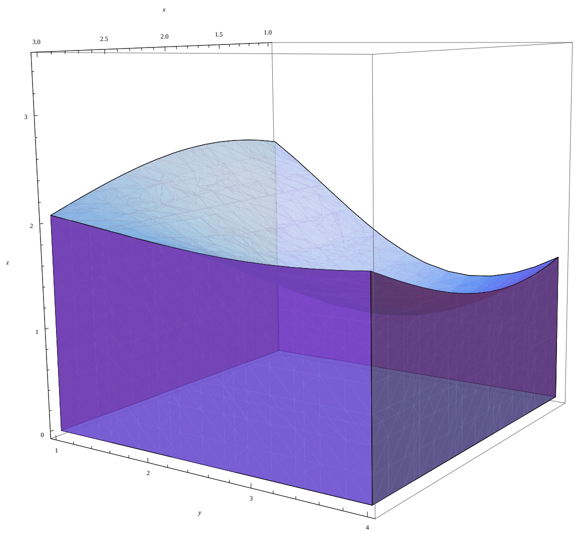


Figure 4.1: A three-dimension solid whose top is $z = f(x, y)$.

So what we'll do is cram several rectangular prisms under the surface $z = f(x, y)$, determine the volume of each prism, and then finally sum up these volumes. See Figure 4.2 on the following page.

Suppose we label these prisms $P_1, P_2, ..., P_n$ and let h_i, ℓ_i, and w_i denote the height, length, and width of each P_i. Then we know that the volume of our object is approximated by

$$\text{Volume} \approx \sum_{i=1}^{n} h_i w_i \ell_i.$$

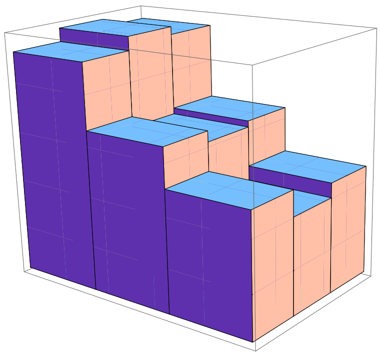
Of course what we want to do is take the limit as we fill the area under the curve with skinnier and skinnier prisms. In order to do this we need to state precisely how these these prisms are placed beneath the surface.

Suppose the four corners of the domain D are $(a, c), (b, c), (b, d), (a, d)$. See Figure 4.3. This means that the rectangle D consists precisely of those points (x, y) where $a \leq x \leq b$ and $c \leq y \leq d$. So we can express D as the set

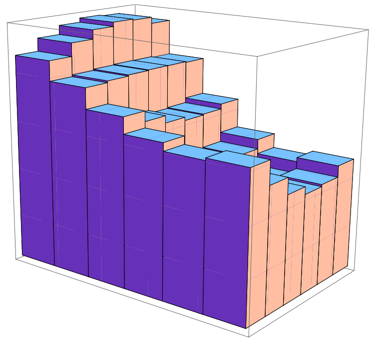
$$D = [a, b] \times [c, d] = \left\{ (x, y) \,\middle|\, a \leq x \leq b,\, c \leq y \leq d \right\}.$$

Our first step is to cut D into finitely many rectangular pieces; each piece will serve as the base of a rectangular prism. To do this we'll cut $[a, b]$ into subintervals by using the partition $\mathcal{P} = \{x_0, x_1, ..., x_m\}$ where

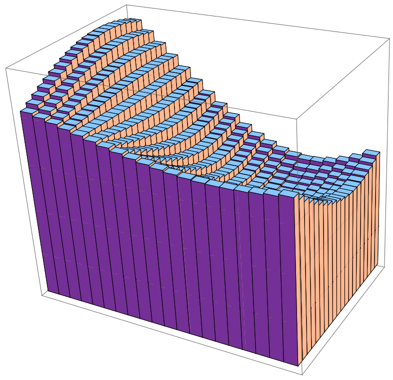
$$a = x_0 < x_1 < ... < x_m = b,$$



(a) 9 prisms



(b) 36 prisms



(c) 400 prisms

Figure 4.2: Approximating volume with prisms.

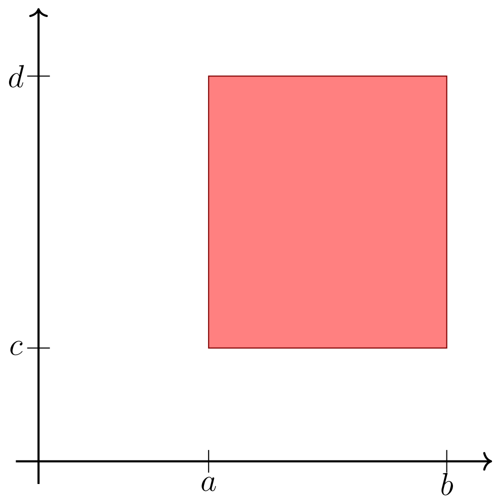


Figure 4.3: The rectangle D is the domain of our function.

and we'll cut $[c, d]$ into subintervals with the partition $\mathcal{Q} = \{y_0, y_1, ..., y_n\}$ where

$$c = y_0 < y_1 < ... < y_n = d.$$

This partitions D into mn subrectangles. We'll let the rectangle in the i-th column and j-th row (ordered left-to-right, bottom-to-top) be denoted D_{ij}. See Figure 4.4.

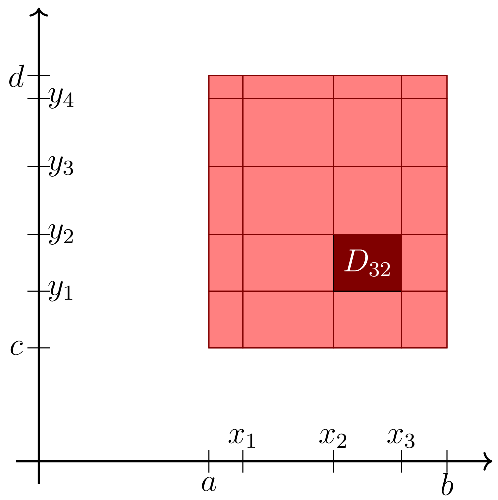


Figure 4.4: The rectangle D is partitioned into subrectangles.

Let's denote the area of the rectangle D_{ij} by ΔA_{ij}. (Notice $\Delta A_{ij} = \Delta x_i \cdot \Delta y_j$.)

Now that we have bases for our rectangular prisms, we just need to determine their height. To do this we let P_{ij}^* denote any point inside of D_{ij}, and then use $f(P_{ij}^*)$ as the height of the prism. Notice that since the x-coordinates of P_{ij}^* are in the i-th subinterval of $[a, b]$, and the y-coordinates of P_{ij}^* are in the j-th subinterval of $[c, d]$, we have $P_{ij}^* = (x_i^*, y_j^*)$. Thus the volume of the ij-th prism is

$$f(x_i^*, y_j^*)\Delta A_{ij}.$$

Summing up the volumes of each of these prisms, we have an estimate for the volume of our solid:

$$\text{Volume} \approx \sum_{i=1}^{m} \sum_{j=1}^{n} f(x_i^*, y_j^*)\Delta A_{ij}.$$

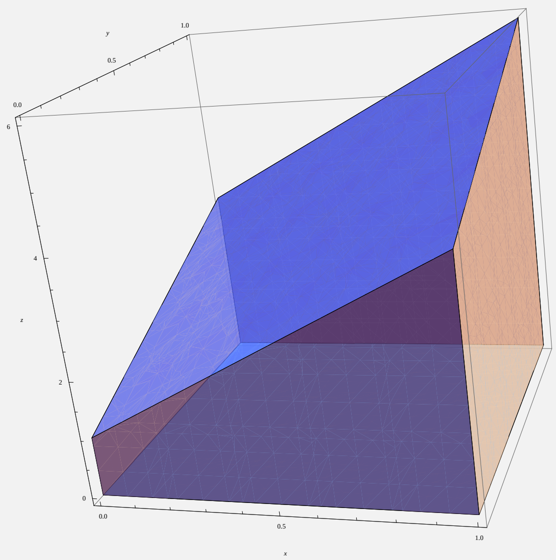
To get a better approximation, stick more, skinnier, prisms underneath the surface. To get the "best" approximation (i.e., the true volume), take the limit as the prisms get arbitrarily skinny. To do this we need that both the widths and lengths of our base rectangles get arbitrarily small. That is, we require $|\mathcal{P}| \to 0$ and $|\mathcal{Q}| \to 0$.

The limit as the rectangles get arbitrarily skinny is called the (*double*) *integral of f over the rectangle D* and is denoted as follows:

$$\iint\limits_D f(x,y)\,dA = \lim_{|\mathcal{P}|\to 0}\lim_{|\mathcal{Q}|\to 0}\sum_{i=1}^{m_{\mathcal{Q}}}\sum_{j=1}^{n_{\mathcal{P}}} f(x_i^*, y_j^*)\,\Delta A_{ij}.$$

As always, we have to worry about whether this limit exists or not. As in the case of one variable, there is a theorem that says that this limit will exist as long as our function f is continuous and D is a rectangle of finite area. (There are other, more general, conditions which guarantee the integral exists, but this is good enough for right now.)

Example 4.1.
Calculate the volume between the surface $z = 3x + 2y + 1$ and the xy-plane over the unit square, $D = [0,1] \times [0,1]$.



To make this process as easy as possible, let's say that we partition both the the horizontal and vertical intervals $[0,1]$ into n subintervals of equal width, and use the upper, right-hand corner of each rectangle as the point where we'll evaluate the function to determine the height of a prism. (As long as our function is continuous we'll get the same value in the end, so we can pick points that are easy to work with.)

This gives us n^2 subrectangles, each of area $1/n^2$, and $x_i^* = 1/n$, $y_j^* = 1/n$. Thus our volume is given by the limit:

$$\iint_{[0,1]\times[0,1]} (3x + 2y + 1)\, dA = \lim_{n\to\infty} \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} (3x_i^* + 2y_j^* + 1) \cdot \frac{1}{n^2}$$

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

Notice that since we cut both the horizontal and vertical intervals into n pieces we have $m = n$, which is why we have two limits as $n \to \infty$ on the first line. Of course, taking the first (inner) limit gives us a number, and so taking the second (outer) limit doesn't do anything, so we can drop one of the limits.

$$\begin{aligned}
\iint\limits_{[0,1]\times[0,1]} (3x+2y+1)\,dA &= \lim_{n\to\infty}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{3i}{n^3}+\frac{2j}{n^3}+\frac{1}{n^2}\right)\\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{3i}{n^3}+\sum_{j=1}^{n}\frac{2j}{n^3}+\sum_{j=1}^{n}\frac{1}{n^2}\right)\\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\left(\frac{3i}{n^2}+\frac{1}{n}+\frac{2}{n^3}\sum_{j=1}^{n}j\right)\\
&= \lim_{n\to\infty}\sum_{i=1}^{n}\left(\frac{3i}{n^2}+\frac{1}{n}+\frac{2}{n^3}\cdot\frac{n^2+n}{2}\right)\\
&= \lim_{n\to\infty}\left(\sum_{i=1}^{n}\frac{3i}{n^2}+\sum_{i=1}^{n}\frac{1}{n}+\sum_{i=1}^{n}\frac{n^2+n}{n^3}\right)\\
&= \lim_{n\to\infty}\left(\frac{3}{n^2}\sum_{i=1}^{n}i+1+\frac{n^2+n}{n^2}\right)\\
&= \lim_{n\to\infty}\left(\frac{3}{n^2}\cdot\frac{n^2+n}{2}+1+\frac{n^2}{n^2}+\frac{n}{n^2}\right)\\
&= \lim_{n\to\infty}\left(\frac{3}{2}\cdot\frac{n^2+n}{n^2}+2+\frac{1}{n}\right)\\
&= \lim_{n\to\infty}\left(2+\frac{3}{2}+\frac{3}{2n}+\frac{1}{n}\right)\\
&=\frac{7}{2}
\end{aligned}$$

Evaluating the limit above we used two facts that you learned in 106 and 108, but may have forgotten about: you can "distribute" summations:

$$\sum_i (a_i+b_i) = \sum_i a_i + \sum_i b_i,$$

and there's a nice formula for the sum $1+2+3+\cdots+n$:

$$\sum_{i=1}^{n} i = \frac{n^2+n}{2}.$$

So the volume between the plane $3x+2y+1$ and the unit square in the xy-plane is $7/2$.

Of course, evaluating a limit such as the one above is a rather tedious thing to do. It'd be nice we if we had some way to turn these complicated double integrals into "normal" integrals of one variable where we could use tools such as integration by parts and u-substitutions. We'll see how this can be done later. For the time being we'll just be content with the fact that we can, at least in principle, evaluate these double integrals by taking limits.

Properties of the Integral

The double integral satisfies several properties analogous to properties of integrals of single variables. Here we mention a few of the most basic ones.

(i) Letting Area(D) denote the area of a rectangle D,

$$\iint_D 1\,dA = \text{Area}(D).$$

This is straight forward to see: the double sum we're taking the limit of is just

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\Delta A_{ij},$$

where ΔA_{ij} is the area of the ij-th subrectangle. However we're summing this over all of the subrectangles, so we just get back the area of D.

Note: Notationally, we sometimes write

$$\iint_D dA = \iint_D 1\,dA.$$

(ii) If $\lambda \in \mathbb{R}$ is a constant and $f : D \to \mathbb{R}$ is continuous, then

$$\iint_D \lambda f(x,y)\,dA = \lambda \iint_D f(x,y)\,dA.$$

This follows from the fact that we can pull the constant λ out of the sums and limits in the definition of the integral.

(iii) If $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are both continuous, then

$$\iint\limits_D (f(x,y) + g(x,y))\, dA = \iint\limits_D f(x,y)\, dA + \iint\limits_D g(x,y)\, dA.$$

This again just follows from the fact that we can split sums and limits up across addition. Notice that this, combined with the previous property, means that we can also split up a subtraction: write $f(x,y) - g(x,y) = f(x,y) + (-1) \cdot g(x,y)$.

(iv) If $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are both continuous and $f(x,y) \leq g(x,y)$ for every $(x,y) \in D$, then

$$\iint\limits_D f(x,y)\, dA \leq \iint\limits_D g(x,y)\, dA.$$

Once more, this follows by the analogous properties for sums and limits.

Applications

Even though we built the double integral above for the purpose of calculating volumes, it's clear that the definition still "makes sense" for functions which may be negative. If the function is negative – i.e., if the surface $z = f(x,y)$ is below the xy-plane – then the corresponding integral we calculate will be negative. This is similar to how $\int_a^b f(x)\, dx < 0$ if $y = f(x)$ stays below the x-axis. In general a function may be above the xy-plane sometimes, and below the xy-plane at other times. When this happens, the integral brakes up into we have positive and negative pieces, and may cancel out. In these cases the double integral doesn't represent a volume, but may still have a concrete, physical meanings.

After we've developed some more tools for calculating integrals, we'll consider the more applications in detail, but it's worthwhile to go ahead and mention some of the things these double integrals can be used for:

Mass of an object

If we have a rectangular "sheet" of some material (metal, plastic, cloth, ...), and if we know the know what the density of this material is at any point, integrating the density gives us the mass of the object. Say our rectangular "sheet" is $w \times \ell$. We can think of this as the rectangle $[0,w] \times [0,\ell]$ in the xy-plane. For any point (x,y)

inside the rectangle, suppose that $\rho(x, y)$ represents the density of the material at that particular point. Then the mass of the sheet is

$$m = \iint\limits_{[0,w]\times[0,\ell]} \rho(x, y)\, dA.$$

If you consider how density is defined in physics, this is almost obvious. Density, in two dimensions, is mass divided by area: $\rho = \frac{m}{A}$. So over a very small subrectangle, D_{ij}, the density is approximately

$$\rho_{D_{ij}} \approx \frac{\text{mass of } D_{ij}}{\text{area of } D_{ij}} = \frac{\text{mass of } D_{ij}}{\Delta A_{ij}}.$$

When we write out the limit of Riemann sums, the ΔA_{ij}'s cancel out and we're just summing up the mass of little pieces of D.

Average value

If $f : D \to \mathbb{R}$ is a continuous function on a rectangle D, then there may be times we want to know what the average value of f is. For example, suppose that D represents the floor in a room, and for each point in the room, the temperature you record at that point is determined by where you're standing in the room – by your xy-coordinates on the floor. (Of course, the temperature in a real room may also depend on how high above the floor you are.) If $T(x, y)$ gives the temperature over the point (x, y), then the average temperature in the room is

$$\text{Avg. temp} = \frac{1}{\text{Area(D)}} \iint\limits_D T(x, y)\, dA.$$

Why is this the average temperature? If the temperature throughout the entire room was constant, say $T(x, y) = C$, then we'd say that the average temperature in the room was C. So to estimate the average temperature, let's partition the room into very small rectangles and suppose that the temperature is constant on each of those rectangles (possibly a different constant on different rectangles).

Now if we wanted to combine all of these average temperatures over small regions together to get the average temperature of the whole region, we'd have to weight those averages by the relative size of the region; that is, by how much proportion of the room is taken up by that region. (Why? Because a 1-inch $\times$ 1-inch region where the temperature is $90°F$ doesn't contribute as much to the average as a 10-ft$\times 10$-ft region where the temperature is $90°F$. If the temperature is really warm over a large region, that counts a lot more for the average than being really warm over a very small region.)

Let's suppose that we call the subrectangles of our partition D_{ij}, with area ΔA_{ij}. The proportion of the room taken up by D_{ij} is $\frac{\Delta A_{ij}}{\text{Area}(D)}$. Say the temperature we use for the constant on D_{ij} is $T(x_i^*, y_j^*)$. So we sum up the values

$$\sum_{i=1}^{m}\sum_{j=1}^{n} T(x_i^*, y_j^*)\frac{\Delta A_{ij}}{\text{Area}(D)}$$

$$=\frac{1}{\text{Area}(D)}\sum_{i=1}^{m}\sum_{j=1}^{n} T(x_i^*, y_j^*)\Delta A_{ij}.$$

Taking the limit gives exactly the integral described above. Notice that it makes perfect sense to talk about an average temperature being negative!

Of course, there's nothing special about the fact that we're talking about temperature above. In general, the *average value* of a continuous function over a rectangle D is

$$\text{Average of } f = \frac{1}{\text{Area}(D)}\iint_D f(x,y)\,dA.$$

4.2 Iterated integrals

In this section we continue studying integration in two variables and introduce *iterated integrals*, which are the primary tool used for calculating integrals of several variables.

Motvation & "Partial Integration"

In the last lecture we defined the double integral of a continuous function as a limit of a double Riemann sum. While this definition makes intuitive sense (approximating a quantity with simpler quantities and taking a limit to get the "best" approximation), it's typically extremely difficult and tedious to use for calculations. Now we want to introduce a way of calculating these quantities which will allow us to apply the tools and techniques from integration in one variable. Before we describe how this is done, we need to make one technical detour.

Recall that the partial derivatives, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, are calculated by keeping one variable constant and differentiating with respect to the other variable. Suppose we instead want a "partial antiderivative" of a function. That is, suppose that $f(x, y)$ is a given function. Can we find functions $G(x, y)$ and $H(x, y)$ such that $\frac{\partial G}{\partial x} = f(x, y)$ and $\frac{\partial H}{\partial y} = f(x, y)$? If we were considering functions of a single variable, then we'd just integrate the function to get its antiderivative. Since partial derivatives are calculated by keeping one variable constant, these "partial antiderivatives" can be calculated the same way: integrate the function by keeping one variable constant.

That is, to calculate $G(x, y)$ we'll integrate $f(x, y)$ with respect to x, pretending that the y in our function is a constant. Similarly, to calculate $H(x, y)$ we integrate $f(x, y)$ with respect to y, pretending x is constant. This is denoted as follows:

$$G(x, y) = \int f(x, y)\, dx$$
$$H(x, y) = \int f(x, y)\, dy$$

There is one caveat here: when we calculate $\int f(x, y)\, dx$ instead of picking up a $+C$, we pick up a $+k(y)$. That is, since y's are constant when we calculate $\frac{\partial G}{\partial x}$, any function of y is also constant. So our $+C$ can be any expression that involves only y's: from the partial derivative point of view these are functions. Similarly, when we calculate $\int f(x, y)\, dy$, we pick up a $+\ell(x)$.

Example 4.2.
Find a $G(x, y)$ such that $\frac{\partial G}{\partial x} = x^2 y - \sin(xy)$. Find a $H(x, y)$ such that $\frac{\partial H}{\partial y} = x^2 y - \sin(xy)$.

We simply integrate, pretending one variable or the other is a constant.

$$\begin{aligned} G(x,y) &= \int \big(x^2y - \sin(xy)\big)\ dx \\ &= \frac{x^3y}{3} + \frac{\cos(xy)}{y} + k(y) \end{aligned}$$

$$\begin{aligned} H(x,y) &= \int \big(x^2y - \sin(xy)\big)\ dy \\ &= \frac{x^2y^2}{2} + \frac{\cos(xy)}{x} + \ell(x) \end{aligned}$$

Now let's just double check that these are the functions we want:

$$\begin{aligned} \frac{\partial G}{\partial x} &= \frac{\partial}{\partial x}\left(\frac{x^3y}{3} + \frac{\cos(xy)}{y} + k(y)\right) \\ &= \frac{3x^2y}{3} + \frac{-\sin(xy)\,y}{y} + 0 \\ &= x^2y - \sin(xy) \end{aligned}$$

$$\begin{aligned} \frac{\partial H}{\partial y} &= \frac{\partial}{\partial y}\left(\frac{x^2y^2}{2} + \frac{\cos(xy)}{x} + \ell(x)\right) \\ &= \frac{2x^2y}{2} + \frac{-\sin(xy)\,x}{x} + 0 \\ &= x^2y - \sin(xy) \end{aligned}$$

Iterated Integrals

To calculate a double integral,

$$\iint\limits_{[a,b]\times[c,d]} f(x,y)\,dA,$$

we will convert the double integrals into two integrals of a single variable, combined together in a particular way. The basic idea is the following: Geometrically, double integrals were developed for calculating volumes. However these is another way to calculate volumes, provided that you know the cross-sectional areas of the solid you're integrating.

Recall that if we have a solid positioned in three-dimensional space so that the x-axis runs through the solid, like a chicken on a rotisserie, then for each plane $x = c$ we denote the area of the intersection of the plane and the solid by $A(x)$. Then the volume of the solid is given by integrating $A(x)$:

$$\text{Volume} = \int_a^b A(x)\, dx$$

For example, in the last lecture we considered the volume of the solid whose top was the plane $3x + 2y + 1$, and whose bottom was the unit square $[0, 1] \times [0, 1]$. Cutting this surface with a plane we see the blue region plotted in Figure 4.5.

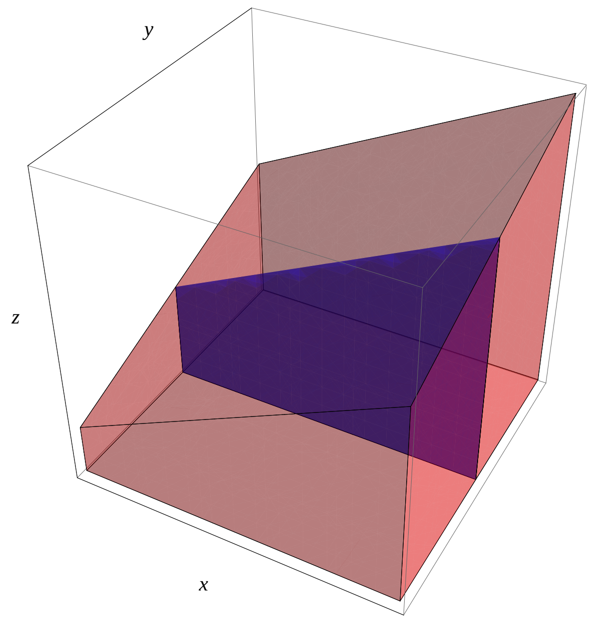


Figure 4.5: Cutting a solid with a plane.

If we could calculate the area, $A(x)$, of this blue slice of the surface, we'd integrate $\int_0^1 A(x)\, dx$ to find the volume. There's nothing special about the x axis here: we could just as easily consider slices of the surface given by planes $y = c$, call $A(y)$ the area of these slices, and then integrate $\int_0^1 A(y)\, dy$ to get the volume.

Calculating these cross-sectional areas is actually very easy because they're just the area under the curve. In particular, the area $A(x)$ is given by

$$A(x) = \int_0^1 (3x + 2y + 1)\ dy.$$

This is just the area of the blue slice because the blue slice is the area underneath the curve $3x + 2y + 1$. Here we've set x to be a constant, so y is the only quantity that changes. Performing the integration we see that this really is just a function of x: the y's get replaced with numbers when we do the integration.

$$\begin{aligned} A(x) &= \int_0^1 (3x + 2y + 1)\ dy \\ &= \left(3xy + y^2 + y\right)\Big|_0^1 \\ &= 3x + 2. \end{aligned}$$

This is the cross-sectional area of our blue slice. Integrating this quantity we get the volume.

$$\begin{aligned} \text{Volume} &= \int_0^1 A(x)\, dx \\ &= \int_0^1 (3x + 2)\ dx \\ &= \left(\frac{3x^2}{2} + 2x\right)\Bigg|_0^1 \\ &= \frac{3}{2} + 2 \\ &= {}^7\!/_2 \end{aligned}$$

Usually we don't bother to write down $A(x)$ as a separate function, and instead just plug our expression for $A(x)$,

$$\int_0^1 (3x + 2y + 1)\ dy,$$

into the integral:

$$\text{Volume} = \int_0^1 \int_0^1 (3x + 2y + 1)\ dy\, dx.$$

To evaluate an expression like this we work "inside-out," starting with the inner-most integral and integrating piece by piece until we've evaluated all of the integrals.

$$\begin{aligned}
\text{Volume} &= \int_0^1 \int_0^1 (3x + 2y + 1)\ dy\, dx \\
&= \int_0^1 \left(3xy + y^2 + 1\right)\Big|_0^1\ dx \\
&= \int_0^1 (3x + 2)\ dx \\
&= \left(\frac{3x^2}{2} + 2x\right)\Big|_0^1 \\
&= \frac{3}{2} + 2 \\
&= {}^7\!/_2
\end{aligned}$$

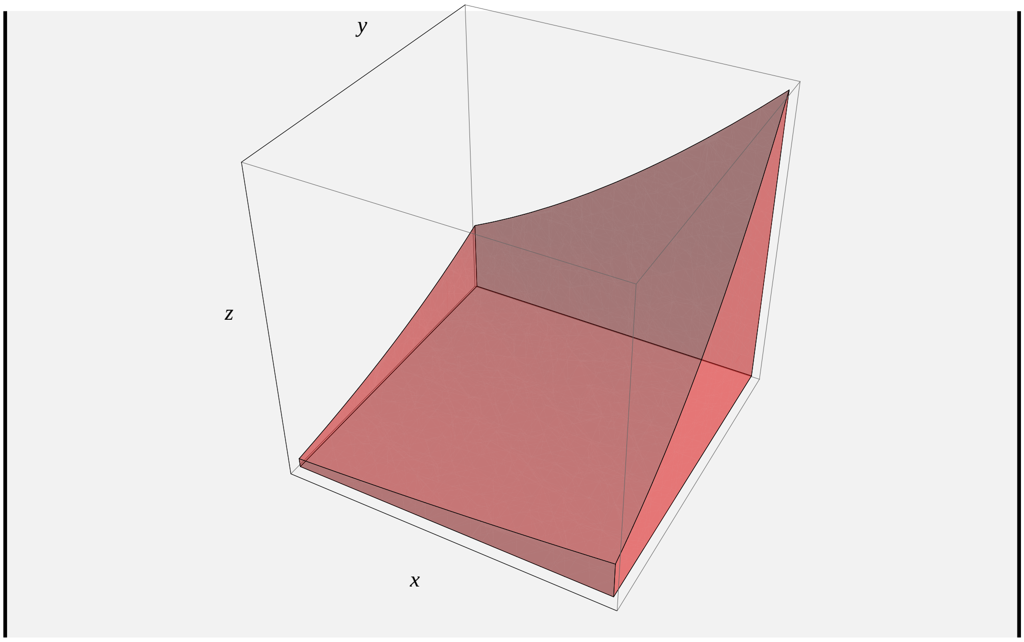
The procedure outlined above is generalized by the following theorem.

Theorem 4.1 (Fubini's theorem)**.**
If $f(x, y)$ is a continuous function defined on the rectangle $D = [a, b] \times [c, d]$, then

$$\iint_D f(x, y)\, dA = \int_a^b \int_c^d f(x, y)\, dy\, dx = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

Example 4.3.
Calculate the integral

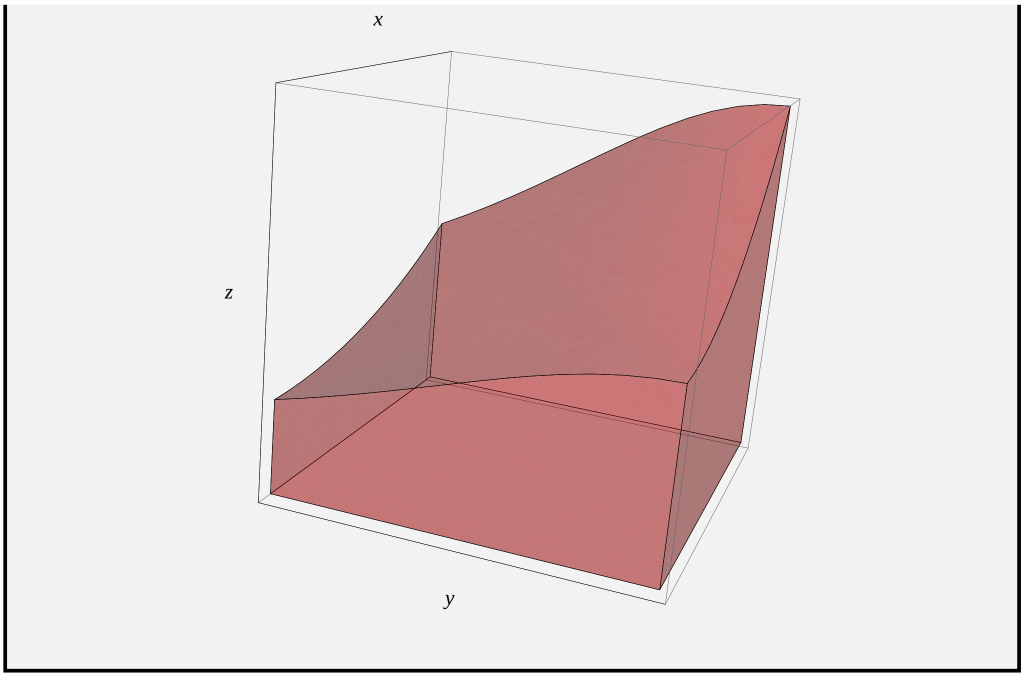
$$\iint_{[2,4]\times[1,2]} \frac{x^2\, y^3}{2}\, dA.$$


y
z
x

$$\begin{aligned}
\iint\limits_{[2,4]\times[1,2]} \frac{x^2\,y^3}{2}\,dA &= \frac{1}{2}\iint\limits_{[2,4]\times[1,2]} x^2\,y^3\,dA \\
&= \frac{1}{2}\int_2^4\int_1^2 x^2\,y^3\,dy\,dx \\
&= \frac{1}{2}\int_2^4 \left.\frac{x^2\,y^4}{4}\right|_1^2 dx \\
&= \frac{1}{2}\int_2^4 \left(4x^2 - \frac{x^2}{4}\right) dx \\
&= \frac{1}{2}\left.\left(\frac{4x^3}{3} - \frac{x^3}{12}\right)\right|_2^4 \\
&= \frac{1}{2}\left(\frac{4\cdot 64}{3} - \frac{64}{12} - \frac{8\cdot 4}{3} + \frac{8}{12}\right) \\
&= \frac{1}{2}\left(\frac{256}{3} - \frac{16}{3} - \frac{32}{3} + \frac{2}{3}\right) \\
&= \frac{210}{6} \\
&= \frac{105}{3} \\
&= 35
\end{aligned}$$

Example 4.4.
Calculate the integral

$$\iint\limits_{[0,1]\times[0,1]} \frac{1+x^2}{1+y2}\,dA$$



$$\iint\limits_{[0,1]\times[0,1]} \frac{1+x^2}{1+y^2}\, dA = \int_0^1 \int_0^1 \frac{1+x^2}{1+y^2}\, dx\, dy$$

Notice that, with respect to x, $\frac{1}{1+y^2}$ is a constant. Hence we can pull it out of the inner-most integral:

$$\int_0^1 \int_0^1 \frac{1+x^2}{1+y^2}\, dx\, dy = \int_0^1 \frac{1}{1+y^2} \int_0^1 \left(1+x^2\right)\, dx\, dy$$

Now once we integrate, the value $\int_0^1 \left(1+x^2\right)\, dx$ is just a number, so we can pull it out of the outer-most integral:

$$\begin{aligned}\int_0^1 \frac{1}{1+y^2} \int_0^1 \left(1+x^2\right)\, dx\, dy &= \int_0^1 \left(1+x^2\right)\, dx \cdot \int_0^1 \frac{1}{1+y^2}\, dy \\ &= \left(x + \frac{x^3}{3}\right)\Bigg|_0^1 \cdot \tan^{-1}(y)\Bigg|_0^1 \\ &= \frac{4}{3} \cdot \frac{\pi}{4} \\ &= \pi/3\end{aligned}$$

4.3 Double integrals over general regions

In the previous two sections, all double integrals were performed over rectangles. In this section we remove this restriction and consider integration over more general regions of $\mathbb{R}^2$.

Motivating Example

A silicon wafer is a large, circular disc made of silicon which is used in the manufacture of computer processors and other electronic devices. Suppose that in the process of fabricating such a wafer some impurities are introduced (dust, water vapor, etc.) so that the wafer isn't pure silicon. If we were able to determine precisely where these impurities lie in the wafer, then we might be able to determine the density of the wafer at a particular point. To figure out the mass of the entire wafer we could then integrate this density. This presents us with a problem in that the wafer is circular (so the domain of our density function is a disc in the plane), whereas we only know how to integrate functions with a rectangular domain. So we need some way of extending our usual double integrals to deal with functions with other sorts of domains.

To associate some actual numbers with the scenario described above, suppose that our wafer has a radius of one meter, and for a point (x, y) in the wafer, the density of the wafer at that point is

$$\rho(x, y) = x^2 \cos(y) + 1.$$

The domain of our function ρ is

$$D = \left\{ (x, y) \,\middle|\, x^2 + y^2 \leq 1 \right\}.$$

Let's notice that if we pick an x-coordinate of a point inside this disc, the y-coordinates we can tack onto this x-coordinate lie between the values $-\sqrt{1 - y^2}$ and $\sqrt{1 - y^2}$. So for example, if we look at all of the (x, y) points inside our disc where the x-coordinate is $1/2$, the y-coordinates have to be between $-\sqrt{1 - 1/4}$ and $\sqrt{1 - 1/4}$.

Recall from the last lecture that to evaluate the integral

$$\iint\limits_D \rho(x, y)\, dA,$$

we integrate a "cross-section" function, $A(x)$. If we knew what the cross-section was, then we'd integrate

$$\iint_D \rho(x,y)\,dA = \int_{-1}^{1} A(x)\,dx,$$

since our x's run from -1 to 1. To calculate the cross-section function last time we just integrated our initial function $\rho(x,y)$ with respect to y over all of the possible y-values. Here our y-values depend on our chosen x, but once we've chosen an x we expect that our cross-section function should be

$$A(x) = \int_{-\sqrt{3}/2}^{\sqrt{3}/2} \rho(x,y)\,dy = \int_{-\sqrt{3}/2}^{\sqrt{3}/2} (x^2\cos(y)+1)\,dy.$$

Of course, there's nothing special about the choice of $x = 1/2$. In general, for any x between -1 and 1, the cross-section is

$$\begin{aligned} A(x) &= \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2\cos(y)+1)\,dy \\ &= (x^2\sin(y)+y)\Big|_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \\ &= x^2\sin(\sqrt{1-x^2}) + \sqrt{1-x^2} - x^2\sin(-\sqrt{1-x^2}) - \sqrt{1-x^2} \\ &= 2x^2\sin(\sqrt{1-x^2}) \end{aligned}$$

Above we used the fact that $\sin\theta$ is an odd function: $\sin(-\theta) = -\sin\theta$.

Notice that in order to find this cross-section, the bounds of our integral had to depend on where we were trying to find the cross-section. Aside from this one modification, our cross-section was found exactly like before. Notice here the bounds for our integral with respect to y were functions of x.

Now that we have the cross-section, we can calculate the integral we initially wanted:

$$\begin{aligned} \iint_D \rho(x,y)\,dA &= \int_{-1}^{1} A(x)\,dx \\ &= \int_{-1}^{1} 2x^2\sin(\sqrt{1-x^2})\,dx \end{aligned}$$

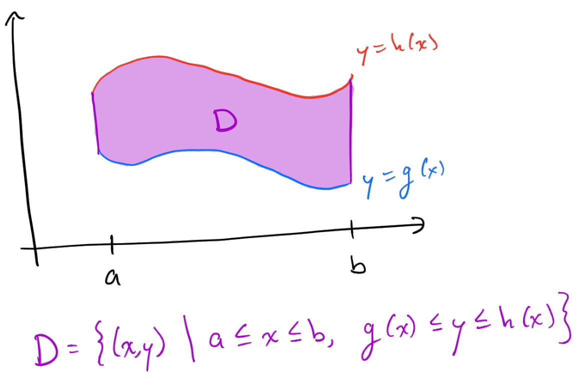
This is a hard integral to solve, so we won't bother to explicitly solve it right now, but just content ourselves with the fact that we can rewrite the integral over a non-rectangular region as an iterated integral:

$$\iint\limits_D (x^2\cos(y)+1)\,dA = \int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2\cos(y)+1)\,dy\,dx.$$

Integrals Over General Regions

We'll say a subset D of the plane $\mathbb{R}^2$ is *type I* if the x-values of points in D stay inside some fixed interval $[a,b]$, but the y-values of points in D are bounded by functions of x. That is, a type I region can be written as

$$D = \left\{(x,y)\,\middle|\, a \le x \le b,\ g(x) \le y \le h(x)\right\}.$$

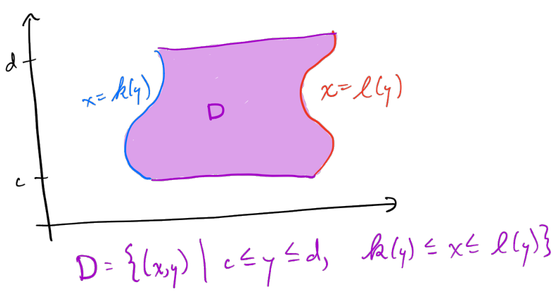


The integral of a continuous function $f(x,y)$ over a type I region D is given by

$$\iint\limits_D f(x,y)\,dA = \int_a^b\int_{g(x)}^{h(x)} f(x,y)\,dy\,dx.$$

We say that D is *type II* if the roles of x and y are switched from that of a type I region: that is, a type II region D can be written as

$$D = \left\{(x,y)\,\middle|\, c \le y \le d,\ k(y) \le x \le \ell(x)\right\}.$$



The integral of a continuous $f(x, y)$ over a type II region D is

$$\iint\limits_D f(x,y)\,dA = \int_c^d \int_{k(y)}^{\ell(y)} f(x,y)\,dx\,dy.$$

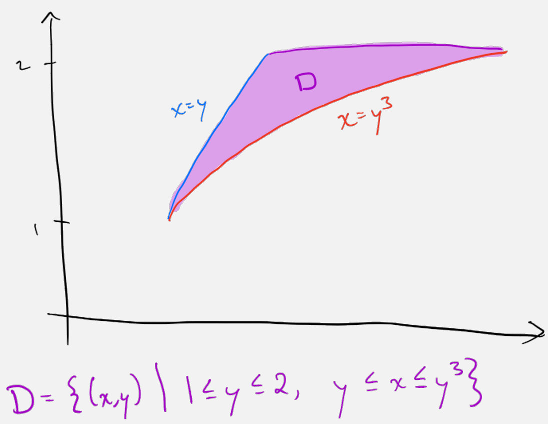
(Notice that some regions are both type I and type II simultaneously. For example, the disc considered above could be considered as type I or type II.)

Example 4.5.
Evaluate the integral

$$\iint\limits_D e^{x/y}\,dA$$

where D is the region

$$D = \left\{(x,y)\,\middle|\,1 \le y \le 2,\ y \le x \le y^3\right\}.$$



Notice that this is a type II region, so we our integral is given by

$$\begin{aligned}
\iint_D e^{x/y}\,dA &= \int_1^2 \int_y^{y^3} e^{x/y}\,dx\,dy \\
&= \int_1^2 \int_y^{y^3} e^{x\cdot 1/y}\,dx\,dy \\
&= \int_1^2 \left(\frac{e^{x/y}}{1/y}\right)\Bigg|_y^{y^3} dy \\
&= \int_1^2 ye^{x/y}\Big|_y^{y^3} dy \\
&= \int_1^2 \left(ye^{y^2} - ye\right) dy \\
&= \int_1^2 ye^{y^2}\,dy - \int_1^2 ye\,dy
\end{aligned}$$

For the integral on the left, perform the substitution $u = y^2$, $du = 2ydy$.

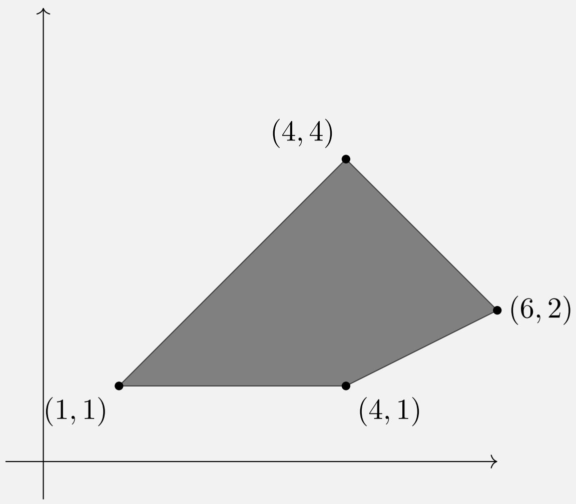
$$\begin{aligned}
\int_1^2 ye^{y^2}\,dy - \int_1^2 ye\,dy =& \frac{1}{2}\int_1^4 e^u\,du - \int_1^2 ye\,dy \\
=& \frac{1}{2}e^u\Big|_1^4 - \frac{ey^2}{2}\Big|_1^2 \\
=& \frac{1}{2}\left(e^4 - e\right) - \left(\frac{4e}{2} - \frac{e}{2}\right) \\
=& \frac{1}{2}(e^4 - e - 3e) \\
=& \frac{e^4 - 4e}{2}
\end{aligned}$$

In general a region may not be expressible as a single type I or type II domain. In such a situation we can cut the region up into several pieces which are type I or type II. We can do this because of the following theorem.

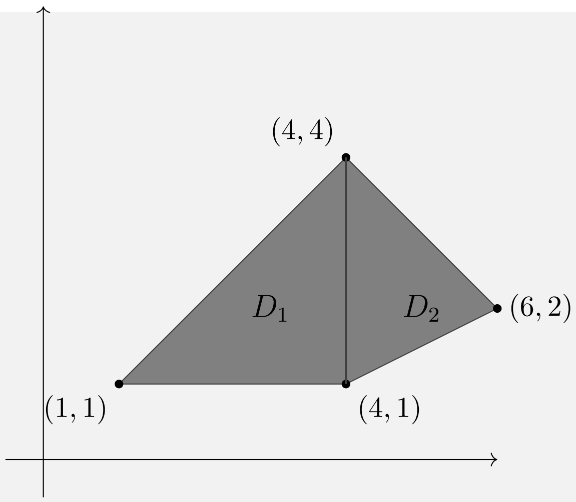
Theorem 4.2.
Suppose that D_1 and D_2 are two regions in $\mathbb{R}^2$ which don't overlap. Then, writing $D = D_1 \cup D_2$,

$$\iint\limits_{D} f(x,y)\,dA = \iint\limits_{D_1} f(x,y)\,dA + \iint\limits_{D_2} f(x,y)\,dA.$$

Example 4.6.
Integrate the function $x + y^2$ over the region indicated in the figure below.



We'll split this up into two regions, each of which is type I.



Our regions are

$$D_1 = \left\{(x, y) \,\middle|\, 1 \le x \le 4,\ 1 \le y \le x\right\}$$
$$D_2 = \left\{(x, y) \,\middle|\, 4 \le x \le 6,\ \frac{x}{2} - 1 \le y \le 8 - x\right\}.$$

Integrating over D_1 we have

$$\begin{aligned}
\iint_{D_1} (x + y)\, dA &= \int_1^4 \int_1^x (x + y)\, dy\, dx \\
&= \int_1^4 \left(xy + \frac{y^2}{2}\right)\Big|_1^x dx \\
&= \int_1^4 \left(x^2 + \frac{x^2}{2} - x - \frac{1}{2}\right) dx \\
&= \int_1^4 \left(\frac{3x^2}{2} - x - \frac{1}{2}\right) dx \\
&= \left(\frac{x^3}{2} - \frac{x^2}{2} - \frac{x}{2}\right)\Big|_1^4 \\
&= 32 - 8 - 2 - \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \\
&= 22.5
\end{aligned}$$

Integrating over D_2,

$$\iint_{D_2} (x+y)\,dA = \int_4^6 \int_{x/2-1}^{8-x} (x+y)\,dy\,dx$$
$$= \int_4^6 \left(xy + \frac{y^2}{2} \right)\Bigg|_{x/2-1}^{8-x} dx$$
$$= \int_4^6 \left[\left(8x - x^2 + \frac{64-16x+x^2}{2} \right) - \left(\frac{x^2}{2} - x + \frac{x^2-4x+4}{8} \right) \right] dx$$

Before integrating, let's simplify the integrand a little bit.

$$\int_4^6 \left[\left(8x - x^2 + \frac{64-16x+x^2}{2} \right) - \left(\frac{x^2}{2} - x + \frac{x^2-4x+4}{8} \right) \right] dx$$
$$= \frac{1}{8} \int_4^6 \left[\left(64x - 8x^2 + 256 - 64x + 4x^2 \right) - \left(4x^2 - 8x + x^2 - 4x + 4 \right) \right] dx$$
$$= \frac{1}{8} \int_4^6 \left[\left(-4x^2 + 256 \right) - \left(5x^2 - 12x + 4 \right) \right] dx$$
$$= \frac{1}{8} \int_4^6 \left(-9x^2 + 12x + 252 \right) dx$$

This simplified integral is much easier to integrate, but let's first notice that each term in the integrand is a multiple of 3, so we can pull the 3 out:

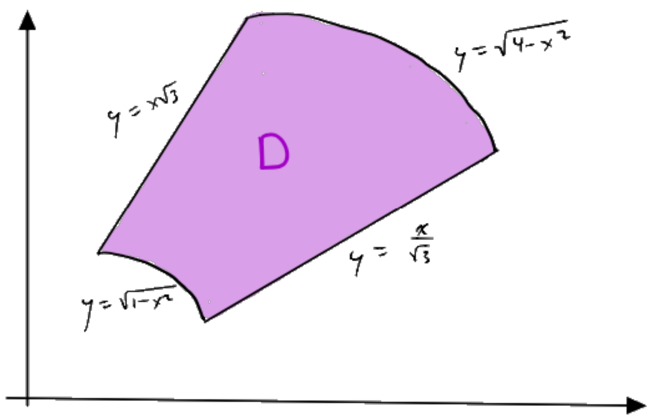
$$\frac{1}{8} \int_4^6 \left(-9x^2 + 12x + 252 \right) dx$$
$$= \frac{3}{8} \int_4^6 \left(-3x^2 + 4x + 84 \right) dx$$
$$= \frac{3}{8} \left(-x^3 + 2x^2 + 84x \right)\Big|_4^6$$
$$= \frac{3}{8} [(-216 + 72 + 504) - (-64 + 32 + 336)]$$
$$= \frac{3}{8}(360 - 304)$$
$$= 21$$

Thus

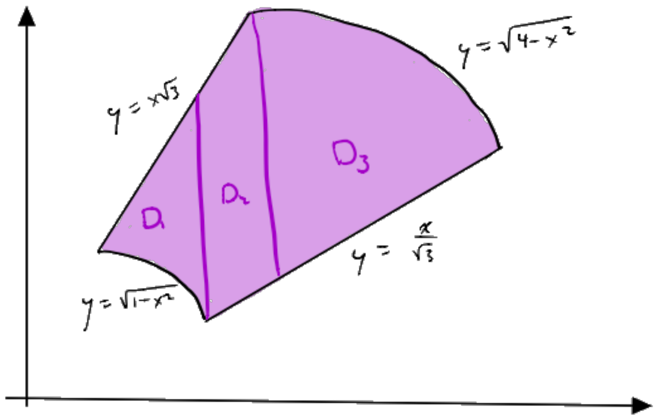
$$\begin{aligned}\iint\limits_D (x+y)\,dA &= \iint\limits_{D_1} (x+y)\,dA + \iint\limits_{D_2} (x+y)\,dA \\ &= \int_1^4 \int_1^x (x+y)\,dy\,dx + \int_4^6 \int_{x/2-1}^{8-x} (x+y)\,dy\,dx \\ &= 22.5 + 21 \\ &= 43.5\end{aligned}$$

4.4 Double integrals in polar coordinates

Suppose that we wanted to integrate the function $f(x, y) = x^2 + y^2$ over the region in the first quadrant bounded by the curves $y = x\sqrt{3}$, $y = \frac{x}{\sqrt{3}}$, $y = \sqrt{1 - x^2}$ and $y = \sqrt{4 - x^2}$.



We can't very easily represent this region as the (x, y)-points where the y-values are sandwiched between two functions of x, or as the region where the x-values are sandwiched between two functions of y. Thus to integrate over this region, we need to split R up into smaller regions where we *can* sandwich y between two functions of x, or x between two functions of y. One way to do this is the following:

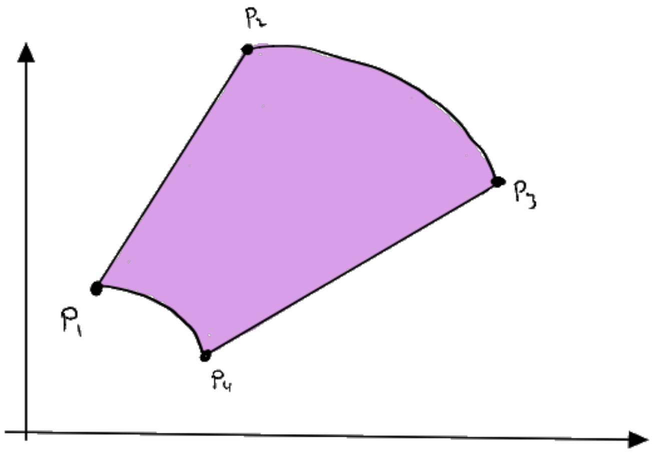


Each of these smaller regions can be represented as type I regions as indicated below:

$$D_1 = \left\{(x,y)\,\middle|\,? \leq x \leq ?,\ \sqrt{1-x^2} \leq y \leq x\sqrt{3}\right\},$$
$$D_2 = \left\{(x,y)\,\middle|\,? \leq x \leq ?,\ \frac{x}{\sqrt{3}} \leq y \leq x\sqrt{3}\right\}, \text{ and}$$
$$D_3 = \left\{(x,y)\,\middle|\,? \leq x \leq ?,\ \frac{x}{\sqrt{3}} \leq y \leq \sqrt{4-x^2}\right\}.$$

Notice we don't yet have any bounds for the x-values in our region. To get these x-values, which we'll need in order to set up our integrals, we have to find the intersection points of our curves.

Supposing we label these intersection points (which are the "corners" of our region D) as follows,



We can find the coordinates of P_1 by setting $x\sqrt{3} = \sqrt{1-x^2}$ (as P_1 occurs at the intersection of these two curves).

$$\begin{aligned} & x\sqrt{3} = \sqrt{1-x^2} \\ \implies & 3x^2 = 1 - x^2 \\ \implies & 4x^2 = 1 \\ \implies & x^2 = \frac{1}{4} \\ \implies & x = \frac{1}{2}. \end{aligned}$$

Plugging $x = \frac{1}{2}$ into either of $x\sqrt{3}$ or $\sqrt{1-x^2}$, we see the y-value of P_1 is $\sqrt{3}/2$; that is, $P_1 = (1/2, \sqrt{3}/2)$.

Similarly, solving $x\sqrt{3} = \sqrt{4-x^2}$, we see that $P_2 = (1, \sqrt{3})$; solving $x/\sqrt{3} = \sqrt{4-x^2}$ gives $P_3 = (\sqrt{3}, 1)$; and, finally, solving $x/\sqrt{3} = \sqrt{1-x^2}$ tells us $P_4 = (\sqrt{3}/2, 1/2)$.

Now as our regions D_1, D_2, and D_3 are disjoint, we can compute

$$\iint_D f(x,y)\,dA = \iint_{D_1} f(x,y)\,dA + \iint_{D_2} f(x,y)\,dA + \iint_{D_3} f(x,y)\,dA.$$

Writing these as iterated integrals we may compute

$$\begin{aligned} \iint_{D_1} f(x,y)\,dA &= \int_{1/2}^{\sqrt{3}/2} \int_{\sqrt{1-x^2}}^{x\sqrt{3}} (x^2+y^2)\,dy\,dx \\ \iint_{D_2} f(x,y)\,dA &= \int_{\sqrt{3}/2}^{1} \int_{x/\sqrt{3}}^{x\sqrt{3}} (x^2+y^2)\,dy\,dx \\ \iint_{D_3} f(x,y)\,dA &= \int_{1}^{\sqrt{3}} \int_{x/\sqrt{3}}^{\sqrt{4-x^2}} (x^2+y^2)\,dy\,dx \end{aligned}$$

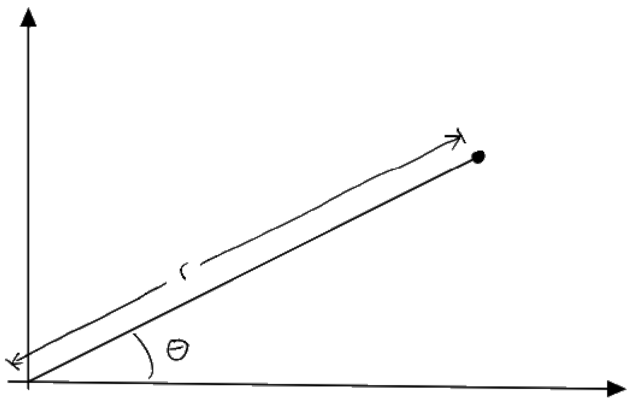
Let's now compute the first of these integrals:

$$\begin{aligned} \iint_{D_1} f(x,y)\,dA &= \int_{1/2}^{\sqrt{3}/2} \int_{\sqrt{1-x^2}}^{x\sqrt{3}} (x^2+y^2)\,dy\,dx \\ &= \int_{1/2}^{\sqrt{3}/2} \left(x^2 y + \frac{y^3}{3} \right)\Bigg|_{\sqrt{1-x^2}}^{x\sqrt{3}} dx \\ &= \int_{1/2}^{\sqrt{3}/2} \left(x^3\sqrt{3} + \frac{x^3\sqrt{27}}{3} - x^2\sqrt{1-x^2} + \frac{(1-x^2)^{3/2}}{3} \right) dx. \end{aligned}$$

Now things start to get hairier: though the first two terms are easy to integrate, the third and fourth terms require a trig substitution, namely $x = \sin(\theta)$, $dx = \cos(\theta)\,d\theta$. After performing this substitution, you still have to do integration by parts and u-substitution to evaluate the integral; and this is only one-third of our original integral! All of this is do-able, of course, but very tedious and time-consuming.

In order to turn this integral into something a bit more tractable, we can try to change our coordinate system from the usual Cartesian coordinates to polar coordinates.

Recall that in polar coordinates we specify points as pairs (r, θ) where the r-value tells us how far from the origin the point is, and θ tells us the angle abofe the x-axis of a line through the origin and our point.



To convert a point (x, y) in Cartesian coordinates to polar, we use the fomulas

$$r = \sqrt{x^2 + y^2}$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right),$$

and to convert a point (r, θ) in polar coordinates into Cartesian coordinates, we use

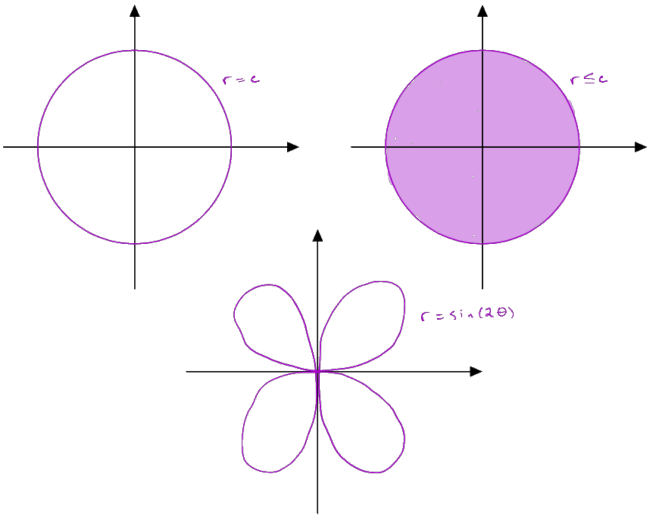
$$x = r\cos(\theta)$$
$$y = r\sin(\theta).$$

For example, the table below describes four points in both Cartesian and polar coordinates.

Cartesian	**Polar**
$(1/2, \sqrt{3}/2)$	$(1, \pi/3)$
$(-2, 5)$	$\left(\sqrt{29}, \tan^{-1}\left(-5/2\right)\right)$
$(3\sqrt{2}/2, 3\sqrt{2}/2)$	$(3, \pi/4)$
$(0, -2)$	$(-2, \pi/2)$

For the last entry in the table above, recall that we adopt the convention that if r is negative, we move away from the origin in the opposite direction. That is $(-1, 5\pi/4)$ and $(1, \pi/4)$ represent the same point in polar coordinates.

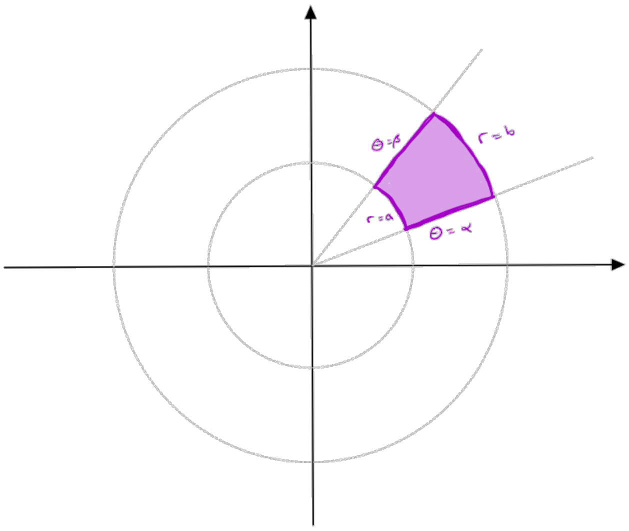
The advantange of using polar coordinates is that some regions become much simpler to represent in polar than they are in Cartesian.



One region we will especially care about is called a ***polar rectangle***. This is the set of all points (r, θ) whose polar coordinates statisfy $a \leq r \leq b$ and $\alpha \leq \theta \leq \beta$.

$$R = \left\{ (r, \theta) \,\middle|\, a \leq r \leq b, \alpha \leq \theta \leq \beta \right\}.$$

Such a region looks something like the following:



We adopt the convention that the polar rectangle above can be denoted $[a, b] \times [\alpha, \beta]$. **Warning**: Notice there is a potential for ambiguity here, as $[a, b] \times [\alpha, \beta]$ in Cartesian and polar represent very different regions!

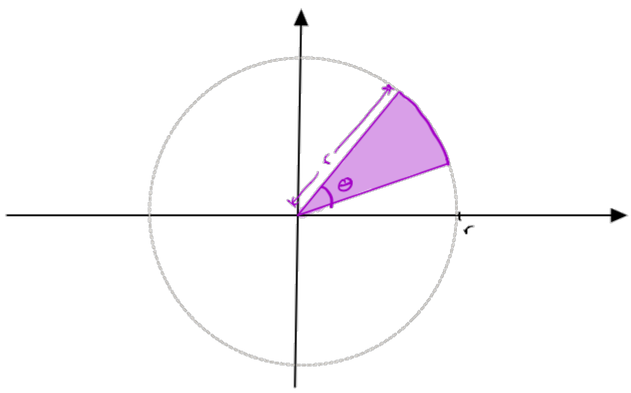
These polar rectangles will be the building blocks for our integrals in polar coordinates.

Recall that to integrate a function over a Cartesian rectangle, we split the rectangle up into smaller subrectangles, and constructed rectangular prisms over these smaller subrectangles, then added up the volumes of all these prisms to estimate the value of our integral.

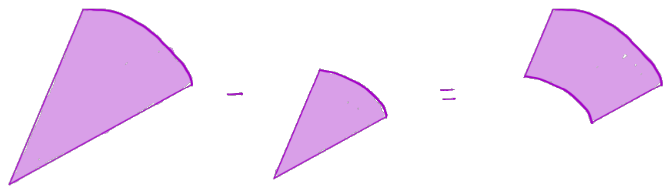
We're going to do something very similar now with polar rectangles, but this will require that we know how to compute the area of a polar rectangle $[a, b] \times [\alpha, \beta]$. To determine the area of this polar rectangle, let's first notice that for a circle of radius r the area of is given by πr^2. If we have a sector of that circle which has angle θ (i.e., a pizza slice shaped region where the tip of the slice makes angle θ), then the area of that slice is

$$\pi r^2 \cdot \frac{\theta}{2\pi} = \frac{\theta r^2}{2}.$$

(The total angle around the circle is 2π and our sector has angle θ, so our sector has proportion $\frac{\theta}{2\pi}$ of the circle. Thus the proportion of the total area πr^2 of the circle we have in the sector is also $\frac{\theta}{2\pi}$.)



For the polar rectangle $[a, b] \times [\alpha, \beta]$, the area we have is given by the area of the "outer" sector minus the area of the inner sector.



Notice the angle of sector is $\beta - \alpha$, so the area of the larger (outer) sector is $\frac{(\beta-\alpha)b^2}{2}$. Similarly, the area of the inner sector is $\frac{(\beta-\alpha)a^2}{2}$. This means the area of our polar rectangle is

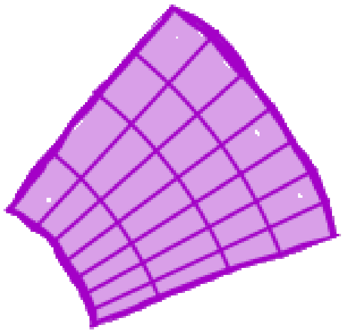
$$\frac{(\beta-\alpha)b^2}{2} - \frac{(\beta-\alpha)a^2}{2}.$$

With a little bit of arithmetic we can rewrite this as follows:

$$\begin{aligned}\frac{(\beta-\alpha)b^2}{2} - \frac{(\beta-\alpha)a^2}{2} &= \frac{\beta-\alpha}{2}\left(b^2 - a^2\right)\\ &= \frac{1}{2}(\beta-\alpha)(b-a)(b+a)\\ &= \frac{b+a}{2}(\beta-\alpha)(b-a).\end{aligned}$$

Now observe that $\beta - \alpha$ is the change in the angle of the sector, so we may refer to this as $\Delta\theta$. Similarly, $b - a$ is the change in the radius of our sectors, so we might call this value Δr. Finally, $\frac{b+a}{2}$ is the point halfway between the radii $r = a$ and $r = b$, so let's refer to this quantity as r^*. With these conventions, the area of our polar rectangle may be written as $r^*\Delta r\Delta\theta$.

Now, to estimate our integral, we first split the polar rectangle into sub-polar-rectangles,



We'll pick a point inside each of these subrectangles, let's call it (r_i^*, θ_j^*), where $r_{i-1} \leq r_i^* \leq r_i$ and $\theta_{j-1} \leq \theta_j^* \leq \theta_j$. When we take limits the choice of r_i^* and θ_j^* we make won't matter, as long as they are in the intervals described above. Since the choices don't matter, we may choose to make some convenient choices, such as picking r_i^* and θ_j^* to be the midpoints of their respective intervals. That is, we may choose

$$r_i^* = \frac{r_i + r_{i-1}}{2}$$
$$\theta_j^* = \frac{\theta_j + \theta_{j-1}}{2}.$$

We wish to plug these values into our function to determine the height of our "polar rectangular prism." If our function was given to us as an expression in polar coordinates, we can directly plug these values in. If, however, our function was given as an expression $f(x, y)$ in Cartesian coordinates, then we need to convert our point (r_i^*, θ_j^*) in polar coordinates into Cartesian coordinates before plugging into the function. Using $x = r\cos(\theta)$ and $y = r\sin(\theta)$, this means the height of our "polar rectangular prism" is $f(r_i^*\cos(\theta_j^*), r_i^*\sin(\theta_j^*))$. Since the area of the base is $r_i^*\Delta r_i\Delta\theta_j$, the volume of our prism is

$$f\left(r_i^*\cos(\theta_j^*), r_i^*\sin(\theta_j^*)\right) r_i^*\Delta r_i\Delta\theta_j.$$

Summing up these volumes, our estimate for the integral of $f(x, y)$ over the polar rectangle $R = [a, b] \times [\alpha, \beta]$ is

$$\iint_R f(x, y)\, dA \approx \sum_{j=1}^{m}\sum_{i=1}^{n} f\left(r_i^*\cos(\theta_j^*), r_i^*\sin(\theta_j^*)\right) r_i^*\Delta r_i\Delta\theta_j.$$

Of course, we'll take the limit of these summations, and the above value becomes

$$\iint\limits_R f(x,y)\,dA = \int_\alpha^\beta \int_a^b f(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta.$$

So, integrating in polar coordinates is almost what you'd expect it to be, in that you replace x with $r\cos(\theta)$ and y with $r\sin(\theta)$, but we pick up an extra factor of r coming from the area of the polar rectangle.

Our initial example, which resulted a complicated mess involving three different integrals, can be greatly simplified if we convert everything to polar.

First we'll need to rewrite the region we're integrating over as a polar rectangle. To do this we again use the equations $y = r\sin(\theta)$ and $x = r\cos(\theta)$. For the inner-most circle given by $y = \sqrt{1-x^2}$, plugging in the polar-to-Cartesian conversion formulas results in the following:

$$\begin{aligned}
&y = \sqrt{1-x^2} \\
\implies &r\sin(\theta) = \sqrt{1-(r\cos(\theta))^2} \\
\implies &r^2\sin^2(\theta) = 1 - r^2\cos^2(\theta) \\
\implies &r^2\sin^2(\theta) + r^2\cos^2(\theta) = 1 \\
\implies &r^2(\sin^2(\theta) + \cos^2(\theta)) = 1 \\
\implies &r^2 = 1 \\
\implies &r = 1.
\end{aligned}$$

(Here it's worth noting that the solution to $r^2 = 1$ is $r = \pm 1$, however, we're talking about the radius of a circle here, so we will ignore the negative value.)

Similarly, we convert the line $y = x/\sqrt{3}$ to polar as follows:

$$\begin{aligned}
&y = \frac{x}{\sqrt{3}} \\
\implies &r\sin(\theta) = \frac{r\cos(\theta)}{\sqrt{3}} \\
\implies &\frac{r\sin(\theta)}{r\cos(\theta)} = \frac{1}{\sqrt{3}} \\
\implies &\tan(\theta) = \frac{1}{\sqrt{3}} \\
\implies &\theta = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \tan^{-1}\left(\frac{1/2}{\sqrt{3}/2}\right).
\end{aligned}$$

As tangent is sine over cosine, we want the angle θ which when plugged into sine gives us $1/2$ and produces $\sqrt{3}/2$ from cosine. This angle is $\theta = \pi/6$.

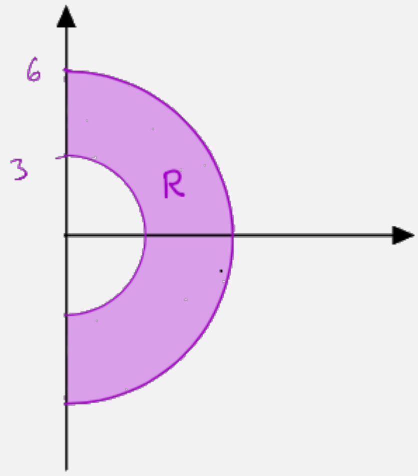
So, our line $y = x/\sqrt{3}$ in polar is simply $\theta = \pi/6$. Doing the same sort of thing for $y = \sqrt{4 - x^2}$ and $y = x\sqrt{3}$ would give us $r = 2$ and $\theta = \pi/3$. Thus the region we are integrating over is the polar rectangle $[1, 2] \times [\pi/6, \pi/3]$.

Integrating $x^2 + y^2$ over this region now greatly simplifies in polar coordinates.

$$\begin{aligned}
\iint\limits_R \left(x^2 + y^2\right)\, dA &= \int_{\pi/6}^{\pi/3} \int_1^2 \left(r^2 \cos^2(\theta) + r^2 \sin^2(\theta)\right)\, r\, dr\, d\theta \\
&= \int_{\pi/6}^{\pi/3} \int_1^2 r^2 \left(\cos^2(\theta) + \sin^2(\theta)\right)\, r\, dr\, d\theta \\
&= \int_{\pi/6}^{\pi/3} \int_1^2 r^2\, r\, dr\, d\theta \\
&= \int_{\pi/6}^{\pi/3} r^3\, dr\, d\theta \\
&= \int_{\pi/6}^{\pi/3} \left.\frac{r^4}{4}\right|_1^2 d\theta \\
&= \int_{\pi/6}^{\pi/3} \left(\frac{16}{4} - \frac{1}{4}\right) d\theta \\
&= \int_{\pi/6}^{\pi/3} \frac{15}{4}\, d\theta \\
&= \frac{15}{4} \cdot \frac{\pi}{6} \\
&= \frac{5\pi}{8}
\end{aligned}$$

Let's finish up our discussion of integration in polar coordinates with a few more examples.

Example 4.7.
Integrate the function $f(x, y) = xy$ over the following region:



First notice that this region is between two half-circles and so is a polar rectangle. The radius starts at 3 and increases to 6, while the angle ranges from $-\pi/2$ to $\pi/2$. Our integral is thus

$$\begin{aligned}\iint\limits_R f(x,y)\,dA &= \int_{-\pi/2}^{\pi/2}\int_3^6 f(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta \\ &= \int_{-\pi/2}^{\pi/2}\int_3^6 r\cos(\theta)\cdot r\sin(\theta)\,r\,dr\,d\theta \\ &= \int_{-\pi/2}^{\pi/2}\int_3^6 r^3\cos(\theta)\sin(\theta)\,dr\,d\theta \\ &= \int_{-\pi/2}^{\pi/2}\cos(\theta)\sin(\theta)\frac{r^4}{4}\bigg|_3^6\,d\theta \\ &= \frac{1215}{4}\int_{-\pi/2}^{\pi/2}\cos(\theta)\sin(\theta)\,d\theta\end{aligned}$$

To finish the integral we will use the substitution $u = \sin(\theta)$, $du = \cos(\theta)\,d\theta$. Notice that our limits will change from $-\pi/2$ and $\pi/2$ to -1

and 1.

$$\frac{1215}{4}\int_{-\pi/2}^{\pi/2}\sin(\theta)\cos(\theta)\,d\theta = \frac{1215}{4}\int_{-1}^{1} u\,du$$
$$= \frac{1215}{4}\frac{u^2}{2}\Big|_{-1}^{1}$$
$$= \frac{1215}{4}\left(\frac{1}{2}-\frac{1}{2}\right)$$
$$= 0$$

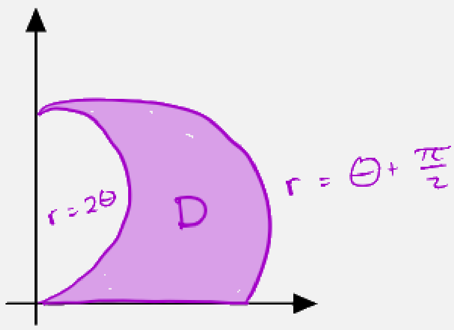
Just as we can integrate functions $f(x, y)$ in Cartesian over regions other than (Cartesian) rectangles, we can likewise perform integration in polar coordinates over more general regions than simply polar rectangles. If we had a region D which in polar coordinates was expressed as

$$D = \{(r,\theta) \,\big|\, \alpha \leq \theta \leq \beta,\ g(\theta) \leq r \leq h(\theta)\}$$

, then we would integrate a function $f(x, y)$ over this region as

$$\iint\limits_D f(x,y)\,dA = \int_{\alpha}^{\beta}\int_{g(\theta)}^{h(\theta)} f(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta.$$

Example 4.8.
Compute the area of the region below.



We can compute the area of a region by integrating the constant function 1 over that region:

$$\text{Area}(D) = \iint_D 1\, dA.$$

For the region we have here, it will be easiest to do this by expressing the region in polar coordinates. Notice in particular that the angles in the region range from $\theta = 0$ to $\theta = {}^{\pi}/_{2}$ since everything in this region is in the first quadrant. We are also given that the curves on the left- and right-hand edges of region are $r = 2\theta$ and $r = \theta + {}^{\pi}/_{2}$, respectively. We can thus compute the area of this region as follows:

$$\begin{aligned}
\text{Area}(D) &= \iint_D 1\, dA \\
&= \int_0^{\pi/2} \int_{2\theta}^{\theta+\pi/2} r\, dr\, d\theta \\
&= \int_0^{\pi/2} \frac{r^2}{2}\bigg|_{2\theta}^{\theta+\pi/2} d\theta \\
&= \int_0^{\pi/2} \left(\frac{(\theta + \pi/2)}{2} - \frac{4\theta^2}{2} \right) d\theta \\
&= \frac{1}{2} \int_0^{\pi/2} \left(\theta^2 + \theta\pi + \frac{\pi^2}{4} - 4\theta^2 \right) d\theta \\
&= \frac{1}{2} \int_0^{\pi/2} \left(-3\theta^2 + \theta\pi + \frac{\pi^2}{4} \right) d\theta \\
&= \frac{1}{2} \left(-\theta^3 + \frac{\theta^2\pi}{2} + \frac{\theta\pi^2}{4} \right)\bigg|_0^{\pi/2} \\
&= \frac{1}{2} \left(\frac{-\pi^3}{8} + \frac{\pi^3}{8} + \frac{\pi^3}{8} \right) \\
&= \frac{\pi^3}{16}
\end{aligned}$$

4.5 Triple integrals

Our original motivation for studying integrals of functions of one variable was to find the area underneath a curve; with of two variables, the motivation was to find the volume under a surface. "Integration" is much more general than just calculating these geometric quantities, and integration should be viewed more as a way of computing a particular type of infinite sum rather than just a "trick" for computing areas and volumes.

We're now going to start studying integration in three variables, where our intuitive notions of "area under a curve" or "volume under a surface" don't apply (or, at least, they don't apply as directly).

So, suppose that B is a rectangular prism in three-dimensional space:

$$B = \{(x, y, z) \mid a \leq x \leq b,\ c \leq y \leq d,\ r \leq z \leq s\}$$

We may sometimes denote this region as $[a, b] \times [c, d] \times [r, s]$.

We will partition $[a, b]$ into ℓ subintervals,

$$a = x_0 < x_1 < \cdots < x_{\ell-1} < x_\ell = b,$$

and similarly partition $[c, d]$ into subintervals

$$c = y_0 < y_1 < \cdots < y_{n-1} < y_n = d,$$

and partition $[r, s]$ into n subintervals

$$r = z_0 < z_1 < \cdots < z_{n-1} < z_n = s.$$

These partitions then split B into ℓmn sub-prisms,

$$B_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k].$$

We pick some point inside each of these sub-prisms, call it $(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})$ and consider the triple Riemann sum

$$\sum_{i=1}^{\ell} \sum_{j=1}^{m} \sum_{k=1}^{n} f(x^*_{ijk}, y^*_{ijk}, z^*_{ijk}) \Delta V_{ijk}$$

where $\Delta V_{ijk} = \Delta x_i \, \Delta y_j \, \Delta z_k$ is the volume of the prism B_{ijk}.

Taking the limit of these prisms as their volumes get arbitrarily small gives us the ***triple integral of*** $f(x, y, z)$ ***over*** B. Letting $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{R}$ be the partitions of the x, y, and z, intervals, we have

$$\iiint\limits_B f(x, y, z)\, dV = \lim_{|\mathcal{P}| \to 0} \lim_{|\mathcal{Q}| \to 0} \lim_{|\mathcal{R}| \to 0} \sum_{i=1}^{\ell_{\mathcal{P}}} \sum_{j=1}^{m_{\mathcal{Q}}} \sum_{k=1}^{n_{\mathcal{R}}} f(x^*_{ijk}, y^*_{ijk}, z^*_{ijk}) \Delta V_{ijk}$$

This limit is guaranteed to exist if f is a continuous function.

Though defining triple integrals as limits of triple Riemann sums makes sense, it's hard to actually compute triple integrals this way. Luckily, Fubini's theorem extends to allow us to write the triple integral as a triple iterated integral:

$$\begin{aligned}
\iiint\limits_B f(x,y,z)\,dV &= \int_r^s \int_c^d \int_a^b f(x,y,z)\,dx\,dy\,dz \\
&= \int_a^b \int_r^s \int_c^d f(x,y,z)\,dy\,dz\,dx \\
&= \int_a^b \int_c^d \int_r^s f(x,y,z)\,dz\,dy\,dz \\
&\vdots
\end{aligned}$$

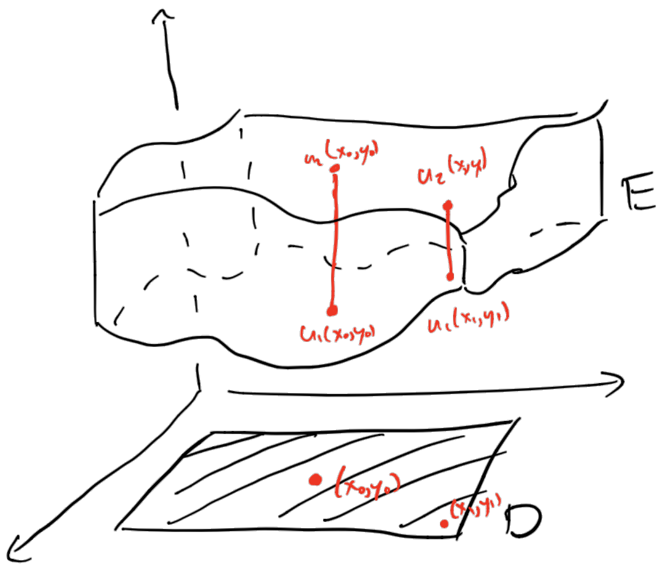
Example 4.9.
Compute the integral of $f(x,y,z) = xy + z^2$ over $B = [0,2] \times [0,1] \times [0,3]$.

$$\begin{aligned}
\iiint\limits_B (xy + z^2)\,dV &= \int_0^3 \int_0^1 \int_0^2 (xy + z^2)\,dx\,dy\,dz \\
&= \int_0^3 \int_0^1 \left(\frac{x^2 y}{2} + xz^2 \right)\Bigg|_0^2 dy\,dz \\
&= \int_0^3 \int_0^1 (2y + 2z^2)\,dy\,dz \\
&= \int_0^3 (y^2 + 2yz^2)\Big|_0^1 dz \\
&= \int_0^3 (1 + 2z^2)\,dz \\
&= \left(z + \frac{2z^3}{3} \right)\Bigg|_0^3 \\
&= 3 + \frac{54}{3} = \frac{9 + 54}{3} = \frac{63}{3} \\
&= 21
\end{aligned}$$

Notice that integrating 1 over B would simply give us the volume of the prism:

$$\begin{aligned}
\iiint\limits_{[a,b]\times[c,d]\times[r,s]} 1\, dV &= \int_r^s \int_c^d \int_a^b 1\, dx\, dy\, dz \\
&= \int_r^s \int_c^d x\Big|_a^b\, dy\, dz \\
&= \int_r^s \int_c^d (b-a)\, dy\, dz \\
&= \int_r^s (b-a)y\Big|_c^d\, dz \\
&= \int_r^s (b-a)(d-c)\, dz \\
&= (b-a)(d-c)z\Big|_r^s \\
&= (b-a)(d-c)(r-s)
\end{aligned}$$

Of course, we may want to integrate over more general regions than simply rectangular prisms. For any three-dimensional solid E we want to integrate over, we could try to project that solid to the xy-plane to obtain some planar region D.



If we can represent E as the set of (x, y, z)-points where for each given point $(x, y) \in D$ the possible z-values for $(x, y, z) \in E$ lying over $(x, y) \in D$ are given by functions of X and y, such as

$$E = \{(x, y, z) \,|\, (x, y) \in D,\ u_1(x, y) \leq z \leq u_2(x, y)\}\,,$$

then we will say E is a ***type I solid*** and write

$$\iiint\limits_E f(x, y, z)\, dV = \iint\limits_D \left(\int_{u_1(x,y)}^{u_2(x,y)} f(x, y, z)\, dz \right) dA$$

Notice that after evaluating the inner integral,

$$\int_{u_1(x,y)}^{u_2(x,y)} f(x, y, z)\, dz$$

the z's will integrate out (they will be replaced by $u_1(x, y)$ and $u_2(x, y)$ after finding the antiderivative of $f(x, y, z)$ with respect to z), and the remaining integration is an integral in two variables.

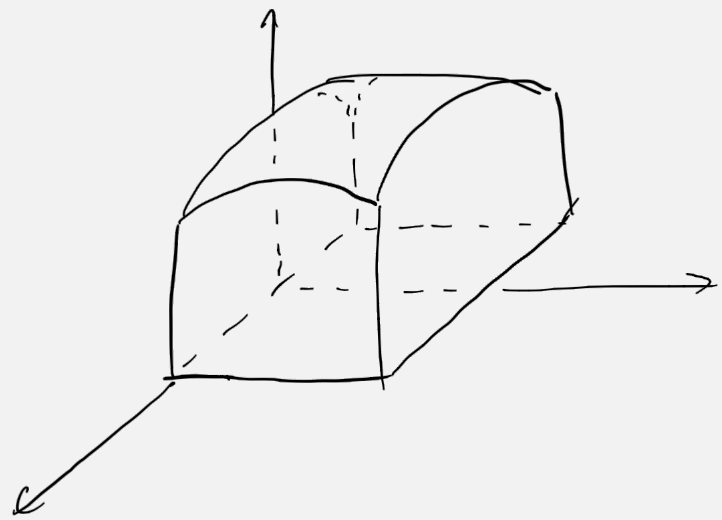
Example 4.10.
Integrate the function $f(x, y, z) = z$ over the region

$$E = \left\{(x, y, z) \,|\, -\sqrt{2} \leq x \leq \sqrt{2},\ 0 \leq y \leq \sqrt{2},\ 0 \leq z \leq \sqrt{4 - x^2 - y^2}\right\}$$

First let's figure out what our domain of integration is. Our z-values are bounded below by 0 and above by $\sqrt{4 - x^2 - y^2}$, where our x-values are between $-\sqrt{2}$ and $\sqrt{2}$, and the y-values are between 0 and $\sqrt{2}$. Notice that $z = \sqrt{4 - x^2 - y^2}$ is the top-half of a sphere of radius 2:

$$\begin{aligned} z &= \sqrt{4 - x^2 - y^2} \\ \implies z^2 &= 4 - x^2 - y^2 \\ \implies x^2 + y^2 + z^2 &= 4 \end{aligned}$$

The last line is the equation of radius 2 centered at the origin. So, the solid we are integrating over looks like



To integrate z over this region we have

$$\iiint\limits_E z\,dV = \iint\limits_D \left(\int_0^{\sqrt{4-x^2-y^2}} z\,dz\right) dA$$
$$= \iint\limits_D \left(\left.\frac{z^2}{2}\right|_0^{\sqrt{4-x^2-y^2}}\right) dA$$
$$= \frac{1}{2}\iint\limits_D \left(4 - x^2 - y^2\right)\,dA.$$

We now write this double integral as an iterated integral:

$$\begin{aligned}
&= \frac{1}{2}\int_0^{\sqrt{2}}\int_{-\sqrt{2}}^{\sqrt{2}} \left(4 - x^2 - y^2\right)\,dx\,dy \\
&= \frac{1}{2}\int_0^{\sqrt{2}} \left(4x - \frac{x^3}{3} - xy^2\right)\Bigg|_{-\sqrt{2}}^{\sqrt{2}} dy \\
&= \frac{1}{2}\int_0^{\sqrt{2}} \left(4\sqrt{2} - \frac{2^{3/2}}{3} - \sqrt{2}y^2 - \left(-4\sqrt{2} + \frac{2^{3/2}}{3} + \sqrt{2}y^2\right)\right) dy \\
&= \frac{1}{2}\int_0^{\sqrt{2}} \left(8\sqrt{2} - \frac{2^{5/2}}{3} - 2\sqrt{2}y^2\right) dy \\
&= \frac{1}{2}\left(8\sqrt{2}y - \frac{2^{5/2}y}{3} - \frac{2\sqrt{2}y^3}{3}\right)\Bigg|_0^{\sqrt{2}} \\
&= \frac{1}{2}\left(16 - \frac{2^{6/2}}{3} - \frac{2\cdot 2^{1/2}\cdot 2^{3/2}}{3}\right) \\
&= \frac{1}{2}\left(16 - \frac{8}{3} - \frac{8}{3}\right) \\
&= 8 - \frac{4}{3} - \frac{4}{3} \\
&= \frac{16}{3}
\end{aligned}$$

The planar region D we have might be more complicated than simply a rectangle. For example, in our problem above, if the region we were integrating over was instead

$$E = \left\{(x, y, z)\,\middle|\, -\sqrt{2} \le x \le \sqrt{2},\ -x \le y \le x^2,\ 0 \le z \le \sqrt{4 - x^2 - y^2}\right\}$$

then the region D that appeared in our double integral would instead be

$$D = \left\{(x, y)\,\middle|\, -\sqrt{2} \le x \le \sqrt{2}, -x \le y \le x^2\right\}.$$

This would change our final answer in the problem because the step that involved the integral

$$\frac{1}{2}\iint\limits_D (4 - x^2 - y^2)\,dA$$

above would be the (slightly) more complicated iterated integral

$$\frac{1}{2}\int_{-\sqrt{2}}^{\sqrt{2}}\int_{-x}^{x^2}(4-x^2-y^2)\,dy\,dx.$$

Normally when computing a triple integral we will avoid writing the step that involves

$$\iint\limits_{D}\left(\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz\right)\,dA$$

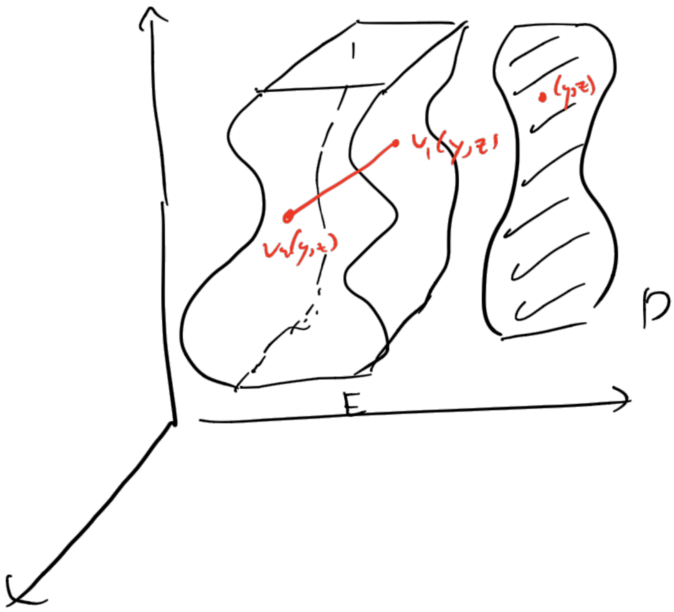
and just immediately write down the triple iterated integral, such as

$$\iiint\limits_{E} z\,dV = \int_{-2}^{\sqrt{2}}\int_{-x}^{x^2}\int_{0}^{\sqrt{4-x^2-y^2}} z\,dz\,dy\,dx.$$

Notice that in doing this, the order of integration is sometimes dictated by the "shape" of the region we are integrating over. In particular, in order to insure that we have a number at the end, we need that the outermost integral has bounds which are individual values; the middle integral may have bounds which are functions of the variable in the outermost integral; and the inner-most integral may have bounds which depend on both of the other variables. These various possibilities basically correspond to the regions we are defining type I (above), type II and type III (below).

We say that a solid is ***type II*** if it has the form

$$E=\left\{(x,y,z)\,\middle|\,(y,z)\in D,\ v_1(y,z)\le x\le v_2(y,z)\right\}$$



In general we would integrate over a type II region using

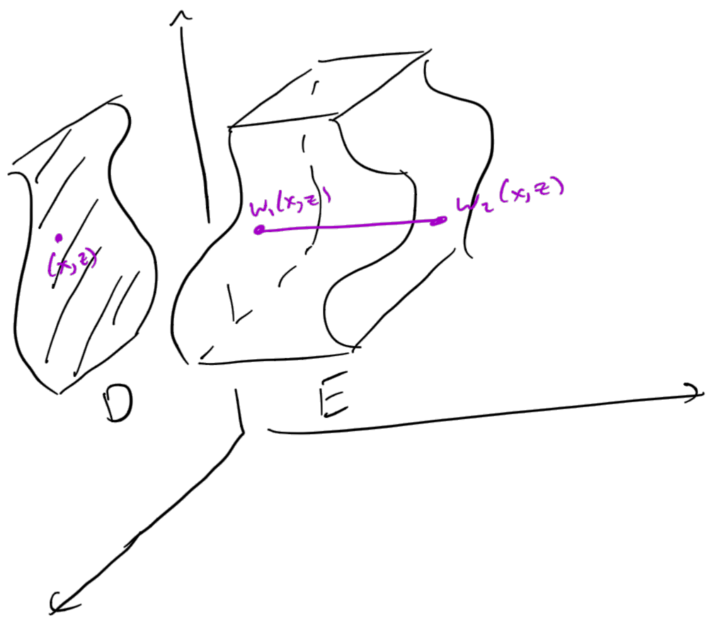
$$\iiint_E f(x,y,z)\,dV = \iint_D \left(\int_{v_1(y,z)}^{v_2(y,z)} f(x,y,z)\,dx \right) dA.$$

When computing the double integral over D, this may become either

$$\int_{y_0}^{y_1} \int_{g_1(y)}^{g_2(y)} \int_{v_1(y,z)}^{v_2(y,z)} f(x,y,z)\,dx\,dz\,dy \quad \text{or} \quad \int_{z_0}^{z_1} \int_{h_1(z)}^{h_2(z)} \int_{v_1(y,z)}^{v_2(y,z)} f(x,y,z)\,dx\,dy\,dz.$$

A ***type III*** solid is of the form

$$E = \{(x,y,z) \,|\, (x,z) \in D, u_1(x,z) \le y \le u_2(x,z)\}$$



The general format for integrating over such a region is

$$\iiint\limits_E f(x,y,z)\,dV = \iint\limits_D \left(\int_{w_1(x,z)}^{w_2(x,z)} f(x,y,z)\,dy \right) dA$$

which may then become

$$\int_{x_0}^{x_1} \int_{g_1(x)}^{g_2(x)} \int_{w_1(x,z)}^{w_2(x,z)} f(x,y,z)\,dy\,dz\,dx \quad \text{or} \quad \int_{z_0}^{z_1} \int_{h_1(z)}^{h_2(z)} \int_{w_1(x,z)}^{w_2(x,z)} f(x,y,z)\,dy\,dx\,dz.$$

We'll end our discussion of triple integrals in Cartesian coordinates by working through one example of integrating over a type III region.

Example 4.11.
Integrate the function

$$f(x,y,z) = x^2 \sin(y)$$

over the solid

$$E = \left\{ (x,y,z) \,\middle|\, 0 \le x \le \sqrt{\pi},\ 0 \le z \le x,\ 0 \le y \le xz \right\}.$$

Since x has definitive bounds, but the other variables have bounds which are functions of x, we must integrate with respect to x last (the outer-most integral). As the bounds for z *do not* depend on y, we may put the z integral in the middle; and since the bounds for y depend on both x and z, we are forced to place the integral with respect to y as the first (inner-most) integral. We then have

$$\begin{aligned}
\iiint_E x^2 \sin(y)\, dV &= \int_0^{\sqrt{\pi}} \int_0^x \int_0^{xz} x^2 \sin(y)\, dy\, dz\, dx \\
&= \int_0^{\sqrt{\pi}} \int_0^x \left(-x^2 \cos(y)\right)\Big|_0^{xz} dz\, dx \\
&= \int_0^{\sqrt{\pi}} \int_0^x \left(-x^2 \cos(xz) + x^2 \cos(0)\right)\, dz\, dx \\
&= \int_0^{\sqrt{\pi}} \int_0^x \left(x^2 - x^2 \cos(xz)\right)\, dz\, dx \\
&= \int_0^{\sqrt{\pi}} \left(x^2 z - x \sin(xz)\right)\Big|_0^x dx \\
&= \int_0^{\sqrt{\pi}} \left(x^3 - x \sin(x^2)\right)\, dx \\
&= \int_0^{\sqrt{\pi}} x^3\, dx - \int_0^{\sqrt{\pi}} x \sin(x^2)\, dx
\end{aligned}$$

The first term if of course simply $\frac{x^4}{4}\Big|_0^{\sqrt{\pi}} = \frac{\pi^2}{4}$. For the second integral we can perform the substitution $u = x^2$, $du = 2x\, dx$ and the integral becomes

$$\frac{1}{2} \int_0^{\pi} \sin(u)\, du = \frac{1}{2} \cos(u)\Big|_0^{\pi} = 1$$

Thus our triple integral equals

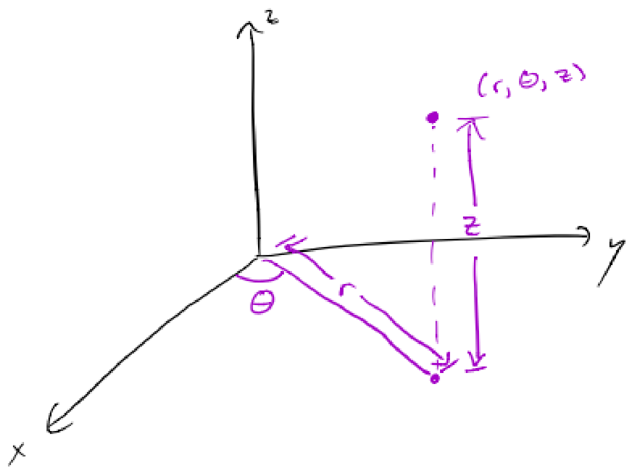
$$\iiint_E x^2 \sin(y)\, dV = \frac{\pi^2}{4} - 1 = \frac{\pi^2 - 4}{4}.$$

4.6 Cylindrical and spherical coordinates

We have seen that when integrating a function of two variables, it is sometimes convenient to rewrite the problem in polar coordinates instead of our familiar Cartesian coordinates. In this section we see the three-variable version of this where we consider two other coordinate systems: cylindrical coordinates and spherical coordinates.

Cylindrical coordinates

Cylindrical coordinates are simply polar coordinates in the xy-plane, together with a z telling us how far above or below the plane we are. That is, in cylindrical coordinates every point in 3-space is represented as a triple (r, θ, z) where r and θ play the same role they played in polar coordinates in the plane, and z tells us how to move away from the xy-plane.



To go back and forth between cylindrical coordinates and Cartesian coordinates, we just convert the given r and θ to x and y exactly as in polar.

Cylindrical to Cartesian	**Cartesian to Cylindrical**
$x = r\cos(\theta)$	$r = \sqrt{x^2 + y^2}$
$y = r\sin(\theta)$	$\theta = \tan^{-1}\left({}^{y}/_{x}\right)$
$z = z$	$z = z$

Now, if we want to integrate over a solid E in 3-space where the projection of E to the xy-plane can be written as

$$D = \left\{(x, y) \,\middle|\, (x, y, z) \in E \text{ for some } z\right\}$$

which can be expressed in polar coordinates as

$$D = \left\{(r, \theta) \,\middle|\, \alpha \leq \theta \leq \beta, h_1(\theta) \leq r \leq h_2(\theta)\right\},$$

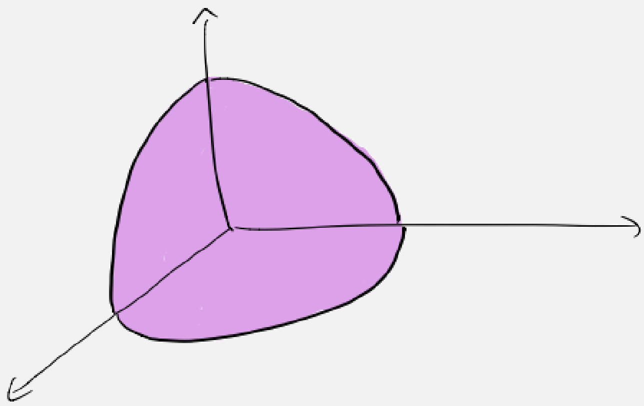
and if the solid E can be written in Cartesian coordinates as

$$E = \left\{(x, y, z) \,\middle|\, (x, y) \in D,\ u_1(x, y) \leq z \leq u_2(x, y)\right\}$$

then we have

$$\begin{aligned}\iiint_E f(x, y, z)\, dV &= \iint_D \left(\int_{u_1(x,y)}^{u_2(x,y)} f(x, y, z)\, dz\right) dA \\ &= \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} \int_{u_1(r\cos(\theta), r\sin(\theta))}^{u_2(r\cos(\theta), r\sin(\theta))} f(r\cos(\theta), r\sin(\theta), z) r\, dz\, dr\, d\theta\end{aligned}$$

Example 4.12.
Setup the iterated integral for integrating $f(x, y, z) = x + y + z$ over the portion of the solid beneath the surface $z = \sqrt{4 - x^2 - y^2}$ in the first octant.



First notice that the region we are integrating over looks like the figure above. Setting $z = 0$, the region D we consider contains (x, y) points satisfying $0 = 4 - x^2 - y^2$, or $x^2 + y^2 = 4$. That is, we have a portion of the circle of radius 2 centered at the origin. However, because we are explicitly only concerned with the first octant in this problem, we only care about the portion of this circle inside the first quadrant in the plane. In polar coordinates this region is represented as

$$D = \left\{ (r, \theta) \,\middle|\, 0 \le r \le 2,\ 0 \le \theta \le {}^{\pi}\!/\!{}_{2} \right\}.$$

Thus our integral becomes

$$\iiint\limits_E f(x,y,z)\, dV = \int_0^{\pi/2} \int_0^2 \int_0^{\sqrt{4-r^2}} (r\cos(\theta) + r\sin(\theta) + z)\, r\, dz\, dr\, d\theta.$$

Sometimes we may initially write an integral in terms of Cartesian coordinates, but then wind up with a very difficult integral to calculate. For example, consider the integral

$$\int_{-2}^{2} \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} \int_{\sqrt{x^2+y^2}}^{2} xz\, dz\, dx\, dy.$$

If you try to immediately start integrating this function, you will realize very quickly that the integration is going to be extremely tedious – doable, but not fun or easy. However, if we rewrite the integral in cylindrical coordinates, it may become easier.

To see that in this particular case, let's first notice that our function $f(x,y,z) = xz$ in cylindrical coordinates would become

$$f(r\cos(\theta), r\sin(\theta), z) = zr\cos(\theta).$$

Projecting our solid to the xy-plane would give us

$$D = \left\{ (x,y) \,\middle|\, -2 \le y \le 2,\ -\sqrt{4-y^2} \le x \le \sqrt{4-y^2} \right\}.$$

Notice this is simply a circle of radius 2 centered at the origin in the plane. In polar coordinates this becomes

$$D = \left\{ (r, \theta) \,\middle|\, 0 \le r \le 2,\ 0 \le \theta \le 2\pi \right\}.$$

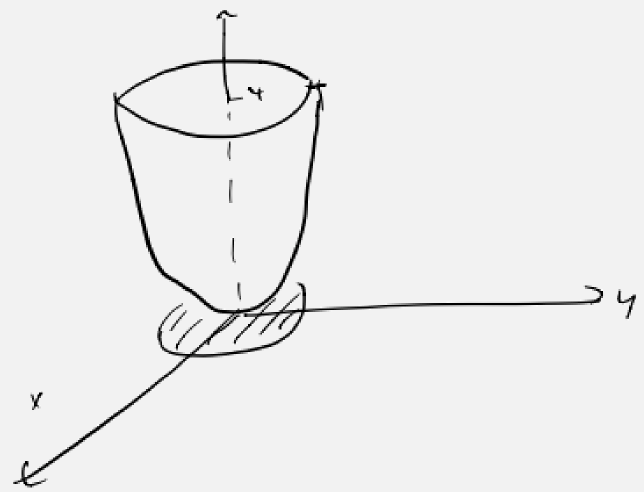
This means our original integral becomes something much more tractable if we rewrite it in cylindrical coordinates:

$$\int_{-2}^{2} \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} \int_{\sqrt{x^2+y^2}}^{2} xz\, dz\, dx\, dy. = \int_0^{2\pi} \int_0^2 \int_r^2 zr\cos(\theta) r\, dz\, dr\, d\theta.$$

Notice this integral is not terribly difficult to compute:

$$\begin{aligned}
\int_0^{2\pi}\int_0^2\int_r^2 zr\cos(\theta)r\,dz\,dr\,d\theta &= \int_0^{2\pi}\int_0^2\int_r^2 r^2 z\cos(\theta)\,dz\,dr\,d\theta \\
&= \int_0^{2\pi}\int_0^2 \left.\frac{r^2z^2\cos(\theta)}{2}\right|_r^2 dr\,d\theta \\
&= \frac{1}{2}\int_0^{2\pi}\int_0^2 \left[4r^2\cos(\theta) - r^4\cos(\theta)\right]\,dr\,d\theta \\
&= \frac{1}{2}\int_0^{2\pi}\left.\left[\frac{4r^3}{3}\cos(\theta) - \frac{r^5}{5}\cos(\theta)\right]\right|_0^2 d\theta \\
&= \frac{32}{15}\int_0^{2\pi}\cos(\theta)\,d\theta \\
&= 0.
\end{aligned}$$

Example 4.13.
As a final example, let's compute the volume of the following solid in 3-space: the solid E will consist of the points (x, y, z) between the paraboloid $z = x^2 + y^2$ and $z = 4$.



In cylindrical coordinates, notice the "shadow" D of the region we are integrating over is a disc of radius 2, which is the polar rectangle $[0, 2] \times [0, 2\pi]$. The z-values begin at the paraboloid $z = x^2 + y^2$ and increase up to $z = r$. Since $x = r\cos(\theta)$ and $y = r\sin(\theta)$, though, we may express the paraboloid in cylindrical coordinates as $z = r^2$.

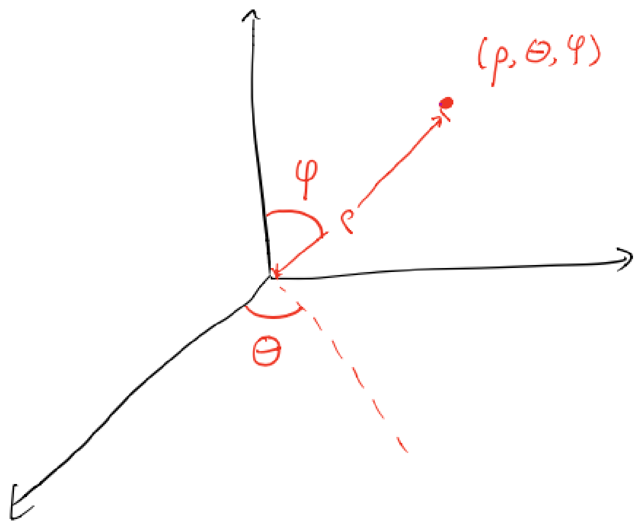
Thus we can now compute the volume of this solid as

$$\begin{aligned}
\mathrm{Vol}(E) &= \iiint_E 1\,dV \\
&= \int_0^2 \int_0^{2\pi} \int_{r^2}^4 r\,dz\,d\theta\,dr \\
&= \int_0^2 \int_0^{2\pi} rz\Big|_{r^2}^4 \,d\theta\,dr \\
&= \int_0^2 \int_0^{2\pi} (4r - r^3)\,d\theta\,dr \\
&= \int_0^2 2\pi(4r - r^3)\,dr \\
&= 2\pi\left(2r^2 - \frac{r^4}{4}\right)\Big|_0^2 \\
&= 2\pi\,(8 - 4) \\
&= 8\pi
\end{aligned}$$

Spherical coordinates

In spherical coordinates we specify a point in 3-space as a triple (ρ, θ, φ) where

- ρ is the distance from the point to the origin;
- θ is the counterclockwise angle in the xy-plane measured from the positive x-axis; and
- φ is the angle in the yz-plane measured from the positive z-axis towards the positive y-axis.



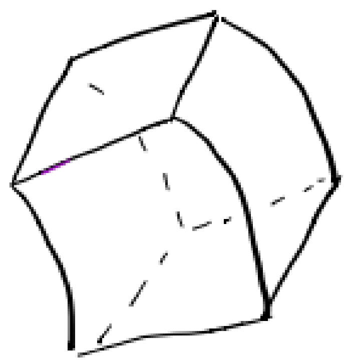
By convention we only consider φ between 0 and π.

Working through the trigonometry involved tells us that these ρ, θ, and φ values in spherical coordinates are related to our usual Cartesian coordinates by the following expressions:

$$x = \rho\sin(\varphi)\cos(\theta),$$
$$y = \rho\sin(\varphi)\sin(\theta), \text{ and}$$
$$z = \rho\cos(\varphi).$$

Notice that the basic object expressed in spherical coordinates as a product of three intervals is a small "chunk" of a solid sphere, referred to as a ***spherical wedge***:

$$E = \left\{(\rho,\theta,\varphi)\,\middle|\, a \le \rho \le b,\ \alpha \le \theta \le \beta,\ c \le \varphi \le d\right\}.$$



In order to integrate over such a region, we split the wedge up into smaller "sub-wedges," pick a point in each of these sub-wedges, evaluate the function at that point, and multiply by the volume of the wedge, and add up all of the pieces. To integrate a function given in Cartesian coordinates as $f(x, y, z)$, this means we would have to consider limits of Riemann sums of the following form:

$$\sum_{i=1}^{\ell}\sum_{j=1}^{m}\sum_{k=1}^{n} f(\rho_i^* \sin(\varphi_k^*)\cos(\theta_j^*),\ \rho_i^* \sin(\varphi_k^*)\sin(\theta_j^*),\ \rho_i^* \cos(\varphi_k^*))\Delta V_{ijk}$$

where ΔV_{ijk} is the volume of the ijk-th subwedge. We will approximate this wedge by a rectangular prism of the same dimensions. (As our wedges get smaller and smaller the error in this approximation will go to zero, so everything will work out when we take limits.)

Some simple geometry will tell us the dimensions of our prism should thus be $\rho_i \sin(\varphi_k)\Delta\theta_j \times \rho_i\Delta\varphi_k \times \Delta\rho_i$. That is, the volume of our prism, which approximates the volume of the wedge, is

$$\begin{aligned}\Delta V_{ijk} &\approx \rho_i \sin(\varphi_k)\Delta\theta_j \rho_i \Delta\varphi_k \Delta\rho_i \\ &= \rho_i^2 \sin(\varphi_k)\Delta\rho\Delta\theta\Delta\varphi.\end{aligned}$$

Putting all of this together, this means the integral of a Cartesian function $f(x, y, z)$ over a spherical wedge can be written in spherical coordinates as a limit of of the following Riemann sums:

$$\sum_{i=1}^{\#\mathcal{P}}\sum_{j=1}^{\#\mathcal{Q}}\sum_{k=1}^{\#\mathcal{R}} f(\rho_i^* \sin(\varphi_k^*)\cos(\theta_j^*),\ \rho_i^* \sin(\varphi_k^*)\sin(\theta_j^*),\ \rho_i^* \cos(\varphi_k^*))(\rho_i^*)^2 \sin(\varphi_k^*)\Delta\rho_i\Delta\theta_j\Delta\varphi_k.$$

Taking the limits, the integral over becomes

$$\iiint_E f(x, y, z)\, dV = \int_c^d \int_\alpha^\beta \int_a^b f(\rho\sin(\varphi)\cos(\theta), \rho\sin(\varphi)\sin(\theta), \rho\cos(\varphi))\rho^2\sin(\varphi)d\rho d\theta d\varphi.$$

While this integral looks atrocious at first glance, in some problems many of the "ugly" expressions that appear above greatly simplify using some trig identities.

Example 4.14.
Integrate the function $f(x, y, z) = (x^2 + y^2 + z^2)^2$ over the sphere of radius 5 centered at the origin.

Let's notice that this sphere *is* in fact a spherical wedge. The radii

ρ range from 0 to 5; the angles θ range from 0 to 2π; and the angles φ range from 0 to π.

Our function actually simplifies quite a lot in spherical coordinates after just a little bit of work:

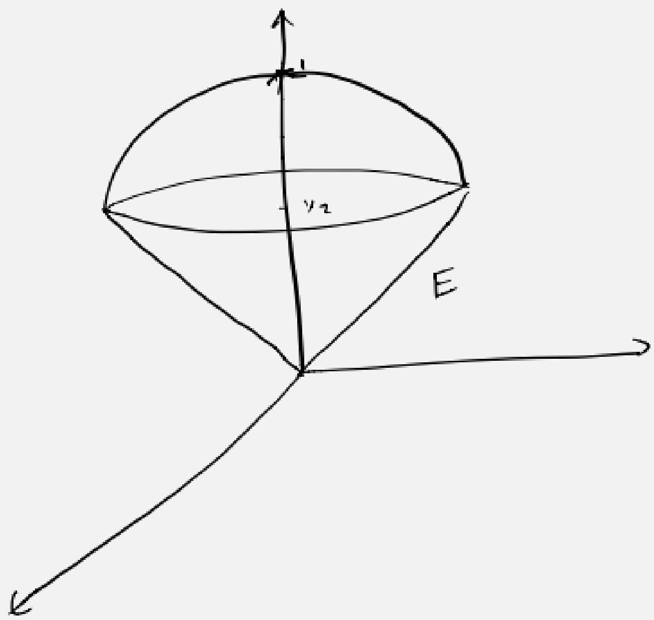
$$\begin{aligned}
&\left(x^2+y^2+z^2\right)^2 \\
=&\left(\rho^2\sin^2(\varphi)\cos^2(\theta)+\rho^2\sin^2(\varphi)\sin^2(\theta)+\rho^2\cos^2(\varphi)\right)^2 \\
=&\left(\rho^2\sin^2(\varphi)(\cos^2(\theta)+\sin^2(\theta))+\rho^2\cos^2(\varphi)\right)^2 \\
=&\left(\rho^2(\sin^2(\varphi)+\cos^2(\varphi)\right)^2 \\
=&\rho^4
\end{aligned}$$

Our integral is thus

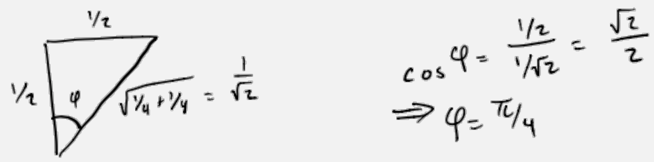
$$\begin{aligned}
\iiint\limits_B f(x,y,z)\,dV &= \int_0^{\pi}\int_0^{2\pi}\int_0^5 \rho^4\rho^2\sin(\varphi)d\rho d\theta d\varphi \\
&= \int_0^{\pi}\int_0^{2\pi}\int_0^5 \rho^6\sin(\varphi)d\rho d\theta d\varphi \\
&= \int_0^{\pi}\int_0^{2\pi} \sin(\varphi)\frac{\rho^7}{7}\bigg|_0^5 d\theta d\varphi \\
&= \int_0^{\pi}\int_0^{2\pi} \frac{5^7}{7}\sin(\varphi)\,d\theta d\varphi \\
&= \frac{5^7}{7}\int_0^{\pi}\sin(\varphi)\theta\bigg|_0^{2\pi} d\varphi \\
&= \frac{5^7}{7}2\pi\int_0^{\pi}\sin(\varphi)d\varphi \\
&= \frac{5^7}{7}2\pi\left(-\cos(\varphi)\bigg|_0^{\pi}\right) \\
&= \frac{5^7}{7}2\pi\cdot(1+1) \\
&= \frac{4\pi 5^7}{7}
\end{aligned}$$

As one final example, let's compute the volume of a certain three-dimensional solid using spherical coordinates.

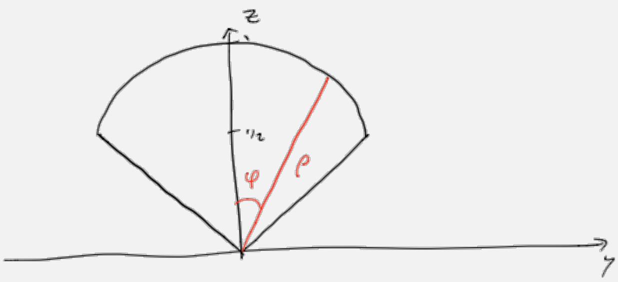
Example 4.15.
Compute the volume of the solid pictured below which is given by the cone the region in 3-space above the $z = \sqrt{x^2 + y^2}$ for $0 \leq z \leq 1/2$, but below the sphere of radius $1/2$ for $1/2 \leq z \leq 1$.



To get the volume we will want to integrate the constant function 1 over this solid. Notice that if we project this region down to the xy-plane we would see a circle. This tells us that the θ values range from 0 to 2π. By considering a vertical cross-section of this object (i.e., intersecting with the yz-plane), and performing a little bit of trigonometry, we can see that the angles φ range from 0 (top of the solid) to $\pi/4$, along the outside cone of the solid.



So, the only remaining question is what is the range of ρ? Again, consider the vertical cross section (intersection with the yz-plane), we see that ρ starts at 0 (the origin is a point in our solid) and extends up until it hits some point on the sphere, which we may interpret as a function of φ.



Since the top of our solid is given by the top hemisphere of the sphere of radius $1/2$ centered at $(0, 0, 1/2)$, we can easily write down the Cartesian equation for this sphere:

$$x^2 + y^2 + (z - 1/2)^2 = 1/4.$$

Converting this to spherical coordinates we have the following:

$$\begin{aligned} & \rho^2 \sin^2(\varphi)\cos^2(\theta) + \rho^2 \sin^2(\varphi)\sin^2(\theta) + (\rho\cos(\varphi) - 1/2)^2 = 1/4 \\ \Longrightarrow\ & \rho^2 \sin^2(\varphi) + \rho^2\cos^2(\varphi) - \rho\cos(\varphi) + 1/4 = 1/4 \\ \Longrightarrow\ & \rho^2 - \rho\cos(\varphi) = 0 \\ \Longrightarrow\ & \rho^2 = \rho\cos(\varphi) \\ \Longrightarrow\ & \rho = \cos(\varphi). \end{aligned}$$

Thus ρ ranges from 0 to $\cos(\varphi)$.

We can now write down our integral and compute the volume of the solid.

$$\begin{aligned}
\mathrm{Vol}(E) &= \iiint\limits_E 1\,dV \\
&= \int_0^{2\pi}\int_0^{\pi/4}\int_0^{\cos(\varphi)} \rho^2\sin(\varphi)d\rho\,d\varphi\,d\theta \\
&= \int_0^{\pi/4}\int_0^{\cos(\varphi)} \rho^2\sin(\varphi)d\rho d\varphi \cdot \int_0^{2\pi} d\theta \\
&= 2\pi\int_0^{\pi/4} \sin(\varphi)\int_0^{\cos(\varphi)} \rho^2 d\rho d\varphi \\
&= 2\pi\int_0^{\pi/4} \cdot\frac{\rho^3}{3}\Big|_0^{\cos(\varphi)} d\varphi \\
&= \frac{2\pi}{3}\int_0^{\pi/4} \cos^3(\varphi)\sin(\varphi)\,d\varphi
\end{aligned}$$

Now we can perform the substitution $u = \cos(\varphi)$, $du = -\sin(\varphi)d\varphi$ and the integral becomes

$$\begin{aligned}
\frac{-2\pi}{3}\int_1^{\sqrt{2}/2} u^3\,du &= \frac{2\pi}{3}\int_{\sqrt{2}/2}^1 u^3\,du \\
&= \frac{2\pi}{3}\cdot\frac{u^4}{4}\Big|_{\sqrt{2}/2}^1 \\
&= \frac{\pi}{6}\left(1 - \frac{4}{16}\right) \\
&- \frac{\pi}{6}\cdot\frac{3}{4} \\
&= \frac{\pi}{8}
\end{aligned}$$

4.7 Change of variables

With polar coordinates for integrals in two variables, and cylindrical and spherical coordinates for integrals in three variables, we saw that the integral necessarily contained some "extra" factors: integrals in polar and cylindrical coordinates contained an factor of r, while integrals in spherical coordinates contained an extra $\rho^2\sin(\varphi)$.

In this section we will introduce the higher-dimensional equivalent of u-substitution and see that the extra factors of r and $\rho^2 \sin(\varphi)$ essentially arise as the "du" when a certain substitution is performed for integrals in two or three variables.

We will taking a few "baby steps" towards the general change of variables principle by first reviewing u-substition in one variable when the u is linear to concretely understand why we need to change th bounds on a definite integral when performing u-substitution and why the du that appears requires we divide the integral by the constant that appears in the substitution. We will then generalize this to understand linear changes of coordinates in multiple variables, and finally generalize to the non-linear situation.

Linear changes of coordinates in one variable

Suppose we wished to integrate a function in one variable of the following form:

$$\int_a^b f(kx)\, dx$$

where k is any constant. For example, perhaps we wish to integrate $\int_0^{\pi/4} \sin(2x)\, dx$. We can simplify our problem by performing a substitution to turn kx into a single variable, which we usually call u. The substitution $u = kx$ will take $f(kx)$ and turn it into $f(u)$. However, we can not simply write our integral as $\int_a^b f(u)\, du$. The reason why is often glossed over in a first or second semester calculus class, but let's take a minute to try to understand what's happening.

In Figure 4.6 below, we have the shaded area which would be computed by $\int_0^{\pi/4} \sin(2x)\, dx$.

Now we wish to perform the substitution $u = 2x$. If we were to simply naively replace $2x$ in the integral with u and dx with du, but not make any other changes, the area we would calculate would correspond to the shaded region in Figure 4.7.

It's clear from the picture this is not the same as the area in Figure 4.6. The reason is that we have introduced new coordinates (the u) which play a different role from our original coordinates (the x). In particular, we need to measure u's differently from the way we measure x's. Since u is twice the value of x, the x-interval $[0, \pi/4]$ is only half as large as it needs to be. That is, when performing calculations in the u-coordinates we have introduced, we need to stretch our x-coordinates out by a factor of two, since each unit of u is two units of x. That is, we need to extend

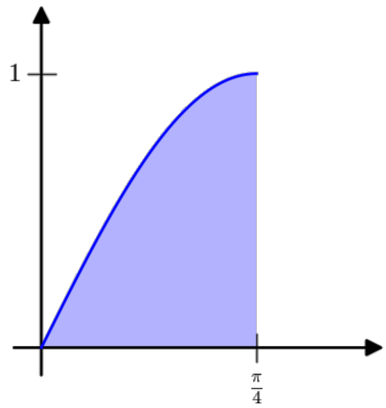


Figure 4.6: The area corresponding to $\int_0^{\pi/4} \sin(2x)\,dx$.

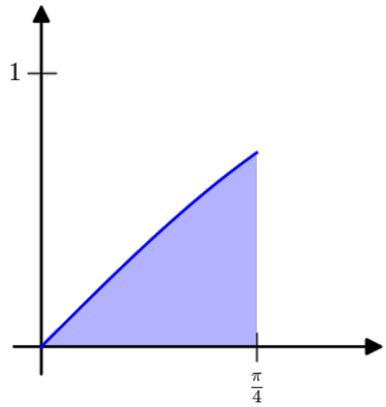


Figure 4.7: The area corresponding to $\int_0^{\pi/4} \sin(u)\,du$.

our interval to $[0, \pi/2]$. Making this change to our integral, we would be computing the area of the shaded region in Figure 4.8.

Now notice the area shown in Figure 4.8 *is not* what we want: it's way too large. If we want our substitution to preserve area (which we definitely *do* want if we want substitution to be useful for calculating integrals), then we need to somehow fix the fact our current area calculation is too large. The issue here is that we stretched our region out in the horizontal direction by a factor of two, while the vertical direction didn't change. It's not too hard to see that this will cause our area to be twice as large as it should be. (For instance, imagine approximating the original region as a Riemann sum by adding up areas of rectangles. If we make every rectangle twice as wide, but don't modify the height, then each rectangle will now have twice the area, so our approximation to the

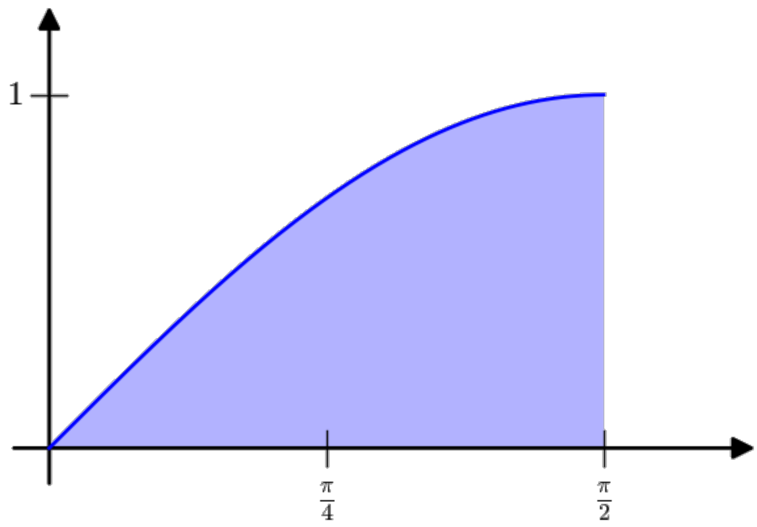


Figure 4.8: The area corresponding to $\int_0^{\pi/2} \sin(u)\,du$.

area grows by a factor of two. Since this happens for *every* Riemann sum, that will mean the limit – our integral – also increases by a factor of two.) Since we stretched horizontally by a factor of two and doubled the area, we will compensate by scaling vertically by one half, and this will correct our area calculation. That is, the correct substitution for $\int_0^{\pi/4} \sin(2x)\,dx$ is $\frac{1}{2}\int_0^{\pi/2} \sin(u)\,du$, which is pictured in Figure 4.9

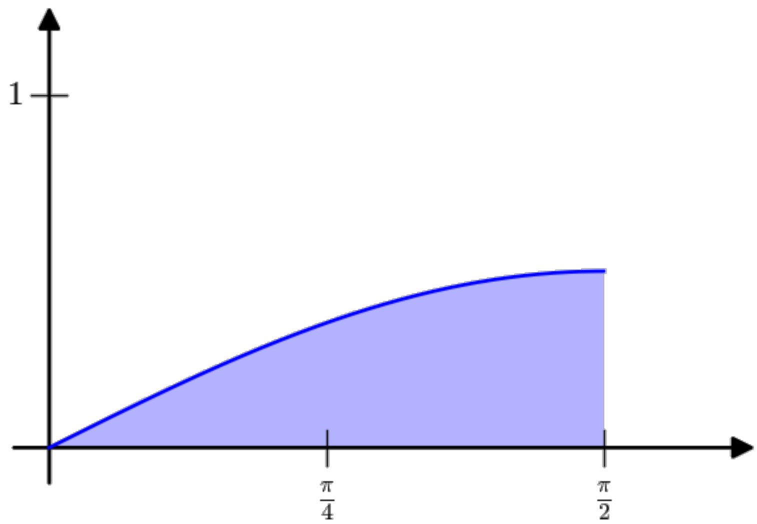


Figure 4.9: The area corresponding to $\frac{1}{2}\int_0^{\pi/2} \sin(u)\,du$.

To be slightly more general, if $T(x) = kx$ is a linear function of one variable, then mimicking the discussion above will tell us that the proper substitution for $\int_a^b f(T(x))\,dx$ is $\frac{1}{\det(T)}\int_{T(a)}^{T(b)} f(u)\,du$. Let's notice, though, that u is really $T^{-1}(x)$, and so we may rewrite this integral in terms of x as $\frac{1}{\det(T)}\int_{T(a)}^{T(b)} f(T^{-1}(x))\,dx$.

Being explicit, since $T(x) = kx$, this is $\frac{1}{k}\int_{ka}^{kb} f(\frac{1}{k}x)\,du$. For instance, in our example above we had $T(x) = 2x$ and $f(x) = \sin(2x)$. This tells us that

$$\int_0^{\pi/4} \sin(2x)\,dx = \frac{1}{2}\int_0^{\pi/2} \sin(\frac{1}{2}2x)\,dx.$$

This is exactly our u-substitution above, just written in terms of x instead of u.

We are writing the expression above in terms of $T(a)$, $T(b)$, and $\frac{1}{\det(T)}$ to see how the expression will generalize to higher dimensions.

Linear change of coordinates in several variables

Suppose now that we have a double integral over a region $D \subseteq \mathbb{R}^2$, $\iint\limits_D f(x,y)\,dA$. What is the two-dimensional equivalent of the (linear) substitution described above? Since we are in two dimensions, notice that "linear" means more than simply "multiply x by a constant." In particular, we have two variables, so we can multiply each one by constants – and these need not be the same constant. We could also add a multiple of one variable to the other. That is, our transformation $T(x,y)$ could give us back the coordinates $(ax+by, cx+dy)$. Interpreting (x,y) as a column vector, $\begin{pmatrix} x \\ y \end{pmatrix}$, notice that this result is the same as if were were to perform a matrix multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

That is, the "linear" T here is a linear transformation, as was described in Section 1.4. Just as applying a one-variable linear transformation in the previous subsection made the size of our region change (e.g., $[0, \pi/4]$ was mapped to $[0, \pi/2]$ which was twice as large), applying a two-variable linear transformation (aka, multiplying by a 2×2 matrix) will make the area of the region change. As we had seen before, the way the area of a region in the plane changes when a linear transformation is applied is that the area scales by the determinant of the matrix,

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

This tells us that the two-variable case, the "correct" substitution is given by

$$\iint_D f(x,y)\,dA = \frac{1}{\det(T)} \iint_{T(D)} f(T^{-1}(x,y))\,dA.$$

Where $T(D)$ refers to the "image" of D under the transformation T. That is, if we applied T to every single point of D, the collection of all outputs is the region we are integrating over on the right-hand side. The $T^{-1}(x,y)$ is the inverse of T applied to (x,y). That is, $T^{-1}(x,y)$ is the point (u,v) such that $T(u,v) = (x,y)$. In terms of 2×2 matrices, the inverse of the matrix determined by $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ is

$$\begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Even in this linear case writing down the exact details of the integral can be tricky, but we will consider one very simple example. Suppose the function we wish to integrate is $f(x,y) = 3x\sin(2y)$ over the rectangle $D = [0,1] \times [0, \pi/4]$. To compare to our integral after performing the substitution, let's go ahead and notice the value of this integral is simply

$$\int_0^{\pi/4} \int_0^1 3x\sin(2y)\,dx\,dy = \int_0^{\pi/4} \sin(2y)\,dy \int_0^1 3x\,dx = \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}$$

Now let's apply the transformation that corresponds to the matrix

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

Notice the determinant of this matrix is 6. Using the formula above, the inverse of our transformation is given by the matrix

$$\begin{pmatrix} \frac{2}{6} & 0 \\ 0 & \frac{3}{6} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

Observe that whereas our original transformation stretched the x by 3 and the y by 2, the inverse contracts the x by $\frac{1}{3}$ and the y by $\frac{1}{2}$. This makes intuitive sense as the inverse T^{-1} is supposed to "undo" the operation of T.

As our transformation T simply stretches horizontally by a factor of 3 and vertically by a factor of 2, the image of our original rectangle $D = [0,1] \times [0, \pi/4]$ is $T(D) = [0,3] \times [0, \pi/2]$. Thus our transformed integral is

$$\begin{aligned}
\frac{1}{6}\int_0^{\pi/2}\int_0^3 f(T^{-1}(x,y))\,dx\,dy &= \frac{1}{6}\int_0^{\pi/2}\int_0^3 f(\frac{1}{3}x, \frac{1}{2}y)\,dx\,dy \\
&= \frac{1}{6}\int_0^{\pi/2}\int_0^3 3\frac{1}{3}x\sin(2\frac{1}{2}y)\,dx\,dy \\
&= \frac{1}{6}\int_0^{\pi/2}\int_0^3 x\sin(y)\,dx\,dy \\
&= \frac{1}{6}\int_0^{\pi/2}\sin(y)\,dy\int_0^3 x\,dx \\
&= \frac{1}{6}\cdot 1\cdot\frac{9}{2} \\
&= \frac{9}{12} \\
&= \frac{3}{4}.
\end{aligned}$$

You might notice that in this case we're essentially applying two substitutions simultaneously: we are effectively performing the substitution $u = 3x$ and also the substitution $v = 2y$. In fact, writing down the iterated integrals of the original double integral, you see these two substitutions can be applied, and the u-substitution would make us multiply our integral by $\frac{1}{3}$ while the v-substitution would have us multiply by $\frac{1}{2}$. Performing both of these together, we ultimately multiply our matrix by $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$, which is the $\frac{1}{\det(T)}$ that appeared in our integral.

The example above was particularly simple because we were integrating over a rectangle and our linear transformation took our rectangle to another rectangle. This need not always happen: linear transformations can also rotate and shear regions, so rectangles don't have to map to rectangles.

A similar formula holds in higher dimensions. In three dimensions, for example, we have

$$\iiint\limits_E f(x,y,z)\,dV = \frac{1}{\det(T)}\iiint\limits_{T(E)} f(T^{-1}(x,y,z))\,dV.$$

where T is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^3$ (aka a 3×3 matrix).

Even though we haven't really dealt with integrals in more than three variables, the pattern continues to any number of dimensions.

Non-linear changes of coordinates in one variable

Earlier we considered a u-substitution which was particularly simple: $u = kx$, and so every x-value was multiplied by the same constant k. Of course, u-substitutions can be more complicated than this. In general, we may wish to perform the substitution $u = g(x)$ to rewrite the integral

$$\int_a^b f(g(x))\,dx$$

as something simpler. Just as with a linear change of coordinates, we can't simply write $\int_a^b f(u)\,du$, since the bounds of our integral will typically change, and the size of our inputs to the function will change as well. It's pretty easy to correct the inputs: the x-coordinate a corresponds to the u-coordinate $g(a)$, and similarly for b. So we expect the bounds on the new integral to be from $g(a)$ to $g(b)$. The integral $\int_{g(a)}^{g(b)} f(u)\,du$ is still not quite what we want: the inputs to our function could scale to get larger which will change the area of our regions, and we'll need to compensate by contracting the outputs. How should we do this, though, since the different x-values may scale by different amounts? The key is to use the derivative to approximate the change on small scales. The derivative $g'(x_0)$ tells us the rate of change of $g(x)$ at the point x_0. That is, it tells us how quickly the outputs of the function change as the inputs are changing. The outputs of $g(x)$ correspond to the inputs to $f(u)$, since $f(u)$ really means the composition $f(g(x))$. That is, at each point x_0 the inputs to f will scale by $g'(x_0)$, so we should divide the output by $g'(x_0)$ to compensate. This value changes from point to point, though; it's not a constant, so we can't pull it out of the integral. But this tells us the change of coordinates $u = g(x)$ turns the integral

$$\int_a^b f(g(x))\,dx$$

into

$$\int_{g(a)}^{g(b)} f(g^{-1}(x))\frac{1}{g'(g^{-1}(x))}\,dx.$$

Notice this really is your usual u-substitution rule, just written down differently. Normally you'd write $u = g(x)$, $du = g'(x)dx$, but these may be rewritten as $x = g^{-1}(u)$ and $dx = \frac{1}{g'(x)}du = \frac{1}{g'(g^{-1}(u))}du$. This tells us

$$\int_{g(a)}^{g(b)} f(u)du = \int_{g(a)}^{g(b)} f(g^{-1}(u))\frac{1}{g'(g^{-1}(u))}\,du = \int_{g(a)}^{g(b)} f(g^{-1}(x))\frac{1}{g'(g^{-1}(x))}\,dx.$$

where the last step follows by simply replacing the "dummy variable" u with x.

As an example, consider $\int_1^{\sqrt{e}} 2x\ln(x^2)\,dx$. Performing the "standard" u-substitution tricks from a first or second semester calculus course, you would say $u = x^2$, $du = 2x\,dx$ and write the integral as

$$\int_1^e \ln(u)\,du = \frac{1}{u}\bigg|_1^e = \frac{1}{e} - 1.$$

To compare this to the new way we have described substitution, we need to interpret $2x\ln(x^2)$ as $f(g(x))$ where $g(x) = x^2$. This means our $f(x)$ is $f(x) = 2\sqrt{x}\ln(x)$. Notice $g'(x) = 2x$ and $g^{-1}(x) = \sqrt{x}$. Using our formula above,

$$\int_{g(a)}^{g(b)} f(g^{-1}(x))\frac{1}{g'(g^{-1}(x))}\,dx$$

we have

$$\begin{aligned}\int_1^e f(\sqrt{x})\frac{1}{2\sqrt{x}}\,dx &= \int_1^e 2\sqrt{x}\ln(x)\cdot\frac{1}{2\sqrt{x}}\,dx\\ &= \int_1^e \ln(x)\,dx\\ &= \frac{1}{x}\bigg|_1^e\\ &= \frac{1}{e} - 1\end{aligned}$$

Non-linear changes of coordinates in several variables

To carry about the above discussion for functions of several variables, we need to understand how a general (*not necessarily linear!*) transformation of several variables distorts areas in two dimensions and volumes in three dimensions. To do this will require a little bit of setup before we can see how our integrals are modified.

In general, a transformation that takes points in the plane to other points in the plane (equivalently, two-dimensional vectors are transformed to other two-dimension vectors) may be written as $\Phi(x, y) = (u(x, y), v(x, y))$, or in vector form

$$\Phi\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}.$$

That is, the function u takes two variables (x and y) as inputs and outputs a single real number which will play the role of the x-coordinate of our new vector, and similarly v takes two variables as inputs and gives a single real number as an output which will be our y-coordinate. In principle these functions can be completely arbitrary: they don't necessarily have to be linear, or continuous, or differentiable, or have any other nice property we like. Understanding general functions is actually much more difficult than understanding functions with "nice" properties though, and so we will always suppose that the components of Φ are ***C^1-functions***, meaning they have continuous first-order partial derivatives with respect to every variable. That is, $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$ all exist and are continuous. As a consequence, the function Φ is differentiable and we can consider its total derivative $D\Phi$. Since Φ has two-dimensional inputs and two-dimensional outputs, its total derivative at every point will be a 2×2 matrix and the terms of this matrix are precisely the partial derivatives listed above:

$$D\Phi = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

This matrix is sometimes also called the ***Jacobian (matrix)*** of Φ.

Recall from Section 3.5 that the total derivative of a function at a point is the linear transformation "best" linear approximation to the function at that point. Thus to understand how Φ changes areas, we should approximate Φ by $D\Phi$ near each point and consider how $D\Phi$ changes areas. As $D\Phi$ is linear, though, we know how $D\Phi$ changes area: the areas scale by the determinant $\det(D\Phi)$. This number, $\det(D\Phi)$ is sometimes also called the ***Jacobian (determinant)*** of Φ.

Remark.
The language is a little bit confusing because sometimes people will simply say "Jacobian" without specifying if they mean the matrix or its determinant. Often if the word "Jacobian" is used without any other quantifiers or context, it will mean the Jacobian determinant. This is sometimes denoted $\operatorname{Jac}(\Phi)$ or $\left|\frac{\partial(u,v)}{\partial(x,y)}\right|$.

Example 4.16.
Compute the total derivative (Jacobian matrix) of the following function at the point $(3, 2)$ and compute its Jacobian determinant:

$$\Phi(x, y) = (\frac{x^3}{3} - y^3 - y, x + \frac{y^3}{3}).$$

In this example we have $u(x, y) = \frac{x^3}{3} - y^3 - y$ and $v(x, y) = x + \frac{y^3}{3}$. We first compute the partial derivatives of our functions:

$$\begin{aligned} u_x(x, y) &= x^2 & u_y(x, y) &= -3y^2 - 1 \\ v_x(x, y) &= 1 & v_y(x, y) &= y^2 \end{aligned}$$

Evaluating these functions at $(3, 2)$ gives us

$$\begin{aligned} u_x(3, 2) &= 9 & u_y(3, 2) &= -13 \\ v_x(3, 2) &= 1 & v_y(3, 2) &= 4 \end{aligned}$$

Thus our Jacobian matrix is

$$\begin{pmatrix} 9 & -13 \\ 1 & 4 \end{pmatrix}$$

and the Jacobian determinant is

$$\det \begin{pmatrix} 9 & -13 \\ 1 & 4 \end{pmatrix} = 9 \cdot 4 - 3 \cdot (-13) = 36 + 13 = 49.$$

What the Jacobian determinant of 75 from Example 4.16 tells us is that if Φ was used as a linear a change of coordinates, then near the point $(3, 2)$ the map Φ will scale areas by a factor of about 75. (The exact value will depend on the exact point under consideration, but since the Jacobian determinant will be a continuous function of (x, y), points near $(3, 2)$ will result in determinants near 75.)

In terms of integrals, this means if we apply a function Φ to transform a region $D \subseteq \mathbb{R}^2$ to $\Phi(D)$, at each point we need to divide out the Jacobian determinant to compensate for the change in areas. That is,

$$\iint_D f(x,y)\,dA = \iint_{\Phi(D)} f(\Phi^{-1}(x,y))\frac{1}{\det(D\Phi(\Phi^{-1}(x,y)))}\,dA.$$

Before doing any examples, let's try to simplify the notation above a little bit. Many texts will avoid the derivations above and jump to the "simpler" formula we are about to describe, but we wanted to describe where everything was coming from before first. The first observation is that the expression

$$\frac{1}{\det(D\Phi(\Phi^{-1}(x,y)))}$$

is really the determinant of $D\Phi^{-1}(x,y)$. (It's easy, albeit tedious, to verify this for 2×2 matrices, and the general result follows easily from basic linear algebra.) This turns our integral above into

$$\iint_D f(x,y)\,dA = \iint_{\Phi(D)} f(\Phi^{-1}(x,y))\det(D\Phi^{-1}(x,y))\,dA.$$

The next observation is that if we apply Φ^{-1} instead of Φ, the roles of Φ and Φ^{-1} in the above swap and we have

$$\iint_D f(x,y)\,dA = \iint_{\Phi^{-1}(D)} f(\Phi(x,y))\det(D\Phi(x,y))\,dA.$$

Now recall that the choice of variables x and y, or α and β, or r and θ is immaterial. For this reason, people sometimes write the right-hand integral with variables u and v, and interpret Φ as mapping from the (u,v)-plane to the (x,y)-plane, and the integral becomes

$$\iint_D f(x,y)\,dA = \iint_{\Phi^{-1}(D)} f(\Phi(u,v))\det(D\Phi(u,v))\,dA.$$

Replacing D by $\Phi(D)$ we have

$$\iint_{\Phi(D)} f(x,y)\,dA = \iint_D f(\Phi(u,v))\det(D\Phi(u,v))\,dA.$$

That is, we think of Φ as taking (u,v)-coordinates and converting them to (x,y)-coordinates. The determinant $\det(D\Phi(u,v))$ is then sometimes denoted

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$

and the integral written

$$\iint_{\Phi(D)} f(x,y)\,dA = \iint_{D} f(\Phi(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dA.$$

Remark.
Though it's sometimes confusing, it's not uncommon to reverse the roles of (x,y) and (u,v) in a problem and write the integral above as

$$\iint_{\Phi(D)} f(u,v)\,dA = \iint_{D} f(\Phi(x,y)) \left|\frac{\partial(u,v)}{\partial(x,y)}\right| dA.$$

These changes really don't matter for the mathematics: the distinctions are more psychological. It probably sounds strange the first time you learn this material, but it can sometimes be easier to think of a problem in one of the several different ways we've described the integral above than it is to think of one of the other integrals. Though the actual math is all the same, it can sometimes be easier to wrap your head around one of these various representations for a given problem.

Example 4.17.
Suppose we have a mapping Φ from the (u,v)-plane to the (x,y)-plane given by $\Phi(u,v) = (u\cos(v), u\sin(v))$. That is, $x = u\cos(v)$ and $y = u\sin(v)$. For a function $f(x,y)$ specified in (x,y)-coordinates, notice the function would become $f(u\cos(v), u\sin(v))$. The Jacobian

would be

$$\begin{aligned}
\left|\frac{\partial(x,y)}{\partial(u,v)}\right| &= \det\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \\
&= \det\begin{pmatrix} \frac{\partial}{\partial u} u\cos(v) & \frac{\partial}{\partial v} u\cos(v) \\ \frac{\partial}{\partial u} u\sin(v) & \frac{\partial}{\partial v} u\sin(v) \end{pmatrix} \\
&= \det\begin{pmatrix} \cos(v) & -u\sin(v) \\ \sin(v) & u\cos(v) \end{pmatrix} \\
&= \cos(v)\cdot u\cos(v) - \sin(v)\cdot(-u\sin(v)) \\
&= u\cos^2(v) + u\sin^2(v) \\
&= u
\end{aligned}$$

That is, the change of variables becomes

$$\iint\limits_{\Phi(D)} f(x,y)\,dA = \iint\limits_{D} f(u\cos(v), u\sin(v))u\,dA$$

If D happens to be a rectangle $[a,b]\times[\alpha,\beta]$, notice the right-hand side becomes

$$\int_\alpha^\beta \int_a^b f(u\cos(v), u\sin(v))u\,du\,dv.$$

You may notice the example above looks strikingly similar to our formula for integration in polar coordinates. It's slightly tedious to work out the details, but if D is a rectangle then $\Phi(D)$ (using the Φ in the example above) will be polar rectangle. As mentioned many times the actual names of our variables don't matter so much, and so if we were to replace u by r and v by θ, the change of variables formula from the example could be rewritten as

$$\iint\limits_{\Phi(D)} f(x,y)\,dA = \int_\alpha^\beta \int_a^b f(r\cos(\theta), r\sin(\theta))r\,dr\,d\theta.$$

That is, we recover the formula for integrating a Cartesian function $f(x,y)$ over a polar rectangle D corresponding to $[a,b]\times[\alpha,\beta]$ in polar coordinates. Notice the r that appears is really the Jacobian determinant of the function Φ that performs the polar-to-Cartesian conversion.

Example 4.18.
Suppose D is the rectangle $[0, 2] \times [-2, 3]$ in the (u, v)-plane and consider the change of coordinates given by $\Phi(u, v) = (u^3, v + 2)$. Use this to rewrite the integral

$$\iint\limits_{\Phi^{-1}(D)} 12u^8 v \, dA.$$

(This is an integral where the change of coordinates is probably overkill as we could compute the integral *without* a change of coordinates, but this will illustrate how the change of variable is used and also emphasize how it's similar to u-substitution in one variable.)

Notice the Jacobian for this problem would be

$$\begin{aligned}\left|\frac{\partial(x, y)}{\partial(u, v)}\right| &= \det\begin{pmatrix} 3u^2 & 0 \\ 0 & 1 \end{pmatrix} \\ &= 3u^2\end{aligned}$$

Conveniently our integral has a factor of $3u^2$ present, so we may write

$$\iint\limits_{D} 12u^8 v \, dA = \iint\limits_{D} 4u^6 v \cdot 3u^2 \, dA.$$

Now notice that $x = u^3$ and $y = v+2$, so the expression $4u^6 v$ in (x, y)-coordinates is $4x^2(y - 2)$. Thus the change of coordinates formula tells us our integral can be written as

$$\iint\limits_{D} 4x^2(y - 2) \, dA.$$

We now need to compute $\Phi(D)$. To do this, we simply apply Φ to each point of $D = [0, 2] \times [-2, 3]$. This gives us the rectangle $D =$

$[0, 8] \times [0, 5]$ and our integral becomes

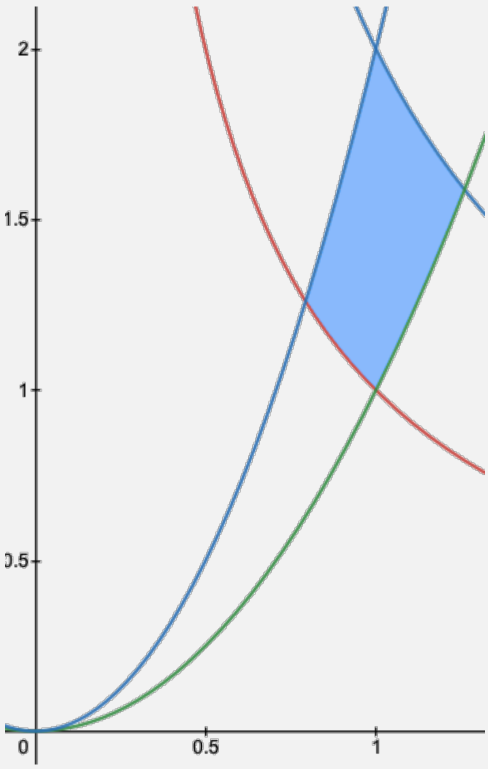
$$\begin{aligned}
\int_0^5 \int_0^8 4x^2(y-2)\,dx\,dy &= \int_0^5 \frac{4x^3}{3}(y-2)\bigg|_0^8 dy \\
&= \int_0^5 \frac{2048}{3}(y-2)\,dy \\
&= \frac{2048}{3}\int_0^5 (y-2)\,dy \\
&= \frac{2048}{3}\left(\frac{y^2}{2}-2y\right)\bigg|_0^5 \\
&= \frac{2048}{3}\left(\frac{25}{2}-10\right) \\
&= \frac{2048}{3}\cdot\frac{5}{2} \\
&= \frac{10240}{6} \\
&= \frac{5120}{3} \approx 1706.6667
\end{aligned}$$

Just to verify this is the true value of the integral we started with, we can compute

$$\begin{aligned}
\iint_D 12u^8 v\,dA &= \int_{-2}^3 \int_0^2 12u^8 v\,du\,dv \\
&= \int_{-2}^3 12v\,dv \int_0^2 u^8\,du \\
&= 6v^2\bigg|_{-2}^3 \frac{u^9}{9}\bigg|_0^2 \\
&= (54-24)\cdot\frac{2^9}{9} \\
&= \frac{30\cdot 2^9}{9} \\
&= \frac{15360}{9} \\
&= \frac{5120}{3} \approx 1706.6667
\end{aligned}$$

The example above was just to illustrate the idea of using a change of variables, but of course the main use is to take difficult problems and make them easier. One way this occurs is if we have a complicated domain, we may be able to change coordinates to get a simpler one.

Example 4.19.
Find the area of the region in the satisfying $1 \leq xy \leq 2$ and $x^2 \leq y \leq 2x^2$.



We can compute this area without using any change of variables, but it is exceedingly annoying. What we will do instead is find an appropriate Φ which will map our region to something which is easier to work with. The simplest possible thing we could hope for would be a rectangle. In particular, if we were to take $u = xy$, notice the inequalities $1 \leq xy \leq 2$ would become $1 \leq u \leq 2$. For the second pair of inequalities, dividing by x^2 would yield $1 \leq \frac{y}{x^2} \leq 2$, so let's take $v = \frac{y}{x^2}$ so we will have $1 \leq v \leq 2$.

In order to apply our change of variables formula as written above, though, we want Φ to convert (u, v) into (x, y). That is, we wish to solve the system

$$\begin{aligned} u &= xy \\ v &= \frac{y}{x^2} \end{aligned}$$

for x and y.

The second equation will let us write $y = x^2 v$, and then the first equation becomes $u = x^3 v$ from which we obtain $x = (u/v)^{1/3}$. Plugging this into $y = x^2 v$ gives us

$$y = \left(\frac{u}{v}\right)^{2/3} v = \sqrt[3]{\frac{u^2}{v^2}}\sqrt[3]{v^3} = \sqrt[3]{\frac{u^2 v^3}{v^2}} = \sqrt[3]{u^2 v}.$$

That is, our Φ is

$$\Phi(u, v) = (u^{1/3} v^{-1/3}, u^{2/3} v^{1/3}).$$

The Jacobian of Φ is then

$$\begin{aligned} &\det \begin{pmatrix} \frac{1}{3} u^{-2/3} v^{-1/3} & \frac{-1}{3} u^{1/3} v^{-4/3} \\ \frac{2}{3} u^{-1/3} v^{1/3} & \frac{1}{3} u^{2/3} v^{-2/3} \end{pmatrix} \\ =&\frac{1}{9}(u^{-2/3} v^{-1/3} u^{2/3} v^{-2/3} + 2u^{-1/3} v^{1/3} u^{1/3} v^{-4/3}) \\ =&\frac{1}{9}\left(\frac{1}{v} + \frac{2}{v}\right) \\ =&\frac{1}{3v} \end{aligned}$$

If D is our original region in the xy-plane, and $\Phi^{-1}(D)$ is its inverse image in the uv-plane, the change of coordinates principle tells us

$$\iint_D 1\, dA = \iint_{\Phi^{-1}(D)} 1 \cdot \frac{1}{3v}\, dA$$

As we had noted above, though, $\Phi^{-1}(D) = [1,2] \times [1,2]$ and the integral becomes

$$\begin{aligned}\int_1^2 \int_1^2 \frac{1}{3v}\, du\, dv &= \frac{1}{3}\int_1^2 du \cdot \int_1^2 \frac{1}{v}\, dv \\ &= \frac{1}{3} \cdot \ln|v| \Big|_1^2 \\ &= \frac{1}{3}(\ln(2) - \ln(1)) \\ &= \frac{\ln(2)}{3}\end{aligned}$$

Above we have only considered changing coordinates for integrals in two variables, but the change of coordinates formula works just as well in three variables: If $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ is a differentiable transformation and $E \subseteq \mathbb{R}^3$ is a region in 3-space where $f(\Phi(u,v,w))$ is defined and integrable, then

$$\iiint\limits_{\Phi(E)} f(x,y,z)\, dV = \iiint\limits_{E} f(\Phi(u,v,w)) \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| dV$$

where $\left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right|$ denotes the Jacobian of Φ, i.e., the determinant of the total derivative $D\Phi$. For functions of three variables, note the Jacobian matrix is

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}$$

The determinant of a 3×3 matrix was described in Section 1.4, and applying the formula there here, we would see

$$\begin{aligned}&\left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| \\ =&\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}\frac{\partial z}{\partial w} + \frac{\partial x}{\partial v}\frac{\partial y}{\partial w}\frac{\partial z}{\partial u} + \frac{\partial x}{\partial w}\frac{\partial y}{\partial u}\frac{\partial z}{\partial v} - \\ &\frac{\partial x}{\partial w}\frac{\partial y}{\partial v}\frac{\partial z}{\partial u} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\frac{\partial z}{\partial w} - \frac{\partial x}{\partial u}\frac{\partial y}{\partial w}\frac{\partial z}{\partial v}\end{aligned}$$

As one example of computing such a Jacobian, consider the following change of coordinates:

$$\Phi(u,v,w) = (u\sin(w)\cos(v), u\sin(w)\sin(v), u\cos(w)).$$

The Jacobian matrix would be

$$\begin{pmatrix} \sin(w)\cos(v) & -u\sin(w)\sin(v) & u\cos(w)\cos(v) \\ \sin(w)\sin(v) & u\sin(w)\cos(v) & u\cos(w)\sin(v) \\ \cos(w) & 0 & -u\sin(w) \end{pmatrix}$$

The determinant is tedious to calculate, but the end result would be

$$\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| = u^2\sin(w).$$

As mentioned many times, the choice of variable names is largely irrelevant. If we replaced u by ρ, v by θ and w by φ, then our transformation would become

$$\Phi(\rho,\theta,\varphi) = (\rho\sin(\varphi)\cos(\theta), \rho\sin(\varphi)\sin(\theta), \rho\cos(\varphi)).$$

Notice the Jacobian determinant is then $\rho^2\sin(\varphi)$. That is, this is simply our spherical coordinates conversion and the $\rho^2\sin(\varphi)$ that appears in spherical integrals is exactly the Jacobian from the change of coordinates equation.

5 Vector Calculus

Wahrlich es ist nicht das Wissen, sondern das Lernen, nicht das Besitzen sondern das Erwerben, nicht das Da-Seyn, sondern das Hinkommen, was den grössten Genuss gewährt.
Truly, it is not knowledge, but learning, not possessing, but acquiring, not being-there, but coming, which gives the greatest enjoyment.

CARL FRIEDRICH GAUSS

5.1 Vector fields

Suppose that D is a subset of the plane. A ***vector field*** on D is a two-dimensional vector-valued function $\vec{F}$ defined on D. That is, for each point (x, y) in D we associate a two-dimensional vector $\vec{F}(x, y)$. Notice this vector may change as the point changes. We visualize these vector fields as a collection of vectors (arrows) in the plane with the tail of $\vec{F}(x, y)$ being placed at the point (x, y) as in Figure 5.1.

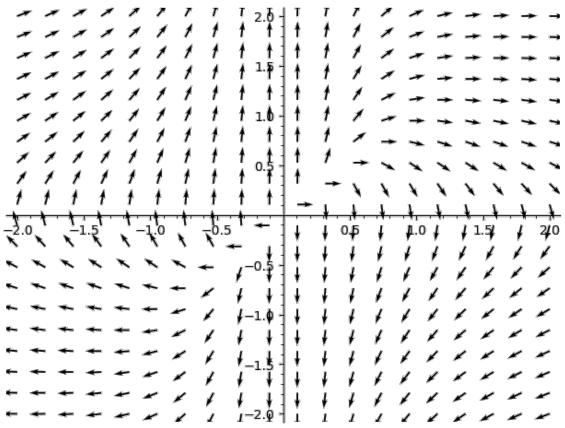


Figure 5.1: The vector field $\vec{F}(x, y) = \frac{1}{\sqrt{(x^2y)^2+(y-x)^2}} \langle x^2y, y - x \rangle$.

Since each vector $\vec{F}(x,y)$ has two components, the vector field is specified by two functions: one tells us the x-components of the vector, and one tells us the y-components. If $P(x,y)$ tells us the x-component of the vector and $Q(x,y)$ tells us the y-component, then we have

$$\vec{F}(x,y) = \langle P(x,y), Q(x,y) \rangle .$$

For example, consider the vector field $\vec{F}(x,y) = \langle -y, x \rangle$. Here the component functions are $P(x,y) = -y$ and $Q(x,y) = x$. Notice that when we plug a given point (x,y) into this vector field, we get back the vector $\langle -y, x \rangle$. This vector field is picture in Figure 5.2 below.

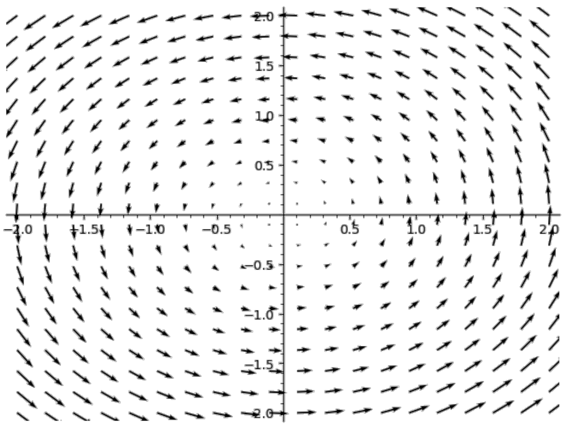


Figure 5.2: The vector field $\vec{F}(x,y) = \langle -y, x \rangle$.

We can similarly define a three-dimensional vector field as follows. Given a subset E of 3-space, for each point $(x,y,z) \in E$ we associate a three-dimensional vector $\vec{F}(x,y,z)$. This vector is determined by three functions, one telling us each of the components of the vectors.

$$\vec{F}(x,y,z) = \langle P(x,y,z), Q(x,y,z), R(x,y,z) \rangle .$$

Vector fields appear naturally in many parts of mathematics and physics. For example, the motion of particles in the plane can be specified with a vector field where the vector at each point represents the velocity of a particle when it is at this particular point. Imagine, for example, a small float placed into a river. The current of the river pushes the float around giving it some velocity at each point. If we could know what the velocity of the float was at every single point in the river, that would give us a vector field.

As another example, Newton's law of universal gravitation says that any two particles (electrons, protons, neutrons, etc.) in the universe are attracted to one another according to an inverse square law. The total force obtained by adding together all forces attracting a particle to all of the other particles in the universe gives us a vector field: one vector corresponding to gravitational force for every point in the universe.

Vector fields that arise in this way in physics are often given names describing what the vectors represent. If the vectors represent velocities, then the vector field is called a ***velocity field***. If the vectors represent forces, the vector field is a ***force field***. If those forces are the gravitational forces, the field is called a ***gravitational field***. If the forces come from the electromagnetic force of charged particles, the field is an ***electromagnetic field***, and so on.

If a vector field represents the velocity of a particle at a point (x, y) in the plane (or (x, y, z) for a three-dimensional vector field), then the curve traced out by a particle as it moves with the prescribed velocities is called a ***flow line*** or ***integral curve***. In the case of our vector field $\vec{F}(x, y) = \langle -y, x\rangle$ above, the integral curves are circles.

To be more precise a flow line / integral curve of a vector field is a parametric curve whose tangent vectors agree with the vectors given by the vector field. That is, if $\vec{F}(x, y) = \langle P(x, y), Q(x, y)\rangle$ is a vector field in the plane, an integral curve of $\vec{F}$ is a parametric curve $\vec{\gamma}(t) = \langle x(t), y(t)\rangle$ with the property that $\vec{\gamma}'(t) = \vec{F}(\gamma(t))$ for every t. Writing this out in components, this simply means that for every t we have

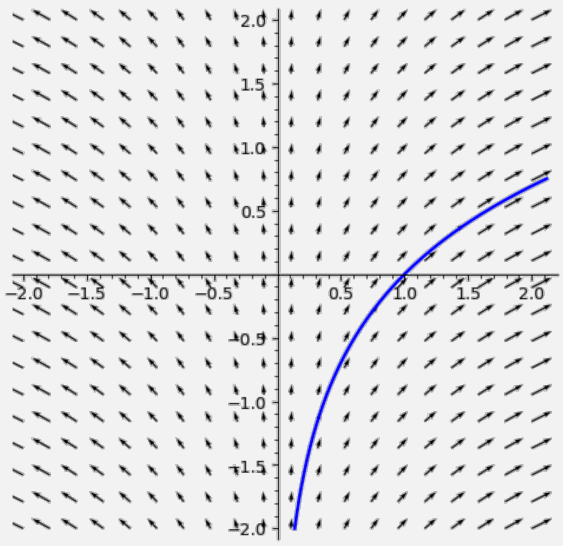
$$x'(t) = P(x(t), y(t))$$
$$y'(t) = Q(x(t), y(t))$$

Determining a flow line to a given vector field essentially involves solving a system of differential equations. As differential equations is not a prerequisite for this course, we will not be concerned with determining the integral curves to a vector field in this course. *However*, we can easily verify if a given parametric curve is an integral curve of a vector field or not by simply computing its derivative and seeing if it agrees with what is given by the vector field.

Example 5.1.
Is the curve $\vec{\gamma}(t) = \langle e^t, t\rangle$ an integral curve for the vector field $\vec{F}(x, y) = \langle x, 1\rangle$?

We need to verify that for each point on our curve the tangent vector $\vec{\gamma}'(t)$ equals the vector given by $\vec{F}$ at that point. Notice the tangent vectors are $\vec{\gamma}'(t) = \langle e^t, 1\rangle$. Given a value $t = t_0$, the corresponding point on the curve is (e^{t_0}, t_0), and so the tangent vector there is $\langle e^{t_0}, 1\rangle$. The corresponding vector from the vector field $\vec{F}$ is $\vec{F}(e^{t_0}, t_0) = \langle e^{t_0}, 1\rangle$. Since these vectors agree for our arbitrarily chosen point on the curve, they agree for every point on the curve, and so $\vec{\gamma}(t)$ is an integral curve of $\vec{F}$.

The vector field and integral curve are plotted below. Notice that the curve "follows" the vector field.



One particular important type of vector field comes from taking the gradient of a function. Recall that given a function $f(x, y)$, its ***gradient*** is defined as

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle .$$

This gives us a vector field because for every point (x_0, y_0) in the plane we have an associated vector, namely $\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle$.

For example, if $f(x, y) = 3ye^x$, then our gradient vector field is

$$\nabla f(x, y) = \langle 3ye^x, 3e^x\rangle .$$

If a given vector field $\vec{F}$ is the gradient of a function (i.e., if there exists a function f so that $\vec{F} = \nabla f$), then the vector field $\vec{F}$ is called ***conservative*** and the function f whose gradient is f is called the ***potential function*** of

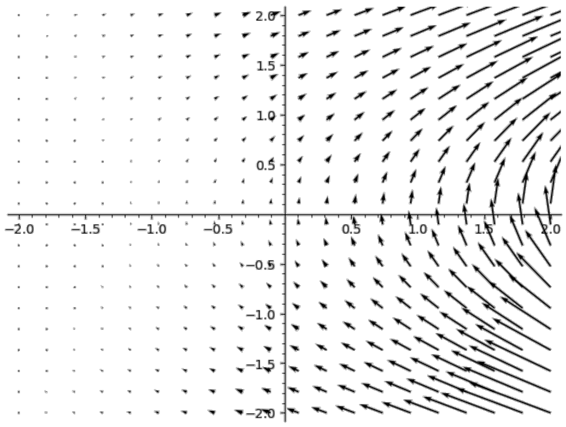


Figure 5.3: The vector field $\langle 3ye^x, 3e^x\rangle$ is conservative with potential function $3ye^x$.

the vector field. For example, the vector field $\langle 3ye^x, 3e^x\rangle$ above is conservative with potential function $3ye^x$.

One particularly important example of a conservative three-dimensional vector field is the gravitational vector field. Newton's law of gravitation says that the gravitational force between two objects of masses m and M which are distance r apart has magnitude ${}^{mMG}/_{r^2}$ where G is a constant.

If we assume the larger of these two objects, with mass M, is placed at the origin in 3-space and the smaller object of mass m is placed at position (x, y, z), notice the displacement vector from the smaller object to the larger object is

$$-\vec{x} = \langle -x, -y, -z\rangle$$

and the distance between the objects is $\|\vec{x}\|$. Thus by Newton the gravitational force at this point is

$$\frac{mMG}{\|\vec{x}\|^2} \cdot \frac{-\vec{x}}{\|\vec{x}\|}.$$

Writing this out in components gives us the following vector field:

$$\vec{F}(x, y, z) = \left\langle \frac{-mMGx}{(x^2+y^2+z^2)^{3/2}}, \frac{-mMGy}{(x^2+y^2+z^2)^{3/2}}, \frac{-mMGz}{(x^2+y^2+z^2)^{3/2}}, \right\rangle$$

We can easily check that the following function is the potential function of this vector field by computing its gradient:

$$f(x, y, z) = \frac{mMG}{\sqrt{x^2+y^2+z^2}}.$$

If you've taken some physics, you may recognize the function above as the general formula for the gravitational potential energy between two masses in 3-space. You may also recall that gravity is a conservative force, meaning that work done against the force from one point to another does not depend on the path between two points. Further more, in this situation the work done can be computed as the difference of the potential function evaluated at the two points. As we will see, a similar statement will hold for abstract conservative vector fields and their potential functions.

One particularly important property of conservative vector fields is that they are always orthogonal to the level curves of their potential function. For example, the vector field $\vec{F}(x, y) = \langle x, 1\rangle$ is conservative with potential function $f(x, y) = \frac{x^2}{2} + y$, and the vector field with some level curves is given in Figure 5.4.

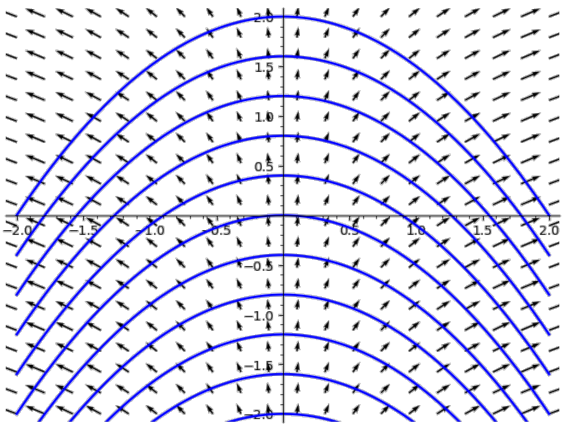


Figure 5.4: The conservative vector field $\langle x, 1\rangle$ together with some of the level curves of its potential function, $x^2/2 + y$.

In general, this means that given any point (x, y), the vector coming from the conservative vector field, $\vec{F}(x, y)$, is orthogonal to the tangent vector of the integral curve through (x, y) at the point (x, y). In the case of $\vec{F}(x, y) = \langle x, 1\rangle$ with potential $f(x, y) = x^2/2 + y$, we can verify this directly. Notice the level curves of the potential are given by solutions to $x^2/2 + y = c$ for constants c. Solving this for y gives us the parabola $y = c - x^2/2$ which we may parametrize as $\vec{\gamma}(t) = \langle t, c - t^2/2\rangle$. Notice the tangent vectors to this curve are $\vec{\gamma}\,'(t) = \langle 1, -t\rangle$. The t value here is precisely the x-value at our point, so we may write these tangent vectors as $\langle 1, -x\rangle$. The corresponding vector from our vector field is $\langle x, 1\rangle$, and the dot product between these vectors is

$$\langle 1, -x \rangle \cdot \langle x, 1 \rangle = x - x = 0,$$

and so the vectors are orthogonal.

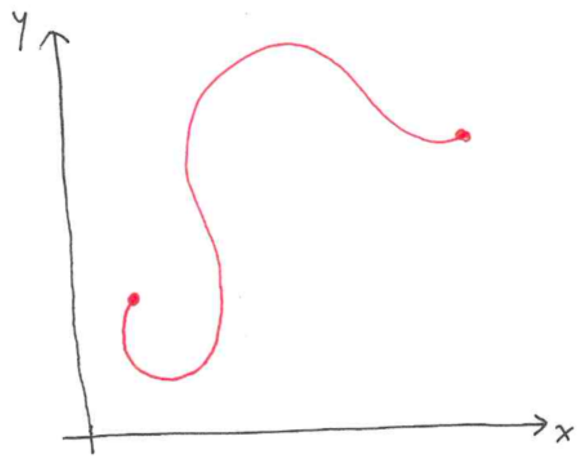
5.2 Line integrals

Intuition

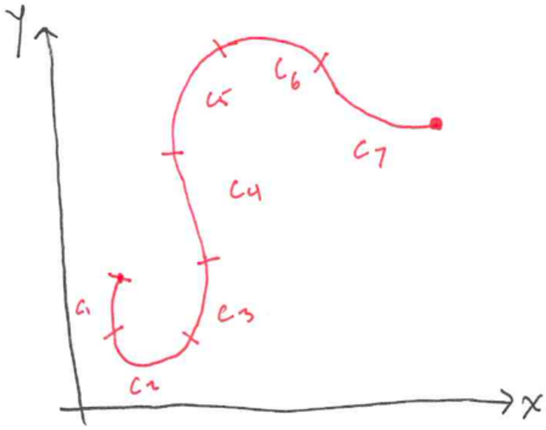
The Riemann integral of a function $f(x, y)$ over an interval $[a, b]$ is defined as a limit of Riemann sums where we cut the interval $[a, b]$ up into n pieces (let's say the pieces are of equal size, Δx) and pick a point x_i^* in each piece, add up the products $f(x_i^*)\Delta x$, and then take a limit,

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

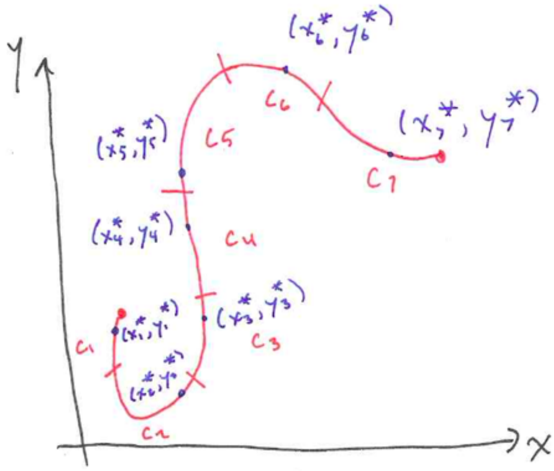
Suppose we wanted to do something similar, but instead of integrating over a subinterval of the real line, we decide to integrate over a curve. That is, suppose C is some curve in the plane,



And suppose that $f(x, y)$ is a function whose domain includes the curve C. Let's cut C up into a bunch of pieces C_1, ..., C_n.



Pick a point (x_i^*, y_i^*) inside of the i-th piece of the curve.



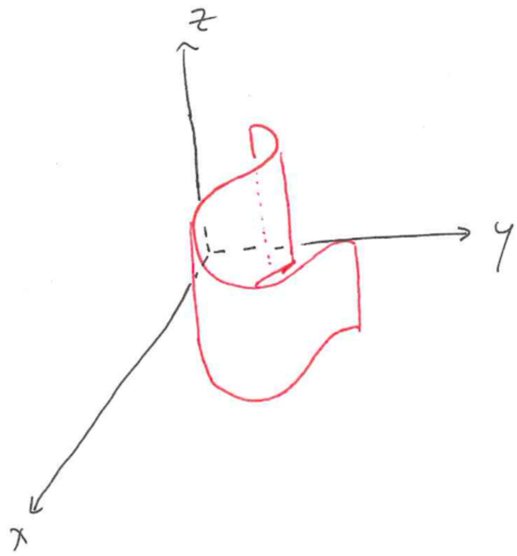
Now evaluate the function at these points, $f(x_i^*, y_i^*)$, multiplying by the arclength of the C_i piece of the curve – call this value Δs_i – and add up all of these terms:

$$\sum_{i=1}^{n} f(x_i^*, y_i^*)\Delta s_i.$$

Taking the limit as n goes to infinity (really, as each C_i piece gets arbitrarily short), we have the *line integral* of f over C:

$$\int_C f(x,y)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*)\Delta s_i.$$

Just to give this idea a geometric interpretation, suppose that $f(x, y) \geq 0$ for all points (x, y) on the curve C. We could then define curve in 3-space by "lifting" the curve C to the surface $z = f(x, y)$. If C is parametrized by $\langle x(t), y(t)\rangle$, then this curve in 3-space would be parametrized by $\langle x(t), y(t), f(x(t), y(t))\rangle$. If we filled in all the space between this curve on the surface $z = f(x, y)$ and the curve in the plane we'd have a surface. (Imagine hanging a sheet from the surface along the curve and letting that sheet hang down to the xy-plane.) The line integral would represent the area of this surface.



We will see that line integrals have many applications and interpretations, so don't get too hung up on this particular interpretation: it's just a quick way to help you see one possible application of line integrals.

Integration with respect to arclength

As is always the case with integration, the "limit of Riemann sums" definition makes some intuitive sense, but it is usually very difficult to make calculations with this definition. To make life a little simpler, we'd like to somehow relate these line integrals back to integrals we already know how to calculate. To get started with this, let's suppose the curve C we are integrating over has a parametrization

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

for $a \leq t \leq b$. Then when we plug in points (x, y) on the curve C into $f(x, y)$, we can just use $f(x(t), y(t))$. Now recall that the length of the curve C is given by

$$\ell(C) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

We can similarly define a function $s(t)$ which gives the arclength of the curve from its initial point (the point $\vec{r}(a) = \langle x(a), y(a)\rangle$) to some other point $\vec{r}(t)$. This is given by

$$s(t) = \int_a^{a+t} \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

Notice by the fundamental theorem of calculus,

$$s'(t) = \sqrt{x'(t)^2 + y'(t)^2},$$

and so we may write the differential ds as

$$\begin{aligned} ds &= s'(t)\, dt \\ &= \sqrt{x'(t)^2 + y'(t)^2}\, dt \end{aligned}$$

We can now use this to simplify our definition of the line integral:

$$\int_C f(x,y) ds = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2}\, dt.$$

As an example, let's calculate the value

$$\int_C y^3\, ds$$

where C is the curve parametrized by $\vec{r}(t) = \langle t^3, t\rangle$ for $0 \leq t \leq 2$.

$$\begin{aligned} \int_C y^3\, ds &= \int_0^2 y(t)^3 \sqrt{x'(t)^2 + y'(t)^2}\, dt \\ &= \int_0^2 t^3 \sqrt{(3t^2)^2 + 1}\, dt \\ &= \int_0^2 t^3 \left(9t^4 + 1\right)^{1/2}\, dt \end{aligned}$$

Now performing the substitution

$$\begin{aligned} u &= 9t^4 + 1 \\ du &= 36t^3\, dt \end{aligned}$$

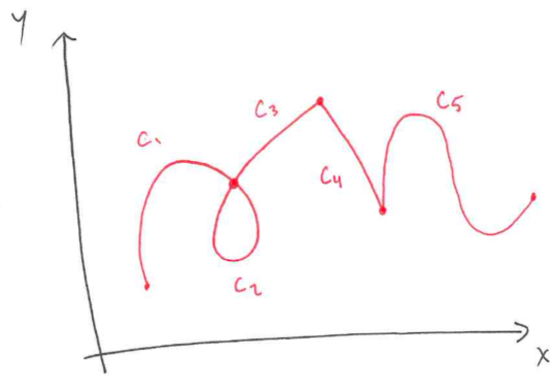
The integral becomes

$$\begin{aligned}\int_C y^3 ds &= \int_1^{145} \frac{\sqrt{u}du}{36} \\ &= \frac{2u^{3/2}}{108}\bigg|_1^{145} \\ &= \frac{1}{54}\left(145^{3/2} - 1\right).\end{aligned}$$

Writing a line integral as

$$\int_C f(x,y)\,ds = \int_a^b f(x(t),y(t))\sqrt{x'(t)^2 + y'(t)^2}\,dt$$

only works if $x'(t)$ and $y'(t)$ are defined, but we may want to integrate over curves where these values aren't defined. (For example, consider a curve with a sharp corner.) Recall that curves where the tangent vectors $\vec{r}\,'(t)$ is defined everywhere, and is never zero, are called *smooth*. The above substitution for our line integral only applies if we integrate over a smooth curve. We can integrate over curves that are not smooth as well, provided we we can break the curve up into smooth pieces. Such a curve is called *piecewise smooth*.



If we could break the curve into, say, n smooth pieces – say C_1 through C_n – then to integrate a function $f(x,y)$ over the curve, we just integrate over each piece and add these integrals together.

$$\int_C f(x,y)\,ds = \sum_{i=1}^{n} \int_{C_i} f(x,y)\,ds.$$

Example 5.2.
Suppose that C is a curve which breaks into two pieces, C_1 which is parametrized by $\langle t, t^2 \rangle$ for $0 \le t \le 1$; and C_2 which is parametrized by $\langle 1, t \rangle$ for $1 \le t \le 2$. And suppose that we want to integrate $f(x, y) = 2x$ over this curve.

$$\begin{aligned}
\int_C 2x\,ds &= \int_{C_1} 2x\,ds + \int_{C_2} 2x\,ds \\
&= \int_0^1 2t\sqrt{1+4t^2}\,dt + \int_1^2 2\sqrt{0^2+1^2}\,dt \\
&= \int_0^1 2t(1+4t^2)^{1/2}\,dt + 2
\end{aligned}$$

Using the substitution $u = 1 + 4t^2$, $du = 8t\,dt$, this becomes

$$\begin{aligned}
\int_C 2x\,ds &= \int_1^5 \frac{1}{4}u^{1/2}\,du + 2 \\
&= \frac{1}{4}\cdot\frac{2}{3}u^{3/2}\Big|_1^5 + 2 \\
&= \frac{1}{6}\left(5^{3/2} - 1\right) + 2 \\
&= \frac{5\sqrt{5}+11}{6}.
\end{aligned}$$

To give all of this a physical interpretation, suppose that C gives the shape of a wire and $\rho(x, y)$ is the density of the write at a point (x, y) on C, then $\int_C \rho(x, y)\,ds$ is the mass of the wire.

Integration with respect to the other variables

In our original definition of the line integral, we defined

$$\int_C f(x,y)\,ds = \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} f(x_i^*, y_i^*)\,\Delta s_i$$

where Δs_i denotes the arclength of a small piece of the curve. There are times, however, where we will instead want to replace the Δs_i's with something else, such as Δx_i or Δy_i. (This will become important when we describe line integrals of vector fields later.)

We define the line integral of $f(x,y)$ over the curve C with respect to x as the value

$$\int_C f(x,y)\,dx = \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} f(x_i^*, y_i^*)\Delta x_i.$$

Assuming C is parametrized by $\vec{r}(t) = \langle x(t), y(t)\rangle$ with $a \leq t \leq b$, this integral may be written as

$$\int_C f(x,y)\,dx = \int_a^b f(x(t),y(t))x'(t)\,dt.$$

Of course, we can also do this integration with respect to y:

$$\int_C f(x,y)\,dy = \int_a^b f(x(t),y(t))y'(t)\,dt.$$

These sorts of integrals will appear together often, and so we'll adopt the following notationalu short-hand:

$$\int_C P(x,y)\,dx + \int_C Q(x,y)\,dy = \int_C P(x,y)\,dx + Q(x,y)\,dy.$$

As an example, let's integrate

$$\int_C \sin(\pi y)\,dy + yx^2\,dx$$

where C is given by the parametrization $\langle 1-t, 4-2t\rangle$ with $0 \leq t \leq 1$.

$$\begin{aligned}
\int_C \sin(\pi y)\,dy + yx^2\,dx &= \int_C \sin(\pi y)\,dy + \int_C yx^2\,dx \\
&= \int_0^1 \sin\left(\pi(4-2t)\right)(-2)dt + \int_0^1 (4-2t)(1-t)^2(-1)dt \\
&= (-2)\int_0^1 \sin\left(4\pi - 2\pi t\right)\,dt - \int_0^1 \left(-2t^3 + 8t^2 - 10t + 4\right)\,dt \\
&= -\frac{1}{\pi}\cos(4\pi - 2\pi t)\Big|_0^1 - \left(-\frac{t^4}{2} + \frac{8t^3}{3} - 5t^2 + 4t\right)\Big|_0^1 \\
&= -\frac{1}{\pi}(1-1) - \left(-\frac{1}{2} + \frac{8}{3} - 5 + 4\right) \\
&= 1 - \frac{8}{3} + \frac{1}{2} \\
&= \frac{6 - 16 + 3}{6} \\
&= \frac{-7}{6}
\end{aligned}$$

Just as we have line integrals of functions of two variables over curves in two-dimensional space, we have line integrals of functions of three variables.

Our definition is just like the two-dimensional case,

$$\int_C f(x,y,z)\,ds = \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} f(x_i^*, y_i^*, z_i^*)\,\Delta s_i$$

and if the curve C is parametrized by $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ over $a \le t \le b$, this becomes

$$\int_C f(x,y,z)\,ds = \int_a^b f(x(t), y(t), z(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\,dt.$$

Remark.
Note that the arclength of the curve C is given by integrating the constant function 1, with respect to arclength, over C:

$$\ell(C) = \int_C ds = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\,dt.$$

Line integrals with respect to x, y, or z are defined similarly to the two-dimensional case:

$$\begin{aligned}\int_C P(x,y,z)\,dx &= \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} P\left(x_i^*, y_i^*, z_i^*\right)\Delta x_i \\ &= \int_a^b P(x(t), y(t), z(t))x'(t)\,dt\end{aligned}$$

$$\begin{aligned}\int_C Q(x,y,z)\,dy &= \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} Q\left(x_i^*, y_i^*, z_i^*\right)\Delta y_i \\ &= \int_a^b Q(x(t), y(t), z(t))y'(t)\,dt\end{aligned}$$

$$\int_C R(x,y,z)\,dz = \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{\#\mathcal{P}} R\left(x_i^*, y_i^*, z_i^*\right)\Delta z_i$$
$$= \int_a^b R(x(t),y(t),z(t))z'(t)\,dt$$

Example 5.3.
Calculate the line integral

$$\int_C xyz\,ds$$

where C is parametrized by $\langle 2\sin(t), t, -2\cos(t)\rangle$ with $0 \le t \le \pi$.

$$\begin{aligned}\int_C xyz\,ds &= \int_0^\pi 2\sin(t)\cdot t\cdot(-2\cos(t))\sqrt{4\cos^2(t)+1+4\sin^2(t)}\,dt \\ &= \int_0^\pi (-4)t\sin(t)\cos(t)\sqrt{4+1}\,dt \\ &= -4\sqrt{5}\int_0^\pi t\sin(t)\cos(t)\,dt.\end{aligned}$$

Now we can use the trig identity

$$\begin{aligned}\sin(2\theta) &= 2\sin\theta\cos\theta \\ \implies \sin\theta\cos\theta &= \frac{\sin(2\theta)}{2}\end{aligned}$$

Now we rewrite this integral as

$$\int_C xyz\,ds = -2\sqrt{5}\int_0^\pi t\sin(2t)\,dt.$$

Now we do integration by parts with

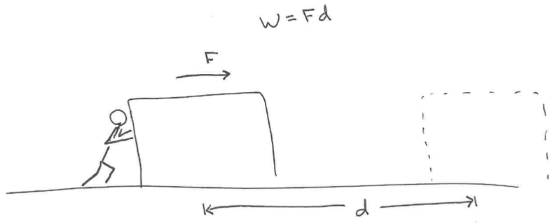
$$\begin{aligned} u &= t & dv &= \sin(2t)\,dt \\ du &= dt & v &= -\frac{\cos(2t)}{2}\end{aligned}$$

to get

$$\begin{aligned}\int_C xyz\,ds &= \sqrt{5}t\cos(2t)\Big|_0^\pi - \sqrt{5}\int_0^\pi \cos(2t)\,dt \\ &= \sqrt{5}\pi\cos(2\pi) - \sqrt{5}\int_0^\pi \cos(2t)\,dt \\ &= \sqrt{5}\left(\pi - \frac{\sin(2t)}{2}\Big|_0^\pi\right) \\ &= \sqrt{5}\left(\pi - \frac{\sin(2\pi)}{2} + \frac{\sin(0)}{2}\right) \\ &= \pi\sqrt{5}.\end{aligned}$$

Work

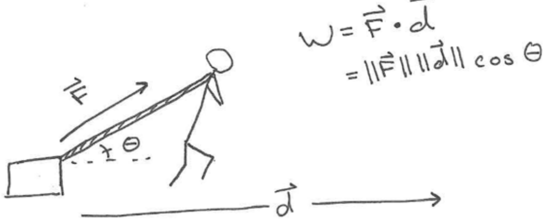
Recall that if a force of magnitude F pushes an object a distance d, then the work done by the force is $W = Fd$.



This calculation makes two key assumptions: the force is constant, and in the direction of motion. We know that if the force is not in the direction of motion we can use dot products to calculate the work. If our force is given a vector $\vec{F}$ and the displacement of the object is given by $\vec{d}$, this works out to be

$$W = \vec{F}\cdot\vec{d} = \|\vec{F}\|\,\|\vec{d}\|\cos\theta$$

where θ is the angle between $\vec{F}$ and $\vec{d}$.



Here we are still supposing that the force being applied is constant, but now we'd like to consider the case of a force which changes.

Suppose the force being applied at a point (x, y, z) is

$$\vec{F}(x, y, z) = \langle P(x, y, z),\ Q(x, y, z),\ R(x, y, z) \rangle .$$

Suppose also that we've moving an object along a curve C in $\mathbb{R}^3$. Then to approximate the work done, we'll cut the curve into several pieces: say C_1, C_2, ..., C_n. We'll suppose that the force is constant on each C_i piece of the curve, and let's denote the arclength of the curve by Δs_i. To estimate the force, we'll pick some point $(x_i^*, y_i^*, z_i^*) \in C_i$ and suppose the force along the curve C_i is given by $\vec{F}(x_i^*, y_i^*, z_i^*)$.

So we have a force, $\vec{F}(x_i^*, y_i^*, z_i^*)$, we also have a length Δs_i, but we don't have a direction. To get a direction we'll use the tangent vector of the curve at (x_i^*, y_i^*, z_i^*). Recall that the unit tangent vector is denoted $\vec{T}(x_i^*, y_i^*, z_i^*)$. However, this a unit vector, and we're assuming the force is constant on all of C_i which has length Δs_i. Thus we'll take our displacement vector to be $\Delta s_i \vec{T}(x_i^*, y_i^*, z_i^*)$. Hence our approximation for the work done over C_i is simply

$$\vec{F}(x_i^*, y_i^*, z_i^*) \cdot \left(\Delta s_i \vec{T}(x_i^*, y_i^*, z_i^*\right) .$$

Since Δs_i is just a scalar, we can move it around to rewrite this as

$$\vec{F}(x_i^*, y_i^*, z_i^*) \cdot \vec{T}(x_i^*, y_i^*, z_i^* \Delta s_i.$$

Summing up these approximations for each piece, we approximate the total work to be

$$W \approx \sum_{i=1}^{n} \vec{F}(x_i^*, y_i^*, z_i^*) \cdot \vec{T}(x_i^*, y_i^*, z_i^*) \Delta s_i.$$

Of course, we want to take the limit as our pieces become arbitrarily small to get the "best" approximation:

$$W = \lim_{\|\mathcal{P}\| \to 0} \sum_{i=1}^{\#\mathcal{P}} \vec{F}(x_i^*, y_i^*, z_i^*) \cdot \vec{T}(x_i^*, y_i^*, z_i^*) \Delta s_i.$$

Notice, however, that this is just the line integral over $\vec{F}(x,y,z)\cdot\vec{T}(x,y,z)$ over C:

$$W=\int_C \vec{F}(x,y,z)\cdot\vec{T}(x,y,z)\,ds,$$

which we'll usually abbreviate to be

$$W=\int_C \vec{F}\cdot\vec{T}\,ds.$$

Supposing our curve is parametrized by

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

for $a\le t\le b$, our unit tangent vectors are simply

$$\vec{T}(t)=\frac{\vec{r}\,'(t)}{\vec{r}\,'(t)}$$

and the ds is simply

$$ds=\|\vec{r}\,'(t)\|\,dt.$$

Writing $\vec{F}(\vec{r}(t))$ as short-hand for $\vec{F}(x(t),y(t),z(t))$, our integral becomes

$$\begin{aligned}W&=\int_a^b \vec{F}(\vec{r}(t))\cdot\frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|}\|\vec{r}\,'(t)\|\,dt\\&=\int_a^b \vec{F}(\vec{r}(t))\cdot\vec{r}\,'(t)\,dt\end{aligned}$$

and this is often abbreviated

$$\int_C \vec{F}\cdot d\vec{r}.$$

The idea of work is our motivation, but in can define the line integral of any continuous vector field $\vec{F}$ over a smooth curve C, parametrized by $\vec{r}(t)$ for $a\le t\le b$, as

$$\int_C \vec{F}\cdot\vec{T}\,ds=\int_C \vec{F}\cdot d\vec{r}=\int_a^b \vec{F}(\vec{r}(t))\cdot\vec{r}\,'(t)\,dt.$$

Example 5.4.
As an example, suppose that $\vec{F}$ is the vector field

$$\vec{F}(x,y,z) = \left\langle x+y, y-z, z^2 \right\rangle$$

and C is the curve parametrized by $\vec{r}(t) = \langle t^3, -t^2, t\rangle$ for $0 \le t \le 1$ and let's calculate $\int_C \vec{F} \cdot d\vec{r}$.

First note that $\vec{r}\,'(t) = \langle 3t^2, -2t, 1\rangle$, and $\vec{F}(\vec{r}(t)) = \langle t^3 - t^2, -t^2 - t, t^2\rangle$, and so our integral is

$$\begin{aligned}
\int_C \vec{F} \cdot d\vec{r} &= \int_0^1 \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\,dt \\
&= \int_0^1 \left\langle t^3 - t^2, -t^2 - t, t^2 \right\rangle \cdot \left\langle 3t^2, -2t, 1 \right\rangle dt \\
&= \int_0^1 \left(3t^2(t^3 - t^2) - 2t(-t^2 - t) + t^2\right)\,dt \\
&= \int_0^1 \left(3t^5 - 3t^4 + 2t^3 + 2t^2 + t^2\right)\,dt \\
&= \int_0^1 \left(3t^5 - 3t^4 + 2t^3 + 3t^2\right)\,dt \\
&= \left(\frac{3t^6}{6} - \frac{3t^5}{5} + \frac{2t^4}{4} + \frac{3t^3}{3}\right)\Big|_0^1 \\
&= \frac{1}{2} - \frac{3}{5} + \frac{1}{2} + 1 \\
&= 2 - \frac{3}{5} \\
&= \frac{7}{5}.
\end{aligned}$$

This seems like a lot of work to do, but we can actually simplify things a little bit. Suppose that

$$\vec{F}(x,y,z) = \langle P(x,y,z),\ Q(x,y,z),\ R(x,y,z)\rangle$$

and that C is parametrized by $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ with $a \le t \le b$. Then

$$\begin{aligned}
\int_C \vec{F} \cdot d\vec{r} &= \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt \\
&= \int_a^b \langle P(\vec{r}(t)), Q(\vec{r}(t)), R(\vec{r}(t)) \rangle \cdot \langle x'(t), y'(t), z'(t) \rangle \; dt \\
&= \int_a^b \left(P(\vec{r}(t))x'(t) + Q(\vec{r}(t))y'(t) + R(\vec{r}(t))z'(t) \right) \; dt \\
&= \int_a^b P(\vec{r}(t))x'(t)\, dt + \int_a^b Q(\vec{r}(t))y'(t)\, dt + \int_a^b R(\vec{r}(t))z'(t)\, dt \\
&= \int_C P(x,y,z)\, dx + \int_C Q(x,y,z)\, dy + \int_C R(x,y,z)\, dz \\
&= \int_C P(x,y,z)\, dx + Q(x,y,z)\, dy + R(x,y,z)\, dz
\end{aligned}$$

5.3 The fundamental theorem of line integrals

The fundamental theorem of calculus tells us that to evalute a definite integral, all we need to do is evaluate an antiderivative at the endpoints of the interval we're integrating over. In this lecture we'll describe the fundamental theorem of line integrals which gives us an analog for line integrals, provided we're integrating a conservative vector field.

Introduction

Recall from single variable calculus that if $F(x)$ is the antiderivative of $f(x)$, then

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

This is the fundamental theorem of calculus. We'd like to have something similar to this for line integrals.

The first thing we need is something to play the role of an antiderivative, and for this we turn to conservative vector fields. Recall that a vector field $\vec{F}$ is called *conservative* if there exists a function f such that $\vec{F} = \nabla f$. We then say f is a *potential function* for $\vec{F}$. Now suppose that C is an oriented curve which starts at P and ends at Q.

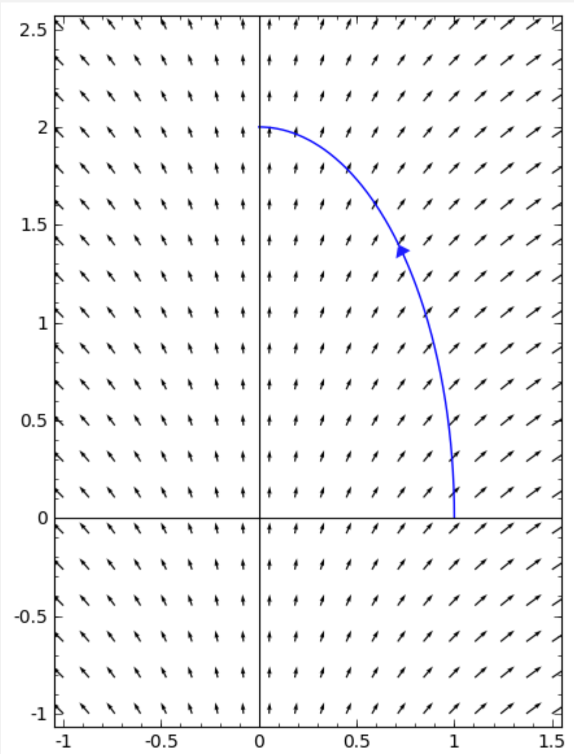
Theorem 5.1 (The Fundamental Theorem of Line Integrals).
Suppose that $\vec{F}$ is a conservative vector field with potential function f, and C is a curve which starts at the point P and ends at the point Q. Then

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \nabla f \cdot d\vec{r} = f(Q) - f(P).$$

Example 5.5.
We had seen previously that the vector field

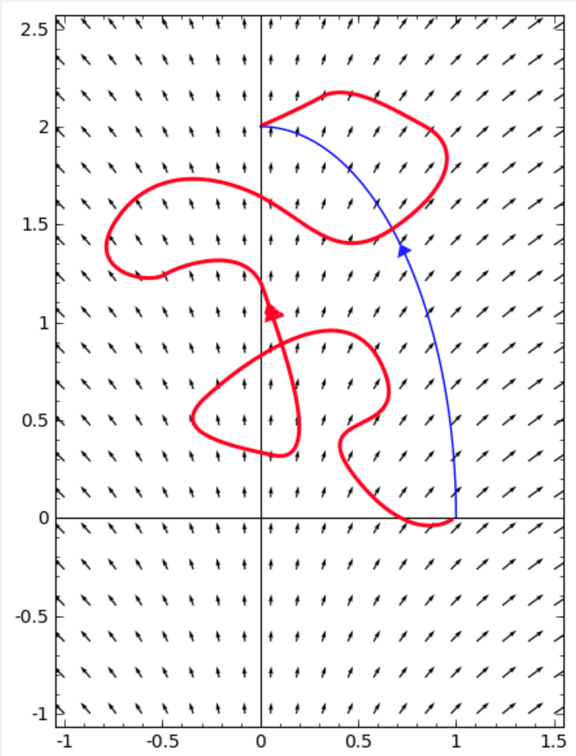
$$\vec{F}(x, y) = \langle x, 1 \rangle$$

is conservative with potential function $f(x, y) = \frac{x^2}{2} + y$. Now we have a really simple way of evaluating integrals of this vector field. Suppose C is a curve parametrized by $\vec{r}(t) = \langle \cos(t), 2\sin(t) \rangle$, for $t \in [0, \pi/2]$, as shown below.



$$\begin{aligned}\int_C \vec{F}\cdot d\vec{r} &= f\left(\cos(\pi/2), 2\sin(\pi/2)\right) - f(cos(0), 2\sin(0))\\ &= 2 - 1/2\\ &= 3/2\end{aligned}$$

Notice if we had *any other* path C from $(1,0)$ to $(0,2)$, we have the same value for the integral of $\vec{F}$ over that curve. Consider, for example, the red curve in the image below.



$$\int_{C_1} \vec{F}\cdot d\vec{r} = \int_{C_2} \vec{F}\cdot d\vec{r} = \int_{C_3} \vec{F}\cdot d\vec{r} = 3/2$$

In general, we say that a line integral $\int_C \vec{F}\cdot d\vec{r}$ is *independent of path* if for any two curves C_1 and C_2 with the same initial and terminal points $\int_{C_1} \vec{F}\cdot d\vec{r} = \int_{C_2} \vec{F}\cdot d\vec{r}$. So our fundamental theorem of line integrals tells us that line integrals of conservative vector fields are independent of path.

A curve is called *closed* if its initial and terminal points are the same;

i.e., if the curve makes a loop. See Figure 5.5 for examples of curves which are closed, and Figure 5.6 for curves which are not closed. Notice that if

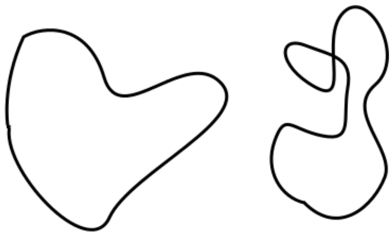


Figure 5.5: Examples of closed curves.

$\int_C \vec{F} \cdot d\vec{r}$ is independent of path, then $\int_C \vec{F} \cdot d\vec{r} = 0$ for any closed curve C. Writing the closed curve C as a *concatenation* of two curves, say C_1 and C_2 with endpoints A and B, we calculate

$$\begin{aligned}\int_C \vec{F} \cdot d\vec{r} &= \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{C_2} \vec{F} \cdot d\vec{r} \\ &= (A - B) + (B - A) \\ &= 0\end{aligned}$$

In fact, saying that a line integrals is independent of path is the same thing as saying that the integral along every closed path is 0. That is, $\int_C \vec{F} \cdot d\vec{r}$ is independent of path if and only if $\int_C \vec{F} \cdot d\vec{r} = 0$ for every closed curve C.

By the fundamental theorem of line integrals, we know line integrals of conservative vector fields are independent of path, but are there any other examples? To answer that question we need to introduce some terminology.

Recall that a set is *open* if you can put a small disk around each point which stays inside the set. A set is *connected* if between any two points in the set there is a continuous curve connecting the points which stays

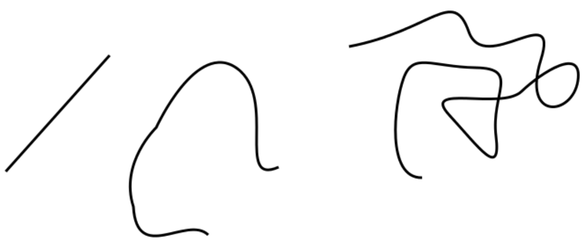


Figure 5.6: Examples of non-closed curves.

inside the set. In Figure 5.7 there are examples of connected sets, while in Figure 5.8 we see an example of a set which is not connected. (Note that a non-connected set can still be decomposed into connected pieces, however.)

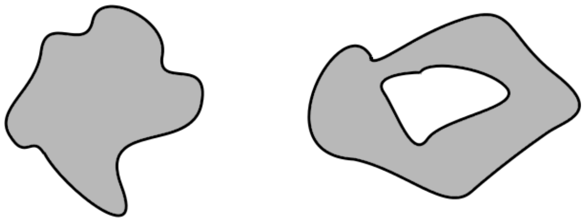


Figure 5.7: Two connected sets.

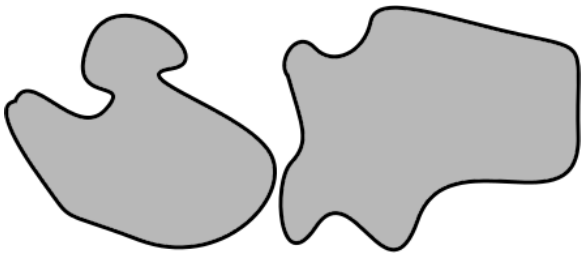


Figure 5.8: A set which is not connected.

Theorem 5.2.
Suppose $\vec{F}$ is a continuous vector field defined on an open connected region D. If $\int_C \vec{F} \cdot d\vec{r}$ is independent of path, then $\vec{F}$ is a conservative vector field.

So anytime we have a vector field where $\int_C \vec{F} \cdot d\vec{r}$ is independent of path (so $\int_C \vec{F} \cdot d\vec{r} = 0$ any time C is a closed curve), then $\vec{F}$ *has* to be a conservative vector field: i.e., there is some function f so that $\vec{F} = \nabla f$.

Theorem 5.3.
If $\vec{F}(x, y)$ is a conservative vector field, say

$$\vec{F}(x, y) = \langle P(x, y), Q(x, y) \rangle$$

Figure 5.9: Some simple curves.

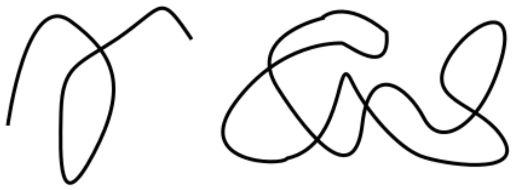


Figure 5.10: Some curves which are not simple.

and if P and Q have continuous first-order partials, then

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

So far everything has assumed $\vec{F}$ is a conservative vector field, but if you're just given a vector field, how can you tell if it's conservative or not?

To answer this, we need to introduce two new ideas:

1. A curve C is called *simple* if it does not intersect itself, except possibly to close up.

2. A region D is called *simply connected* if it's connected and every simple curve lying in D only bounds points in D. Equivalently, every closed curve in D can be continuously shrunk down to a single point, without ever leaving D.

Theorem 5.4.
Suppose $\vec{F} = \langle P, Q \rangle$ is a vector field on a simply connected region D. Then

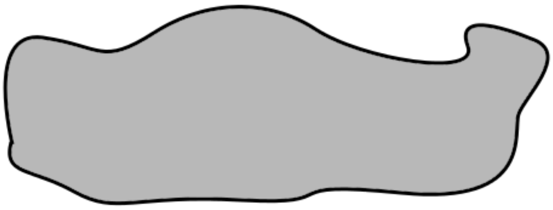


Figure 5.11: A simply connected set.

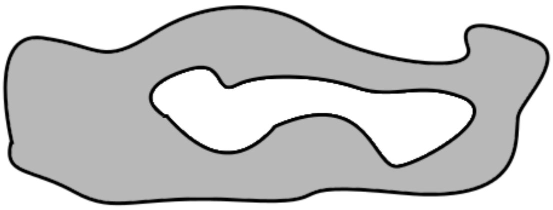


Figure 5.12: A set which is not simply connected.

if P and Q have continuous first-order partials and

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

then $\vec{F}$ is conservative.

The proof of this theorem will have to wait until after we've learnt Green's theorem.

Example 5.6.
Is the vector field $\vec{F}(x, y) = \langle e^x \sin(y), e^x \cos(y) \rangle$ conservative?
Notice that our vector field is defined on all of $\mathbb{R}^2$, which is obviously simply connected. Hence we can apply the above theorem:

$$\frac{\partial P}{\partial y} = e^x \cos(y) = \frac{\partial Q}{\partial x}$$

Hence this vector field is conservative.

Example 5.7.
Is the vector field $\vec{F}(x, y) = \langle x^2, y^2 \rangle$ conservative? If so, what is its potential function?

To check for conservativity, we apply our theorem above:

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y} x^2 = 0 = \frac{\partial}{\partial x} y^2 = \frac{\partial Q}{\partial x}$$

so $\vec{F}$ is conservative. To find the potential function, we need to find a function $f(x, y)$ such that

$$f_x(x, y) = x^2$$
$$f_y(x, y) = y^2$$

Integrating, we get

$$f(x, y) = \int x^2 \, dx = \frac{x^3}{3} + g(y)$$
$$f(x, y) = \int y^2 \, dy = \frac{y^3}{3} + h(x).$$

Comparing the two we see

$$g(y) = \frac{y^3}{3} + K$$
$$h(x) = \frac{x^3}{3} + K$$

So

$$f(x, y) = \frac{1}{3}\left(x^3 + y^3\right) + K.$$

Now if we want to integrate $\int_C \vec{F} \cdot d\vec{r}$, it's easy. Say C is the line segment $(1,2)$ to $(3,4)$. Then we have

$$\begin{aligned}\int_C \vec{F}\cdot d\vec{r} &= \int_C \nabla f \cdot d\vec{r}\\ &= f(3,4) - f(1,2)\\ &= \frac{1}{3}\left(3^3+4^3\right) + K - \frac{1}{3}\left(1^3+2^3\right) - K\\ &= \frac{1}{3}\left(-1^3-2^3+3^3+4^3\right)\\ &= \frac{1}{3}\left(-1-8+27+64\right)\\ &= \frac{82}{3}\end{aligned}$$

Kinetic and Potential Energy

To finish off this section, we'll explain where the terms "conservative" and "potential" come from. Suppose we want to move an object of mass M from point A to point B, along a curve C parametrized by $\vec{r}(t)$ for $a \leq t \leq b$. Recalling $\vec{F}(t) = m\vec{a}(t)$, $\vec{a}(t) = \vec{v}\,'(t)$, and $\vec{v}(t) = \vec{r}\,'(t)$, we know that the force used at a point $\vec{r}(t)$ on the curve is

$$\vec{F}(\vec{r}(t)) = m\vec{r}\,''(t)$$

so the work done is

$$\begin{aligned}W &= \int_C \vec{F}\cdot d\vec{r}\\ &= \int_a^b \vec{F}(\vec{r}(t))\cdot \vec{r}\,'(t)\,dt\\ &= \int_a^b m\vec{r}\,''(t)\cdot \vec{r}\,'(t)\,dt.\end{aligned}$$

Now we can simplify this by recalling the product rule for derivatives involving dot products:

$$\frac{d}{dt}\vec{u}(t)\cdot\vec{v}(t) = \vec{u}\,'(t)\cdot\vec{v}(t) + \vec{u}(t)\cdot\vec{v}\,'(t).$$

Now notice

$$\begin{aligned}\frac{d}{dt}\vec{r}'(t)\cdot\vec{r}'(t) &= \vec{r}''(t)\cdot\vec{r}'(t)+\vec{r}'(t)\cdot\vec{r}''(t)\\ &= 2\vec{r}''(t)\cdot\vec{r}'(t).\end{aligned}$$

So we can rewrite our above integral as

$$\begin{aligned}W &= \int_C \vec{F}\cdot d\vec{r}\\ &= \int_a^b m\vec{r}''(t)\cdot\vec{r}'(t)\,dt\\ &= \frac{m}{2}\int_a^b \frac{d}{dt}\vec{r}'(t)\cdot\vec{r}'(t)\,dt\\ &= \frac{m}{2}\int_a^b \frac{d}{dt}\|\vec{r}'(t)\|^2\,dt\\ &= \frac{m}{2}\|\vec{r}'(t)\|^2\Big|_a^b\\ &= \frac{m}{2}\|\vec{r}'(b)\|^2-\frac{m}{2}\|\vec{r}'(a)\|^2\\ &= \frac{m}{2}\|\vec{v}(b)\|-\frac{m}{2}\|\vec{v}(a)\|.\end{aligned}$$

In physics, the quantity $\frac{mv^2}{2}$ is called the *kinetic energy* of an object at aparticular point, which we might denote by K. Hence we may write the work as

$$W = K(B) - K(A).$$

Suppose $\vec{F}$ is a conservative vector field, so $\vec{F} = \nabla f$ for some function f. Physicists call the function $-f(x, y, z)$ the *potential energy* of the object and often write $P(x, y, z) = -f(x, y, z)$. Thus we have

$$\vec{F} = \nabla f = -\nabla P.$$

By our fundamental theorem for line integrals we thus have

$$\begin{aligned}W &= \int_C \vec{F}\cdot d\vec{r}\\ &= \int_C -\nabla P\cdot d\vec{r}\\ &= -P(B)+P(A)\\ &= P(A)-P(B).\end{aligned}$$

However, we already know that $W = K(B) - K(A)$. Equating these we have

$$K(B) - K(A) = P(A) - P(B),$$

which we may rewrite as

$$P(B) + K(B) = P(A) + K(A)$$

and this is called the *law of conservation of energy*: the sum of the kinetic and potential energies never changes!

5.4 Green's theorem

In this section we will discuss *Green's theorem* which tells us that for certain types of functions, the double integral over a domain D in the plane will be equal to the line integral around a curve bounding D. That is, Green's theorem will give us a tool for certain types of double integrals into line integrals and vice versa. This can be extremely useful, as we will see, since this will sometimes allow us to rewrite a complicated integral as a simpler one.

Before jumping into Green's theorem, though, we need to discuss a few preliminary ideas.

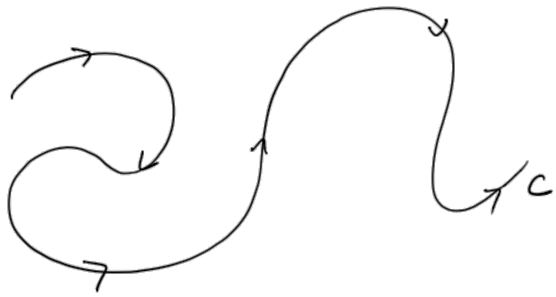
Orientation

A curve in the plane is simply a certain collection of points, but for some problems it is helpful to assign the curve a notion of direction. That is, in certain problems we don't care simply about the collection of points that make up the curve, but we care about the direction in which the curve is traversed. For example, if our curve represents the trajectory of a particle as it moves around and we want to compute the velocity of the particle, we need to know if the particle was traversing the curve from left-to-right or from right-to-left. An "orientation" for a curve is simply a choice of such a direction.

This idea of "direction" makes intuitive sense for orientation, but there are other equivalent ways we can define an orientation which will be easier to generalize later. One equivalent way of defining orientation is as a choice of which side of the curve is the "right-hand" side. If you imagine walking along the curve and sticking your right hand straight out to the right, then this distinguishes one side of the curve. Similarly, if you distinguished one side of the curve as being on the right, there would be only way to walk along the curve so that your right hand stayed on the side which was distinguished as the "right-hand side."

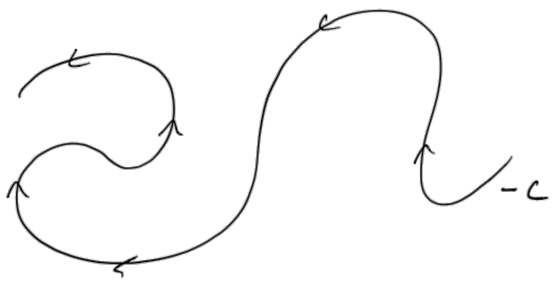
We can make this definition of orientation more precise as follows. At every point P of a curve C, choose a vector $\vec{N}(P)$ which is orthogonal to the tangent line of the curve at that point, but in such a way that $\vec{N}(P)$ is a continuous function. The continuity condition will force all of these normal vectors (the $\vec{N}(P)$ vectors) to lie on one one side of the curve. This is the side we will consider to be the right-hand side. (This definition might seem overly cumbersome, but we mention it now because it will be helpful to consider when we define the orientation of a surface later.)

Usually when we have particular curve C drawn in the plane, we will simply indicate the orientation by arrows. These arrows tell us how we're walking along the curve; equivalently, they tell us which side of the curve is the right-hand side.

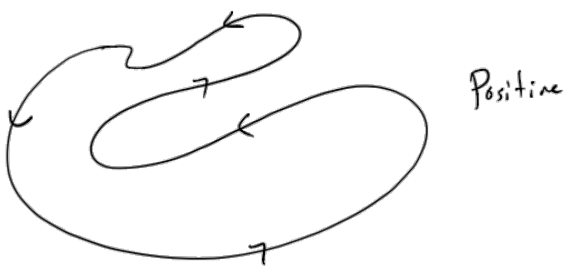


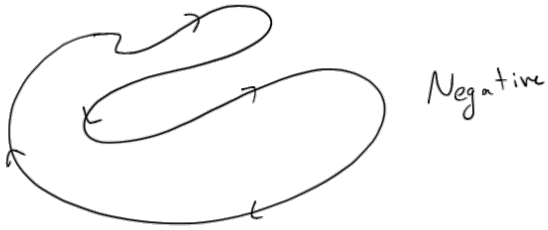
Notice that if C is given a parametrization, that parametrization supplies C with an orientation as it specifies an order in which C is traversed.

If C is an oriented curve, we will let $-C$ denote the same curve (the same set of points in the plane), but with the opposite orientation.

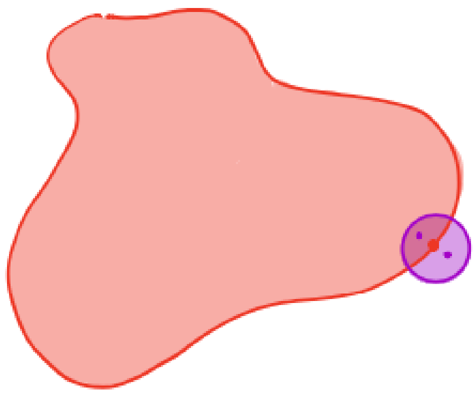


When discussing simple closed curves in the plane, we will also adopt the following convention: if the curve is traversed in the counter-clockwise direction, we will say the curve has ***positive orientation***; if the curve is traverse in the clockwise direction, it has ***negative orientation***.

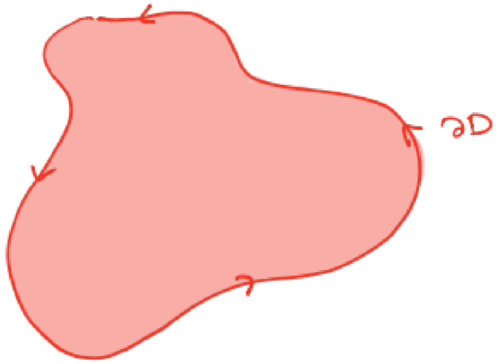




Recall that if $D \subseteq \mathbb{R}^2$ is a subset of the plane, then the ***boundary*** of D is the set of points where for every ball B centered at that point, B contains points in D and outside of D. The boundary is the collection of all such points and is denoted ∂D.



By convention, if ∂D is a single simple closed curve, we will assume it has the positive orientation. If for some reason we'd rather consider the negative orientation we will write $-\partial D$.



When C is a simple closed curve, we will write

$$\oint_C P\,dx + Q\,dy \quad \text{or} \quad \ointctrclockwise_C P\,dx + Q\,dy$$

To indicate that we are integrating over C with the positive orientation.

Example 5.8.
Let C be the circle of radius 3 centered at the point $(0,0)$. Compute

$$\oint_C x\,dy - y\,dx.$$

Since we are are told (via the $\oint$ notation) that C is positively oriented, we have to be a little careful that we orient the curve correctly. Fortunately it's easy to see that our "usual" parametrization of a circle,

$$x(t) = 3\cos(t),\; y(t) = 3\sin(t) \qquad 0 \le t \le 2\pi$$

does traverse the circle counterclockwise. We then simply compute

$$\begin{aligned}\oint_C x\,dy - y\,dx &= \int_0^{2\pi} \left(3\cos(t)\cdot 3\cos(t) - 3\sin(t)\cdot 3(-\sin(t))\right)\,dt \\ &= 9\int_0^{2\pi} \left(\cos^2(t) + \sin^2(t)\right)\,dt \\ &= 18\pi\end{aligned}$$

Green's theorem

We are now in a position to state Green's theorem.

Theorem 5.5 (Green's theorem)**.**
Suppose D is a simply connected region of the plane with piecewise smooth boundary ∂D. If $P(x,y)$ and $Q(x,y)$ both C^1 (i.e., all of the first-order partial derivatives of the functions exist and are continuous), then

$$\oint_{\partial D} P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\,dA$$

Example 5.9.
As an example of why Green's theorem is helpful suppose that C was a circle of radius 2 centered at the origin and we wished to compute

$$\oint_C (x-y)\,dx + (x+y)\,dy.$$

One way to do this would be to choose a parametrization of C agreeing with the positive orientation, then rewrite our integral in terms of the parametrization. Using the parametrization $(2\cos(t), 2\sin(t))$ with $0 \le t \le 2\pi$, for example, the integral could be computed as

$$\int_0^{2\pi} \left[(2\cos(t) - 2\sin(t)) \cdot (-2\sin(t)) + (2\cos(t) + 2\sin(t)) \cdot (2\cos(t))\right]\,dt$$

We *can* compute this integral, but it's very tedious. However, Green's theorem tells us that we may write this as

$$\iint\limits_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\, dA$$

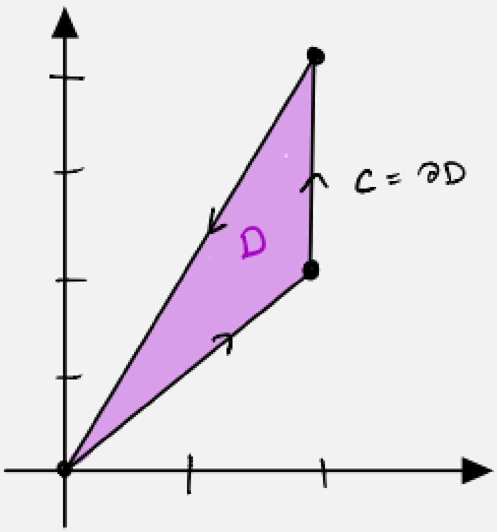
where $P(x, y) = x - y$ and $Q(x, y) = x + y$, and D is the disc of radius 2 centered at the origin. This would give us the following:

$$\begin{aligned}\oint\limits_C (x-y)\,dx + (x+y)\,dy &= \iint\limits_D (1 - (-1))\, dA \\ &= 2\iint\limits_D dA \\ &= 2\text{Area}(D) \\ &= 8\pi.\end{aligned}$$

Example 5.10.
As another example, let's use Green's theorem to evaluate

$$\oint\limits_C xy^2\,dx + 2x^2y\,dy$$

where C is the triangle with vertices $(0, 0)$, $(2, 2)$, and $(2, 4)$.



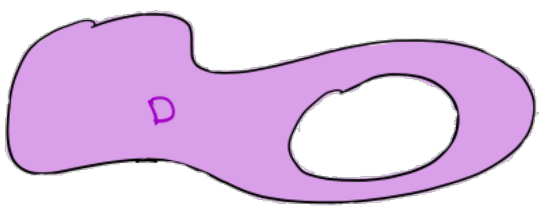
Notice that without Green's theorem we would have to break our integral up into three pieces, one piece for each edge of the triangle, and then determine a parametrization for each of those edges. Using Green's theorem, however, all we need to do is consider the "interior" of that triangle. If that is called D, then we easily see

$$D = \left\{ (x, y) \,\middle|\, 0 \leq x \leq 2,\ x \leq y \leq 2x \right\}.$$

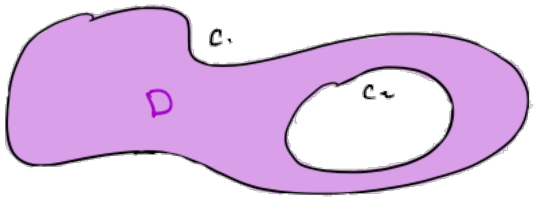
Now Green's theorem tells us the following:

$$\begin{aligned}
\oint_C xy^2\,dx + 2x^2y\,dy &= \iint_D (4xy - 2xy)\,dA \\
&= 2\iint_D xy\,dA \\
&= 2\int_0^2 \int_x^{2x} xy\,dy\,dx \\
&= 2\int_0^2 \left.\frac{xy^2}{2}\right|_x^{2x} dx \\
&= \int_0^2 \left(4x^3 - x^2\right)\,dx \\
&= \left.\left(x^4 - \frac{x^3}{3}\right)\right|_0^2 \\
&= 16 - {}^8\!/_3 \\
&= \frac{40}{3}.
\end{aligned}$$

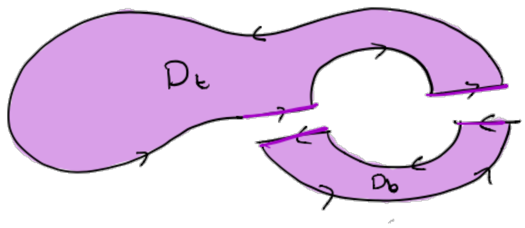
So far we've assumed the region we were integrating over was simply connected (i.e., doesn't have any "holes"). We can use Green's theorem for regions containing holes as well, but we have to be alittle bit careful about how the "interior boundaries" of thre gion are oriented. To understand this, suppose we had a region D in the plane which had a hole in it.



Call the "exterior boundary" of this region C_1 and the "interior region" of the boundary C_2 as indicated in the figure.



Now imagine that we cut our non-simply connected region D into two simply connected regions by cutting along line segments connecting the interior and exterior boundaries. For the region indicated in the figure, we may call the top region D_t and the bottom region D_b.



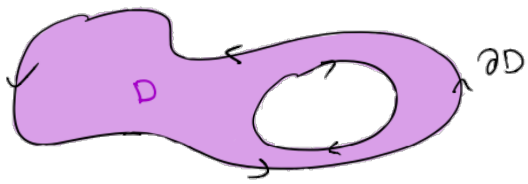
Now notice that we may write an integral over D as

$$\iint\limits_{D} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint\limits_{D_t} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA + \iint\limits_{D_b} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

When integrating over D_t and D_b we may apply Green's theorem to write these integrals as

$$p \oint\limits_{\partial D_t} P\,dx + Q\,dy + \oint\limits_{\partial D_b} P\,dx + Q\,dy$$

Notice that giving ∂D_t and ∂D_b have positive orientations induces an orientation on the portions of the interior boundaries of D contained in those regions. Also notice that the portions of ∂D_t and ∂D_b containing those slits connecting the interior and exterior regions have opposite orientations, thus the contributions those portions of the boundary make to the total integral will cancel when we add the integrals back together. This means Green's theorem will still be true for non-simply connected regions provided we orient the interior boundaries correctly, which will actually give those boundary components the negative (clockwise) orientation.



Making this convention our statement of Green's theorem,

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA = \oint_{\partial D} P\,dx + Q\,dy$$

still holds.

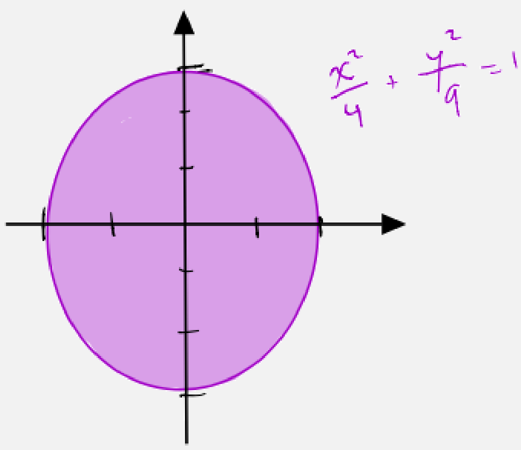
Computing area with Green's theorem

One interesting application of Green's theorem is the computation of area of a region by integrating around it's boundary. To see this, recall that the area of a region D can be computed as $\iint_D 1\,dA$. If we can find two functions P and Q so that $Q_x - P_y = 1$, then by Green's theorem we can compute the area of D by integrating $\oint_{\partial D} P\,dx + Q\,dy$. There are several options for such functions, but three easy ones are

- $P(x, y) = 0$ and $Q(x, y) = x$.
- $P(x, y) = -y$ and $Q(x, y) = 0$.
- $P(x, y) = {}^{-y}/_2$ and $Q(x, y) = {}^{x}/_2$.

For any of these we have $Q_x - P_y = 1$ and this gives us an alternative way of computing area.

Example 5.11.
Area of ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$.



By Green's theorem we may compute this area as

$$\text{Area}(D) = \iint_D 1\, dA = \oint_{\partial D} \frac{-y}{2}\, dx + \frac{x}{2}\, dy.$$

To perform this calculation we need to orient the ellipse, which is easily done with

$$x(t) = 2\cos(t),\ y(t) = 3\sin(t) \qquad 0 \le t \le 2\pi.$$

Now we simply compute

$$\begin{aligned}\oint_{\partial D} \frac{-y}{2}\, dx + \frac{x}{2}\, dy &= \int_0^{2\pi} \left(\frac{-3\sin(t)}{2} \cdot (-2\sin(t)) + \frac{2\cos(t)}{2} \cdot 3\cos(t) \right)\, dt \\ &= 3\int_0^{2\pi} \left(\sin^2(t) + \cos^2(t)\right)\, dt \\ &= 6\pi\end{aligned}$$

Given any complicated region in the plane this gives us a reasonable way of computing the area of that region if we can parametrize (or approximately parametrize) that region's boundary. For example, perhaps we want to know the area of a large body of water. If we can have a drone fly around the boundary of that region and record its position using GPS, we can approximate a parametrization to the boundary of the region (e.g., using a piecewise linear path connecting the coordinates of several points recorded by the drone as it flew around the region). We can then compute an integral such as $\oint_{\partial D} x\,dy$ (this is very easy to do if the boundary consists of line segments) and compute (or approximate) the area of that body of water.

5.5 Curl and divergence

We now turn our attention to operations we can perform on three-dimensional vector fields called *curl* and *divergence*. As we will see, curl gives us a way of measuring how a vector field "spins" around a point, while the divergence gives a notion of how the vector field "spreads out" from a point.

Intuition for Curl

Before describing the curl technically, let's first describe the quantity we're trying to measure.Imagine that a vector field represents velocities of particles moving in a fluid. (E.g., think of a river of water, and the vectors tell you the speed and direction of particles that get pushed around by the river.) Now suppose that we take a small ball, like a balloon or beach ball, and place it in the water. Instead of just letting the river push the ball like normal, let's say we're somehow able to anchor the ball so that the center of the ball stays fixed at a point in the river; however, we allow the ball to spin about this fixed center point. We'd like to describe how the ball spins as the water pushes the sides of the ball. This is essentially what the curl of a vector field measures.

In order to measure how the ball spins as the water pushes it, we need to know three pieces of information: what axis is it spinning around, which way is it spinning (clockwise or counter-clockwise), and how quickly is it spinning. We can put these three pieces of information to obtain a vector: the axis of rotation will tell us a line parallel to the direction of the vector, and the speed will tell us the magnitude of this vector. We now have to pick a sign for the vector (e.g., if we know the axis of rotation is vertical, should the vector point up or should it point down)? Here we'll apply a convention called the *right-hand rule*. If we take our right hand and curl our fingers in the direction of rotation, our thumb will determine the direction of the vector.

The ∇-operator

Before beginning our study of curl, we'll introduce some notation that will simplify things a little bit later on. Recall that if $f(x, y, z)$ is a real-valued function whose first-order partial derivatives exist, then the gradient of f is the vector field

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

We'll be performing several manipulations that involve first-order partial derivatives, and so we'll introduce an operator which picks off these derivatives. This operator is denoted ∇ and called "del." We think of the del operator as a vector

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

and will manipulate ∇ in the same way that we manipulate vectors (e.g., taking cross products and dot products) to construct new vector fields.

Curl

Suppose that $\vec{F}$ is a three-dimensional vector-field,

$$\vec{F}(x, y, z) = \langle P(x, y, z),\ Q(x, y, z),\ R(x, y, z) \rangle ,$$

and further suppose that the first-order partial derivatives of each component function, P, Q, and R, exist and our continuous. We then define an operation called "curl" which takes our vector field, $\vec{F}$, and produces a new vector field, denoted $\mathrm{curl}(\vec{F})$, as follows:

$$\mathrm{curl}(\vec{F}) = \left\langle \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right), \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right), \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \right\rangle .$$

Luckily this complicated-looking formula can be easily remembered by using the ∇ operator. Thinking of $\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle$ and $\vec{F} = \langle P, Q, R \rangle$ as vectors, we may write

$$\begin{aligned}
\operatorname{curl}(\vec{F}) &= \nabla \times \vec{F} \\
&= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{bmatrix} \\
&= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{\imath} - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right)\vec{\jmath} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k} \\
&= \left\langle \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right), \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right), \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \right\rangle.
\end{aligned}$$

The reason for calling this the "curl" of a vector field is that it represents how the vector field spins about a point.

Given a point (a, b, c), if $\vec{F}$ spins in a disk around (a, b, c), then $\operatorname{curl}(\vec{F})(a, b, c)$ is a vector which is perpendicular to the disk (with the direction given by the right-hand rule), and the magnitude tells us how quickly the vector field is rotating.

To explain this, consider a solid disk rotating about a fixed axis. If

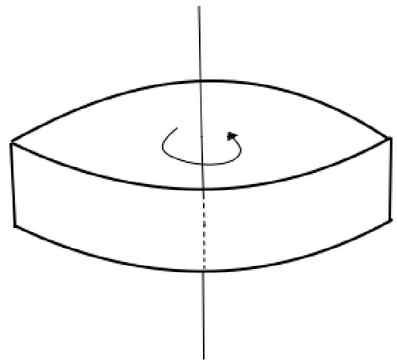


Figure 5.13: A disc rotating about a fixed axis.

this fixed axis is the positive z-axis, then we can represent the rotation of the disk by the vector, using our conventions described at the start of the lecture, $\vec{w} = \omega\vec{k} = \langle 0, 0, \omega\rangle$, where ω is the angular speed of rotation. Let's pick a point $P = (x, y, z)$ on the disk, which we'll suppose is d units from the axis of rotation. Let $vecr = \vec{OP} = \langle x, y, z\rangle$. If $\vec{v}$ is the tangential velocity of P as we spin around the z-axis, the angular speed is $\frac{\|\vec{v}\|}{d}$ – which is what we're calling ω.

Recall that $\vec{F}(x, y) = \langle -y, x\rangle$ was the vector field giving rotation around the origin, So $\vec{F}(x, y, z) = \langle -y, x, 0\rangle$ is the vector field giving rotation

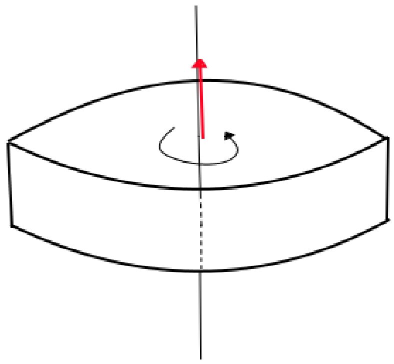


Figure 5.14: A disc rotating about a fixed axis, with a vector describing the rotation.

about the z-axis, if we want this rotation to have speed $\|\vec{v}\|$ we'd use

$$\frac{\|\vec{v}\|}{d}\langle -y, x, 0\rangle$$

where $d = \sqrt{x^2+y^2}$. So the tangential velocity at P is

$$\frac{\|\vec{v}\|}{d}\langle -b, a, 0\rangle .$$

Notice this equals $\vec{v} = \vec{w}\times\vec{r}$:

$$\begin{aligned}\vec{w}\times\vec{r} &= \left\langle 0, 0, \frac{\|\vec{v}\|}{d}\right\rangle \times \langle a, b, c\rangle \\ &= \det\begin{bmatrix}\vec{\imath} & \vec{\jmath} & \vec{k} \\ 0 & 0 & \frac{\|\vec{v}\|}{d} \\ a & b & c\end{bmatrix} \\ &= -\frac{\|\vec{v}\|}{d}b\vec{\imath} + \frac{\|\vec{v}\|}{d}a\vec{\jmath} + 0\vec{k} \\ &= \left\langle -\frac{\|\vec{v}\|}{d}b, \frac{\|\vec{v}\|}{d}a, 0\right\rangle \\ &= \frac{\|\vec{v}\|}{d}\langle -b, a, 0\rangle \\ &= \vec{v}.\end{aligned}$$

In general, if we have angular speed ω, the tangential velocity at (x, y, z) is

$$\vec{v} = \langle -\omega y, \omega x, 0\rangle .$$

Note

$$\begin{aligned}
\operatorname{curl}(\vec{v}) &= \nabla \times \vec{v} \\
&= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{bmatrix} \\
&= \left(-\frac{\partial}{\partial z}\omega x\right)\vec{\imath} - \left(\frac{\partial}{\partial z}\omega y\right)\vec{\jmath} + \left(\frac{\partial}{\partial x}\omega x + \frac{\partial}{\partial y}\omega y\right) \\
&= 2\omega\vec{k} \\
&= 2\vec{w}
\end{aligned}$$

Example 5.12.
Find the curl of the vector field

$$\vec{F}(x,y,z) = \langle xye^z, 0, yze^x\rangle \,.$$

$$\begin{aligned}
\operatorname{curl}\vec{F} &= \nabla \times \vec{F} \\
&= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xye^z & 0 & yze^x \end{bmatrix} \\
&= \left(\frac{\partial}{\partial y}yze^x\right)\vec{\imath} - \left(\frac{\partial}{\partial x}yze^x - \frac{\partial}{\partial y}xye^z\right)\vec{\jmath} + \left(-\frac{\partial}{\partial y}xye^z\right)\vec{k} \\
&= \langle ze^x, xe^z - yze^x, -xe^z\rangle \,.
\end{aligned}$$

One of the crucial properties of the curl is that the curl of a conservative vector field is always zero.

Theorem 5.6.
If f is a continuous function of three variables with second-order continuous partials, then $\operatorname{curl}(\nabla f) = 0$.

This gives us an easy way to see if a vector field of three variables is conservative or not: if $\text{curl}\vec{F} \neq 0$, then $\vec{F}$ can not be conservative.

For example, consider

$$\vec{F}(x, y, z) = \left\langle xz, xyz, -y^2 \right\rangle.$$

Computing the curl,

$$\begin{aligned}
\text{curl}\vec{F} &= \nabla \times \vec{F} \\
&= \det \begin{bmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz & xyz & -y^2 \end{bmatrix} \\
&= \left(\frac{\partial}{\partial y} \left(-y^2\right) - \frac{\partial}{\partial z} xyz \right) \vec{\imath} + \left(\frac{\partial}{\partial z} xz + \frac{\partial}{\partial y^2} \right) \vec{\jmath} + \left(\frac{\partial}{\partial x} xyz - \frac{\partial}{\partial y} xz \right) \vec{k} \\
&= (-2y - xy)\vec{\imath} + (x + 0)\vec{\jmath} + (yz - 0)\vec{k} \\
\neq 0.
\end{aligned}$$

Hence $\vec{F}$ can not be conservative.

In general the converse is not true: if $\text{curl}\vec{F} = 0$, that does not imply that $\vec{F}$ is conservative. It is true, however, if the domain of $\vec{F}$ is all of $\mathbb{R}^3$.

Theorem 5.7.
If $\vec{F}$ is a defined on all of $\mathbb{R}^3$ and if $\text{curl}\vec{F} = 0$, then $\vec{F}$ is conservative.

This is a very nice theorem to know because it tells you precisely when a vector field is conservative – without having to calculate a potential function!

Divergence

The divergence of a vector field gives us a way of measuring how the vector field pushes away from a given point. In particular, if have a particular point $P_0 = (x_0, y_0)$, we can imagine placing a small circle C_r of radius r around P_0 and at each point (x, y) on this circle we let $\vec{N}(x, y)$ denote the unit normal of the circle which points away from P_0. The dot product of $\vec{F}$ and $\vec{N}$ at this point then gives us a number, $\vec{F} \cdot \vec{N} = \|\vec{F}\| \cos(\theta)$ where

θ is the angle between the vectors determined by $\vec{F}$ and $\vec{N}$. If the angle between $\vec{F}$ and $\vec{N}$ is less than π, this quantity is positive. That is, $\vec{F} \cdot \vec{N}$ gives us a way of measuring "how much" of $\vec{F}$ points in the direction of $\vec{N}$ (away from P_0). We then compute the average value of these quantities on the circle as

$$\frac{1}{2\pi r} \oint_{C_r} \vec{F} \cdot \vec{N} \, ds.$$

(The $\frac{1}{2\pi r}$ arises since this is dividing the integral by the arclength of the circle to get an average value.) The divergence of $\vec{F}$ at P_0 is defined as the limit of these values as the radius of the circle r goes to zero, and gives us an "infinitesimal" measurement of how the vector field $\vec{F}$ pushes away from the point P_0.

Though the definition of divergence is in terms of limits of integrals, it can actually be computed quite easily:

Theorem 5.8.
If $\vec{F}(x, y) = \langle P(x, y), Q(x, y) \rangle$ is a vector field whose components are C^1, then the divergence is equal to

$$\operatorname{div}(\vec{F}) = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}.$$

Using the del operator ∇ from above, it's easy to see that we can also write

$$\operatorname{div}(\vec{F}) = \nabla \cdot \vec{F}.$$

(Here we are using the two-dimensional version of the del operator, $\nabla = \left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\rangle$.)

Example 5.13.
The divergence of the vector field $\vec{F}(x, y) = \langle x^3 y + y^2, xy - x^2 + y \rangle$

is

$$\operatorname{div}(\vec{F}) = \frac{\partial}{\partial x}\left(x^3y + y^2\right) + \frac{\partial}{\partial y}\left(xy - x^2 + y\right) = 3x^2y + y + 1$$

Notice that the divergence of a vector field is a scalar-valued quantity that varies from point to point: i.e., a function. At points where this function is positive this means the vector field is pushing away from that point more than it is pulling in; at points where the function is negative the vector field is pulling into the point more than it is pushing away; and where the divergence is zero the vector field is pushing away and pulling in at the same rate.

We can define divergence for three-dimensional vector fields in a way similar to the definition for two-dimensional vector fields above, but that would require us to discuss "surface integrals" which have not yet done. Despite this, we can still compute divergence of three-dimensional vector fields using the formula above, $\operatorname{div}(\vec{F}) = \nabla \cdot \vec{F}$.

Example 5.14.
The divergence of

$$\vec{F}(x, y, z) = \left\langle x^2y, xy - z, xy \right\rangle$$

is equal to

$$\begin{aligned} &\operatorname{div}(\vec{F}) \\ =&\nabla \cdot \vec{F} \\ =&\left\langle \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right\rangle \cdot \langle x^2y, xy - z, xy\rangle \\ =&\frac{\partial}{\partial x}x^2y + \frac{\partial}{\partial y}(xy - z) + \frac{\partial}{\partial z}xy \\ =&2xy + x \end{aligned}$$

When a vector field $\vec{F}$ has zero divergence, $\operatorname{div}(\vec{F}) = \nabla \cdot \vec{F} = 0$, we say the vector field has ***incompressible***.

Relating gradients, curl, and div: Next steps

Notice we have the following operations we can perform:

- Given a function f we can take its gradient to obtain a vector field ∇f.
- Given a vector field $\vec{F}$ we can take its curl to compute a new vector field $\text{curl}(\vec{F})$.
- Given a vector field $\vec{F}$ we can take its divergence to compute a function $\text{div}(\vec{F})$.

Is there any relationship between these operations? We have already seen above that if we compute the curl of a conservative vector field we must obtain the zero vector field: $\text{curl}(\nabla f) = \vec{0}$. It can be shown as well that the divergence of a the curl of a vector field is always the zero function: $\text{div}(\text{curl}(\vec{F})) = 0$. That is, we have a sequence of operations which whenever we apply two "consecutive" operations we obtain zero. Though it's hard to appreciate right now, this is the start of a lot of very interesting mathematics. If there were a follow-up to this course (e.g., a fourth semester of calculus), exploring and generalizing these ideas would be a natural place to start and would lead us to the theory of differential forms which unifies much of the material presented in this last chapter of the notes. The interested student is encouraged to read about these topics in more advanced books such as *Differential Forms* by Guillemin and Haine, *From Calculus to Cohomology* by Madsen and Tornehave, *Differential Topology* by Guillemin and Pollack, and *Calculus on Manifolds* by Spivak.

Flux and the vector form of Green's theorem

In defining divergence above we considered the limit of integrals of the form $\oint\limits_{C_r} \vec{F} \cdot \vec{N}\, ds$ where $\vec{N}$ was the unit normal vector to the circle C_r at each point. More generally, line integrals of the form

$$\int\limits_C \vec{F} \cdot \vec{N}\, ds$$

for any curve C (not just circles) are sometimes referred to as the ***flux*** of the vector field across C. This is a measurement of how much of the vector field flows across the curve C. That is, if the vector field $\vec{F}$ represented the velocity of particles in a fluid, this quantity tells us how much of the fluid flows across the curve per unit time.

One subtlety in this calculation is that there are two possibilities for what $\vec{N}$ could be since our curve C has two possible orientations. The different orientations only affect the sign of the flux, however: if $\vec{N}_1$ corresponds to one orientation and $\vec{N}_2$ is the opposite orientation, the two different flux calculations are related by

$$\int\limits_C \vec{F} \cdot \vec{N}_2 \, ds = -\int\limits_C \vec{F} \cdot \vec{N}_1 \, ds.$$

Example 5.15.
Compute the flux of $\vec{F}(x, y) = \langle x^2 y, y + 3x \rangle$ across the right-hand half of the unit circle centered at the origin, with the unit normal vectors pointing away from the origin.

To do the computation, we need to determine the unit normal vector at each point on our semicircle, and for this it's helpful to have a parametrization. If we parametrize our semicircle by

$$\vec{r}(t) = \langle x(t), y(t) \rangle = \langle \cos(t), \sin(t) \rangle, \quad -\pi/2 \leq t \leq \pi/2$$

then the tangent vectors would be

$$\vec{r}\,'(t) = \langle x'(t), y'(t) \rangle = \langle -\sin(t), \cos(t) \rangle.$$

Conveniently these are already unit vectors, $\|\vec{r}\,'(t)\| = 1$; so $\vec{T}(t) = \vec{r}\,'(t)$ in this example. The derivative is

$$\vec{T}\,'(t) = \langle -\cos(t), -\sin(t) \rangle.$$

Again, these are already unit vectors, so we don't need to worry about scaling them. However, these vectors point *towards* the origin, whereas we want to do the computation with the normal vectors pointing away from the origin. Thus we will take

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

We can now compute the flux as

$$\begin{aligned}
&\int_C \vec{F} \cdot \vec{N}\, ds \\
=&\int_{-\pi/2}^{\pi/2} \left\langle \cos^2(t)\sin(t), \sin(t) + 3\cos(t)\right\rangle \cdot \langle \cos(t), \sin(t)\rangle \sqrt{\cos^2(t) + \sin^2(t)}\, dt \\
=&\int_{-\pi/2}^{\pi/2} \left(\cos^3(t)\sin(t) + \sin^2(t) + 3\cos(t)\sin(t)\right)\, dt \\
=&{}^{\pi}\!/_2 \approx 1.5708
\end{aligned}$$

The interpretation of this number is that our vector field is flowing across the curve C at a rate of 1.5708 units of area (since the vector field is two-dimensional) per unit time.

If D is a region in the plane, we may be interested in the flux of $\vec{F}$ through the boundary ∂D,

$$\oint_{\partial D} \vec{F} \cdot \vec{N}\, ds.$$

To compute this, we may at first be tempted to use a parametrization for ∂D as in the example above. However, in this case there is a simpler method. First suppose $\vec{r}(t) = \langle x(t), y(t)\rangle$ is a parametrization for ∂D (with the positive orientation) with $a \leq t \leq b$. It's not hard to compute that the outward pointing normals of ∂D are given by

$$\vec{N}(t) = \left\langle \frac{y'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}, \frac{-x'(t)}{\sqrt{x'(t)^2 + y'(t)^2}} \right\rangle.$$

(Since $\vec{N}(t)$ and $\vec{T}(t)$ are always orthogonal, this can be obtained from $\vec{T}(t)$ by a $90°$ counter-clockwise rotation.) Now, supposing $\vec{F}(x, y) =$

$\langle P(x,y), Q(x,y)\rangle$ we may write the integral for the flux as

$$\begin{aligned}
&\oint_{\partial D} \vec{F}\cdot\vec{N}\,ds\\
&= \int_a^b \langle P(x(t),y(t)), Q(x(t),y(t))\rangle \cdot \left\langle \frac{y'(t)}{\sqrt{x'(t)^2+y'(t)^2}}, \frac{-x'(t)}{\sqrt{x'(t)^2+y'(t)^2}} \right\rangle \sqrt{x'(t)^2+y'(t)^2}\,dt\\
&= \int_a^b \langle P(x(t),y(t)), Q(x(t),y(t))\rangle \cdot \langle y'(t), -x'(t)\rangle\ dt\\
&= \int_a^b (P(x(t),y(t))y'(t) - Q(x(t),y(t))x'(t))\ dt\\
&= \oint_{\partial D} -Q(x,y)dx + P(x,y)\,dy.
\end{aligned}$$

Though the order of P and Q is backwards from the way we have written them before, this is an integral where we can apply Green's theorem, which would tell us this integral equals

$$\iint_D \left(\frac{\partial}{\partial x}P(x,y) - \frac{\partial}{\partial y}(-Q(x,y))\right)\,dA.$$

(Notice this looks slightly different from our earlier form of Green's theorem because the roles of P and Q are reversed, and we have an extra negative.) Of course, this can be written more simply as

$$\iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right)\,dA,$$

but the integrand here is precisely the divergence of the vector field $\vec{F}$. That is, if $\vec{N}$ represents is the outward pointing unit normals on the boundary of D, then the flux of $\vec{F}$ across the boundary of D can be computed as the the integral of the divergence of $\vec{F}$ over the region D:

$$\oint_{\partial D} \vec{F}\cdot\vec{N}\,ds = \iint_D \mathrm{div}(\vec{F})\,dA.$$

This is sometimes referred to as ***the vector form of Green's theorem***.

Electromagnetism

We end this section by briefly mentioning that the these notions of vector fields, gradients, divergence, etc. are the language for the modern formulation of electromagnetism in physics. In particular, the equations that govern the relationships between electric fields and magnetic fields are summed up by the following four equations, known collectively as *Maxwell's equations*:

$$\begin{aligned}
\nabla \cdot \vec{E} &= \frac{\rho}{\varepsilon_0} \\
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{E} &= \frac{-\partial \vec{B}}{\partial t} \\
\nabla \times \vec{B} &= \mu_0 \left(\vec{J} + \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \right).
\end{aligned}$$

Since this is a course in calculus and not electrodynamics, we won't take any time to explain what these equations mean, but will simply mention them to point out that all of this material does have important applications. Interested students should read *Div, Grad, Curl and All That* by Schey, *A Student's Guide to Maxwell's Equations* by Fleisch, and *The Feynman Lectures on Physics, Volume II* by Feynman.

A

Proofs of results from the text

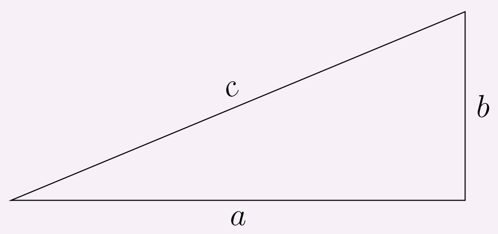
In this appendix we collect proofs for many of the theorems given in the main part of the text. The proofs are collected here for the sake of completeness, and are not required reading. Students planning in majoring in mathematics, though, may be interested in trying to understand some of the proofs that appear here.

A.1 Proofs from Chapter 1

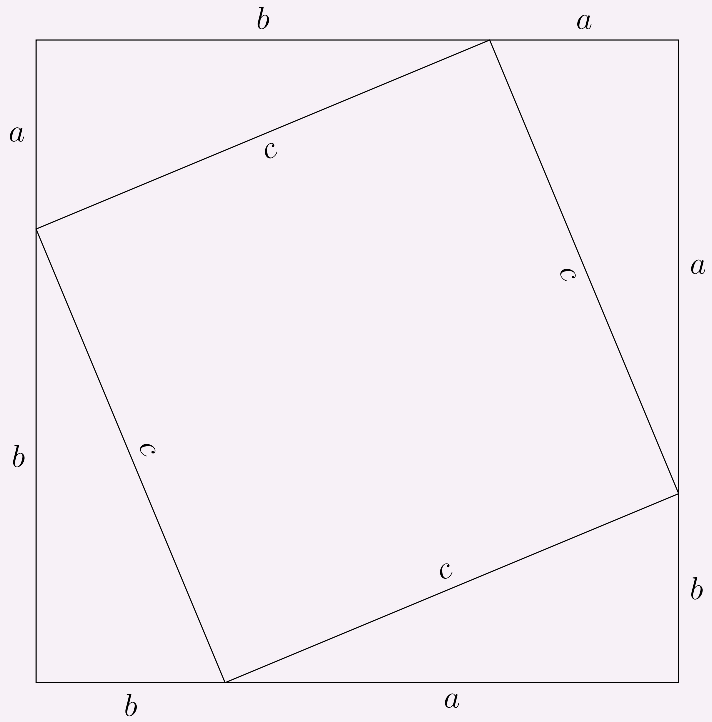
Proofs from Section 1.1, Three-dimensional space

Proof of Theorem 1.1 on page 9, The Pythagorean Theorem.
There are several ways to prove the Pythagorean theorem, but we will give what might be the simplest to understand proof using simple algebra and geometry.

Suppose we have a right triangle whose legs have lengths a and b and whose hypotenuse has length c.



Now consider taking four copies of this triangle and arranging them as indicated below. Notice This results in two squares: the outer square is has dimensions $(a+b)\times(a+b)$ and the inner square has dimensions $c \times c$.



The outer square has area $(a+b)^2$ and there are two different ways we could think about the area of the inner square: the inner square must have area c^2 since it is a $c \times c$ square, but it must also be equal to the area of the outer square minus the areas of our four triangles on the corners which is

$$(a+b)^2 - 4 \cdot \frac{1}{2}ab = a^2 + 2ab + b^2 - 2ab = a^2 + b^2.$$

Both of these quantities, c^2 and a^2+b^2, represent the area of the inner square, so they must be equal: $c^2 = a^2 + b^2$. □

Proof of Theorem *1.2 on page 15**, Completing the Square.*
We must show that $x^2 + bx + c$ is equal to $(x + {}^b\!/\!{}_2)^2 - {}^{b^2}\!/\!{}_4 + c$. This is

easily done by expanding and canceling:

$$\left(x+\frac{b}{2}\right)^2-\frac{b^2}{4}+c = x^2+2x\frac{b}{2}+\frac{b^2}{4}-\frac{b^2}{4}+c$$
$$= x^2+bx+c$$

□

Proofs from Section 1.3, Vectors

Proof of Theorem 1.4 on page 36.
We'll just show the third property in the case of two-dimensional vectors to outline the basic logic. All of the other properties are proved similarly.

$$\begin{aligned}(\lambda+\mu)\vec{v} =& (\lambda+\mu)\langle v_1, v_2\rangle\\ =& \langle(\lambda+\mu)v_1, (\lambda+\mu)v_2\rangle\\ =& \langle\lambda v_1+\mu v_1, \lambda v_2+\mu v_2\rangle\\ =& \langle\lambda v_1, \lambda v_2\rangle+\langle\mu v_1, \mu v_2\rangle\\ =& \lambda\langle v_1, v_2\rangle+\mu\langle v_1, v_2\rangle\\ =& \lambda\vec{v}+\mu\vec{v}.\end{aligned}$$

□

Proofs for Section 1.4, Linear transformations and matrices

Proof of Theorem 1.5 on page 52.
Let $\vec{u}$ and $\vec{v}$ be n-dimensional vectors. Since T is assumed to be linear, we have that $T(\vec{u}+\vec{v}) = T(\vec{u})+T(\vec{v})$. Notice these are both m-dimensional vectors, and so their sum is an m-dimensional vector. We can thus write

$$S(T(\vec{u})+T(\vec{v})) = S(T(\vec{u}))+S(T(\vec{v}))$$

Similarly, since T is linear, for every scalar λ we have $T(\lambda\vec{v}) = \lambda T(\vec{v})$. Since S is linear, we also have

$$S(\lambda T(\vec{v})) = \lambda S(T(\vec{v})).$$

We have thus shown

$$S \circ T(\vec{u} + \vec{v}) = S(T(\vec{u} + \vec{v}))S(T(\vec{u})) + S(T(\vec{v})) = S \circ T(\vec{u}) + S \circ T(\vec{v})$$
$$S \circ T(\lambda\vec{v}) = S(T(\lambda\vec{v})) = S(\lambda T(\vec{v})) = \lambda S(T(\vec{v})) = \lambda S \circ T(\vec{v}),$$

and so $S \circ T$ is a linear transformation. □

Proof of Theorem 1.6 on page 56.
Suppose A is an $m \times n$ matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}$$

and $\vec{u}$ and $\vec{v}$ are n-dimensional vectors,

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{pmatrix}, \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix},$$

and let λ be a scalar.

Now we simply compute $A(\vec{u} + \vec{v})$ and rearrange the terms to see this equals $A\vec{u} + A\vec{v}$.

$$
\begin{aligned}
&A(\vec{u}+\vec{v})\\
=&\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{pmatrix}\left(\begin{pmatrix}u_1\\u_2\\u_3\\\vdots\\u_n\end{pmatrix}+\begin{pmatrix}v_1\\v_2\\v_3\\\vdots\\v_n\end{pmatrix}\right)\\
=&\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}u_1+v_1\\u_2+v_2\\u_3+v_3\\\vdots\\u_n+v_n\end{pmatrix}\\
=&(u_1+v_1)\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+(u_2+v_2)\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots+(u_n+v_n)\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}\\
=&\begin{pmatrix}u_1a_{11}+v_1a_{11}\\u_1a_{21}+v_1a_{21}\\\vdots\\u_1a_{m1}+v_1a_{m1}\end{pmatrix}+\begin{pmatrix}u_2a_{12}+v_2a_{12}\\u_2a_{22}+v_2a_{22}\\\vdots\\u_2a_{m2}+v_2a_{m2}\end{pmatrix}+\cdots+\begin{pmatrix}u_na_{1n}+v_na_{1n}\\u_na_{2n}+v_na_{2n}\\\vdots\\u_na_{mn}+v_na_{mn}\end{pmatrix}\\
=&\begin{pmatrix}u_1a_{11}\\u_1a_{21}\\\vdots\\u_1a_{m1}\end{pmatrix}+\begin{pmatrix}v_1a_{11}\\v_1a_{21}\\\vdots\\v_1a_{m1}\end{pmatrix}+\begin{pmatrix}u_2a_{12}\\u_2a_{22}\\\vdots\\u_2a_{m2}\end{pmatrix}+\begin{pmatrix}v_2a_{12}\\v_2a_{22}\\\vdots\\v_2a_{m2}\end{pmatrix}+\cdots\begin{pmatrix}u_na_{1n}\\u_na_{2n}\\\vdots\\u_na_{mn}\end{pmatrix}+\begin{pmatrix}v_na_{1n}\\v_na_{2n}\\\vdots\\v_na_{mn}\end{pmatrix}\\
=&u_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+v_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+u_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+v_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots u_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}+v_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}\\
=&u_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+u_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots u_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}+v_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+v_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots v_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}\\
=&\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}u_1\\u_2\\u_3\\\vdots\\u_n\end{pmatrix}+\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}v_1\\v_2\\v_3\\\vdots\\v_n\end{pmatrix}\\
=&A\vec{u}+A\vec{v}
\end{aligned}
$$

We perform a similar set of manipulations to see that $A(\lambda\vec{v}) = \lambda(A\vec{v})$.

$$
\begin{aligned}
A(\lambda\vec{v}) &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \lambda v_3 \\ \vdots \\ \lambda v_n \end{pmatrix} \\
&= \lambda v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \lambda v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots \lambda v_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \\
&= \lambda \left(v_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + v_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots v_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right) \\
&= \lambda(A\vec{v})
\end{aligned}
$$

□

Proof of Theorem *1.7 on page 57*.

Let $\vec{v}$ be an arbitrary n-dimensional vector,

$$
\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix}.
$$

Notice that we may write $\vec{v}$ as follows:

$$\vec{v} = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots v_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Since T is a linear transformation we then have

$$\begin{aligned} T(\vec{v}) &= T\left(v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots v_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right) \\ &= v_1 T \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 T \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_3 T \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots v_n T \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \end{aligned}$$

Notice, however, that this is exactly the same as multiplying the vector $\vec{v}$ by the matrix with columns

$$T \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, T \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, T \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, T \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

$\square$

Proof of Theorem 1.8 on page 70.
We'll write out the proof for properties (i) and (ii) in the case of two-dimensions. The other properties are proven similarly. Suppose $\vec{u} = \langle u_1, u_2 \rangle$, $\vec{v} = \langle v_1, v_2 \rangle$, and $\vec{w} = \langle w_1, w_2 \rangle$.

(i)

$$\begin{aligned}\vec{v} \cdot \vec{v} &= \langle v_1, v_2 \rangle \cdot \langle v_1, v_2 \rangle \\ &= v_1 v_1 + v_2 v_2 \\ &= v_1^2 + v_2^2 \\ &= \left(\sqrt{v_1^2 + v_2^2}\right)^2 \\ &= \|\vec{v}\|^2\end{aligned}$$

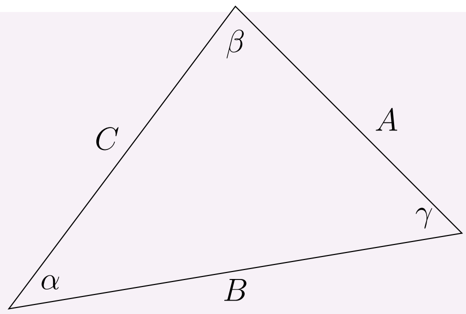
(ii)

$$\begin{aligned}\vec{u} \cdot (\vec{v} + \vec{w}) &= \langle u_1, u_2 \rangle \cdot \langle v_1 + w_1, v_2 + w_2 \rangle \\ &= u_1(v_1 + w_1) + u_2(v_2 + w_2) \\ &= u_1 v_1 + u_1 w_1 + u_2 v_2 + u_2 w_2 \\ &= u_1 v_1 + u_2 v_2 + u_1 w_1 + u_2 w_1 \\ &= \langle u_1, u_2 \rangle \cdot \langle v_1, v_2 \rangle + \langle u_1, u_2 \rangle \cdot \langle w_1, w_2 \rangle \\ &= \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}\end{aligned}$$

□

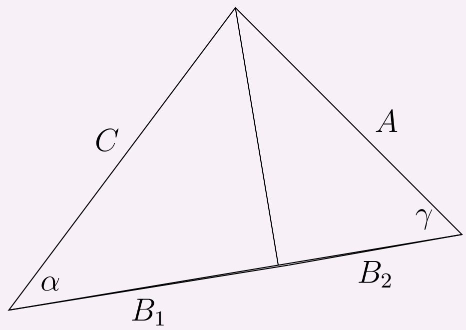
Proofs from Section 1.5, The dot product

Proof of Theorem 1.9 on page 72.
Consider the triangle below,



We will show that $A^2 = B^2 + C^2 - 2BC\cos(\alpha)$. The other two parts of the law of cosines are proven in exactly the same way by simply making the appropriate changes to the construction we are about to describe.

We begin by drawing a line from the vertex with angle β in the triangle above down to the side of length B, such that the the line we draw will intersect the side of our triangle orthogonally.



Now we can compute the values of D and B_1 using some simple trigonometry. In particular, the triangle with sides of length C, D, and B_1 is a right triangle, and so we know know that $\cos(\alpha)$ is equal to the length of the adjacent side (B_1) over the length of the hypotenuse (C), and so

$$\cos(\alpha) = \frac{B_1}{C}$$
$$\implies B_1 = C\cos(\alpha).$$

Similarly, we know that $\sin(\alpha)$ equals ${}^{D}/{}_{C}$, and so

$$D = C\sin(\alpha).$$

Since $B_1 + B_2 = B$, we can compute

$$B_2 = B - B_1 = B - C\cos(\alpha).$$

Now applying the Pythagorean theorem to the right triangle which has A as its hypotenuse and B_2 and D as its other sides we have

$$\begin{aligned} A^2 &= D^2 + B_2^2 \\ &= C^2\sin^2(\alpha) + (B - C\cos(\alpha))^2 \\ &= C^2\sin^2(\alpha) + B^2 - 2BC\cos(\alpha) + C^2\cos^2(\alpha) \\ &= B^2 + C^2(\sin^2(\alpha) + \cos^2(\alpha)) - 2BC\cos(\alpha) \\ &= B^2 + C^2 - 2BC\cos(\alpha). \end{aligned}$$

□

Proof of Theorem 1.10 on page 74.
The proof of this theorem is given in the main part of the text, just above where Theorem 1.10 appears on page 74. □

Proofs from Section 1.6, The cross product

Proof of Theorem 1.11 on page 82.
Consider the square of $\|\vec{u} \times \vec{v}\|$:

$$\begin{aligned} \|\vec{u} \times \vec{v}\|^2 =& (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 \\ =& (u_2v_3)^2 - 2u_2u_3v_2v_3 + (u_3v_2)^2 + (u_3v_1)^2 - 2u_1u_3v_1v_3 + (u_1v_3)^2 + \\ & (u_1v_2)^2 - 2u_1u_2v_1v_2 + (u_1v_2)^2 \\ =& \|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2 \\ =& \|\vec{u}\|^2\|\vec{v}\|^2 - \|\vec{u}\|^2\|\vec{v}\|^2\cos^2\theta \\ =& \|\vec{u}\|^2\|\vec{v}\|^2(1 - \cos^2\theta) \\ =& \|\vec{u}\|^2\|\vec{v}\|^2\sin^2\theta \end{aligned}$$

Taking square roots of both sides gives us the result. □

Proof of Theorem *1.12 on page 83*.

(i) We can check for orthogonality by taking dot products. Using our formula above we know

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

Thus

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

The same thing happens if we look at $(\vec{u} \times \vec{v}) \cdot \vec{v}$.

(ii) This is clear if our vectors, $\vec{u}$ and $\vec{v}$, lie in the xy-plane. But given any two vectors, we could always rotate things so that our vectors lie in the xy-plane, and rotating clearly doesn't change the area of the parallelogram with sides $\vec{u}$ and $\vec{v}$.

□

Proof of Theorem *1.13 on page 84*.
As with all of the vector properties we've discussed so far, these can be shown by writing the vectors out in coordinates and then wading through the algebra. We show property (iii) below.

Let $\vec{u} = \langle u_1, u_2, u_3 \rangle$, $\vec{v} = \langle v_1, v_2, v_3 \rangle$, $\vec{w} = \langle w_1, w_2, w_3 \rangle$. We then have

$$\begin{aligned}
\vec{u} \times (\vec{v} + \vec{w}) &= \langle u_1, u_2, u_3 \rangle \times (\langle v_1 + w_1, v_2 + w_2, v_3 + w_3 \rangle) \\
&= \det \begin{pmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 + w_1 & v_2 + w_2 & v_3 + w_3 \end{pmatrix} \\
&= \vec{\imath} u_2(v_3 + w_3) + \vec{\jmath} u_3(v_1 + w_1) + \vec{k} u_1(v_2 + w_2) \\
&\quad - \vec{\imath} u_3(v_2 + w_2) - \vec{\jmath} u_1(v_3 + w_2) - \vec{k} u_2(v_1 + w_1) \\
&= \langle u_2v_3 + u_2w_3 - u_3v_2 - u_3w_2, \\
&\quad u_3v_1 + u_3w_1 - u_1v_3 - u_1w_2, \\
&\quad u_1v_2 + u_1w_2 - u_2v_1 - u_2w_1 \rangle \\
&= \langle u_2v_3 - u_3v_2 + u_2w_3 - u_3w_2, \\
&\quad u_3v_1 - u_1v_3 + u_3w_1 - u_1w_2, \\
&\quad u_1v_2 - u_2v_1 + u_1w_2 - u_2w_1 \rangle \\
&= \langle u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1 \rangle + \\
&\quad \langle u_2w_3 - u_3w_2, u_3w_1 - u_1w_2, u_1w_2 - u_2w_1 \rangle \\
&= \left(\vec{\imath} u_2v_3 - \vec{\imath} u_3v_2 + \vec{\jmath} u_3v_1 - \vec{\jmath} u_1v_3 + \vec{k} u_1v_2 - \vec{k} u_2v_1 \right) + \\
&\quad \left(\vec{\imath} u_2w_3 - \vec{\imath} u_3w_2 + \vec{\jmath} u_3w_1 - \vec{\jmath} u_1w_3 + \vec{k} u_1w_2 - \vec{k} u_2w_1 \right) \\
&= \left(\vec{\imath} u_2v_3 + \vec{\jmath} u_3v_1 + \vec{k} u_1v_2 - \vec{\imath} u_3v_2 - \vec{\jmath} u_1v_3 - \vec{k} u_2v_1 \right) + \\
&\quad \left(\vec{\imath} u_2w_3 + \vec{\jmath} u_3w_1 + \vec{k} u_1w_2 - \vec{\imath} u_3w_2 - \vec{\jmath} u_1w_3 - \vec{k} u_2w_1 \right) \\
&= \det \begin{pmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} + \det \begin{pmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \\
&= \vec{u} \times \vec{v} + \vec{u} \times \vec{w}
\end{aligned}$$

Proofs of the other properties follow similarly. □

B

Sets

> *A set is a Many that allows itself to be thought of as a One.*
>
> GEORG CANTOR

"Sets" are abstract tools used for collecting information, and provide a sort of universal language for expressing mathematics[1]. Even if you've never seen or heard of sets before, they've been lurking in the background of most of the mathematics you've done in other courses. It can be convenient sometimes to know a little bit of the basic language of sets, as this will allow us to easily and compactly describe certain mathematical objects that we'll encounter.

The goal of this appendix is to introduce the basic ideas, language, and notation of the theory of sets. Some of this may seem strange and abstract if you've never seen it before, so lots of examples are included to hopefully elucidate anything that initially seems unintuitive. The main thing you should take away from this appendix is to have a general idea of what a set is; to learn some of the common notation; and obtain a basic understanding how certain simple geometric objects (such as lines and circles) can be represented as sets.

Remark.
This appendix on sets is included entirely for the sake of completeness and to sate the curiosity if any student that wants to know about what this word "set" we use from time to time means. This notion is a bit abstract and you can safely avoid reading this appendix if you want to.

[1] This idea of sets being a "universal language" for mathematics isn't *really* correct, but it's a convenient way of thinking about sets. "Most" mathematicians most of the time think of most mathematics as being described in terms of sets, but there are exceptions.

B.1 Introduction

Throughout this course we will be concerned with building a collection of tools that we can use to do calculus in several dimensions. You already know how to do calculus in one dimension, and calculus in several dimensions is a natural extension, but with several caveats. In order to deal with the niceties of multiple dimensions, we'll need to learn some general background.

In this appendix, the background we'll be concerned with the theory of sets. Sets are simple, though abstract, objects that let us group mathematical entities together; set theory forms one of the basic building blocks of modern mathematics by giving us a primitive set of basic operations that apply to (almost) all areas of math. For this reason set theory is worth learning if you have any interest in pursuing deeper mathematics or science.

What's the point of this?

For our purposes, sets provide a common and convenient language that we can use to describe some of the higher-dimensional objects that we'll care about later in the course. Using sets we can make vague notions precise, and so we want to invest some time learning the basics. Some of this material may seem esoteric or overly pedantic at first glance, but the basic ideas will be important for us later in this course, and later in your mathematical careers.

B.2 Basic Ideas

A *set* is an unordered collection of "things," and these "things" are called the *elements* of the set. When there are only a few elements in a set, we'll usually list all of the elements in the set inside of a pair of curly braces (the symbols $\{$ and $\}$), with the elements separated by commas. This definition may seem a bit vague, but a few examples should clear things up.

Example B.1.
Below is a set containing three elements, the numbers two, four, and

six:

$$\{2, 4, 6\}.$$

As stated above, a set is an *unordered* collection of things, so the set above is the same as the set $\{4, 2, 6\}$ or $\{6, 4, 2\}$. The thing that really matters about sets is what's inside of them, and this is why we don't care about the order. (Think of a set as a box containing some objects you own. You typically don't care about where things sit inside of the box, but you do care about whether something is in the box or not.)

The sets we will consider will typically have very many elements, and it's inconvenient to write down everything in the set each time we refer to that set. For this reason we'll usually give sets a name – just an easier way of referring to the set. To give the set $\{2, 4, 6\}$ above the name A, we just write $A = \{2, 4, 6\}$. This way whenever we want to refer to the set we can just write A instead of $\{2, 4, 6\}$.

We use the symbol $\in$ to denote that something is in a set, and the symbol $\notin$ to denote that something is not in a set. For example, if we write $2 \in A$, this means that 2 is an element of the set A. Writing $3 \notin A$, however, means that 3 is not an element of the set A.

If we have two sets, say A and B, and if every element of A is contained in B, then we say that A is a *subset* of B and write $A \subseteq B$.

Example B.2.
Consider the set $A = \{2, 4, 6\}$ from before and the set $B = \{-1, 2, 3, 4, 6, 17\}$. Notice that every element of A is contained in B. So A is a subset of B; notationally, $A \subseteq B$.

Remark.
In the few examples above, we considered sets of numbers, but the definition of a set is simply an unordered collection of "things." These "things" do not have to be numbers, and can in fact be essentially

anything. For example, you could have a set of words,

$$\{\text{mathematics, football, beer, jabberwocky, penguin}\};$$

or a set of symbols,

$$\{\heartsuit, ☺, \flat, ©\};$$

or even a set of sets,

$$\{\{1, 2, 3\}, \{4\}, \{5, 6\}\}.$$

For our purposes sets will usually consist of numbers, points in space, or vectors (which we'll define in a future lecture).

Cardinality

The number of distinct elements in a set is called the sets *cardinality*. This is just a fancy word that means how many things are in the set. The set $A = \{-13, 4, \pi\}$, for example, has cardinality three because there are three things in A. Symbolically, we sometimes denote the cardinality of a set A by $\#A$. In the case of $A = \{-13, 4, \pi\}$, we have $\#A = 3$.

One point here that may seem a little bit weird is that we used the word *distinct* when defining cardinality. To make this clear, consider the set $B = \{2, 2, 2\}$. What's the cardinality of B? Well, there's only one *distinct* element in B, so $\#B = 1$. Because we only care about distinct elements, we generally think of elements in sets as not being repeat. That is, $B = \{2, 2, 2\}$ is really just the set that contains only the element 2 – we don't care how many times 2 is written down, just whether 2 is in B or not. For this reason, we consider $B = \{2, 2, 2\}$ and $C = \{2\}$ to be the same set. This may take some getting used to, but just remember we only care about whether something is in a set or not – we don't care about anything else, and in particular we don't care about how many "times" something appears in a set. It appears or it doesn't, that's all we care about, so we think of $\{2, 2, 2\}$ as the same as $\{2\}$ as the same as $\{2, 2, 2, 2, 2, 2, 2, 2, 2, 2\}$.

It is possible for a set to contain nothing at all: this is called the empty set and denoted $\emptyset$ or $\{\,\}$. This is the unique set that doesn't have anything in it. Thinking of sets as boxes containing "things," the empty set corresponds to a an empty box.

Remark.
There are some funny things about the cardinality that we're not mentioning here. In particular, sets that have infinitely-many elements can have different cardinalities; there are different "sizes" of infinity. Furthermore, some collections of things are actually "too large" to have a notion of cardinality, even infinite cardinality – some things are "too big" to be sets. While those ideas are very interesting, making sense of them could easily take up an entire course, so we won't say much about them except to mention that there is a big, interesting mathematical world out there that most people never have an opportunity to learn about. If you're curious, though, you should always feel free to ask any mathematics professor!

B.3 Unions and Intersections

Given two sets, A and B, there are several ways we can combine these sets to produce a third set. Here we'll discuss three of them: unions, intersections, and Cartesian products. Unions and intersections are straightforward so we'll discuss those now, and then discuss Cartesian products later.

The union of two sets

If we're given two sets, say A and B, their *union* is yet another set which contains all of the elements of A, all of the elements of B, and nothing else. This union is denoted $A \cup B$.

Example B.3.
Suppose $A = \{2, 4, 6\}$ and $B = \{1, 3, 5\}$. Then their union, $A \cup B$, is the set which contains everything in A and everything in B, but nothing else. That is,

$$A \cup B = \{1, 2, 3, 4, 5, 6\}.$$

Example B.4.
Suppose now that $A = \{1, 2, 3\}$, and $B = \{1, 3, 7\}$. The union of A and B is then

$$A \cup B = \{1, 2, 3, 7\}.$$

Notice that every element of A is an element of $A \cup B$, and every element of B is an element of $A \cup B$. The fact that some elements (like 1 and 3) appear in both sets doesn't change the union.

Example B.5.
Suppose that $A = \{-1, \pi, 4\}$ and $B = \{-1, 0, 1, 3, \pi, 4\}$. (Notice that $A \subseteq B$.) What is $A \cup B$ in this case? Since everything in A is already an element of B, the union $A \cup B$ is just B: intuitively, A doesn't add anything to the union because B already contains everything.

This is a general fact: if $A \subseteq B$, then $A \cup B = B$.

Example B.6.
As one last example of the union of two sets, let's suppose that $A = \emptyset$ and B is any other set. What is the union, $A \cup B$? Here, just as in the last example, A doesn't contribute anything to the union (since A doesn't have anything to contribute), so $A \cup B = B$ yet again.

If you think about the union of two sets as being something like the sum of two numbers, unioning with the empty set is like adding zero: it doesn't do anything.

The intersection of two sets

If A and B are two sets, their *intersection* is the set of elements that are in both A and B. The intersection is denoted $A \cap B$.

Example B.7.
Suppose $A = \{☺, ☹, ♫, ♪\}$ and $B = \{♂, ♀, ☺, \Upsilon, ♪\}$. Then the intersection of A and B is

$$A \cap B = \{☺, ♪\}.$$

Example B.8.
Recall Example B.5 above where $A = \{-1, \pi, 4\}$ and $B = \{-1, 0, 1, 3, \pi, 4\}$. In this case $A \cap B$ contains everything that's in both A and B. Since everything that's an element of A is already an element of B, $A \cap B = A$.

In general if $A \subseteq B$, then $A \cap B = A$.

Example B.9.
Let's consider an example where A and B have nothing in common. Say $A = \{1, 3, 5\}$ and $B = \{2, 4, 6\}$. Remember that $A \cap B$ is supposed to be the set of everything that's in *both* A and B. In this example there's nothing that's an element of both A and B. This means $A \cap B$ contains nothing: $A \cap B$ is the empty set, $A \cap B = \emptyset$.

Example B.10.
Extending the last example, let's suppose that $A = \emptyset$ and B is any other set. What should $A \cap B$ be? Since A doesn't have any elements, it can't have any elements in common with B (regardless of what B is). Just like in the last example, this means $A \cap B = \emptyset$.

We saw in the previous section that $\emptyset \cup B = B$ and said that if you think of unions as being analogous to addition, unioning with the empty set was like adding zero. Similarly, $\emptyset \cap B = \emptyset$. If you thought about intersections as being analogous to multiplication, intersecting with the empty set is like multiplying by zero: it always gives you back the empty set.

(This notion of unions being similar to addition, and intersections being similarly to multiplication, isn't a mere coincidence. You can actually devise a method of "encoding" numbers as sets so that the usual arithmetic operations become operations on sets. This is an interesting thing to think about, but it would take us too far afield to discuss.)

Unions and intersections of more than two sets

Above we described unions and intersections of two sets, but we can actually define unions and intersections for lots of sets. That is, if we had three sets we could talk the union of all three, $A \cup B \cup C$, or the intersection of all three $A \cap B \cap C$. We could also do this for four sets, or five sets, or ten billion sets, and the idea is exactly the same.

To make this precise, let's say that we have n sets where n is any positive whole number. Let's suppose $n \geq 2$ just to make things easy. Suppose our n sets are listed as A_1, A_2, A_3, ..., A_{n-1}, A_n. Then we'll define the union

$$A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_{n-1} \cup A_n$$

as the set containing every element that's in any of the A_k sets.

Example B.11.
Let's consider a simple example where we take the union of the following four sets:

$$\begin{aligned} A_1 &= \{2, 4, 6, 8, 10\} \\ A_2 &= \{3, 6, 9, 12, 15, 18, 21\} \\ A_3 &= \{5, 10, 15, 20\} \\ A_4 &= \{7, 14, 21\} \end{aligned}$$

Then the union of the four sets is

$$A_1 \cup A_2 \cup A_3 \cup A_4 = \{2, 3, 4, 6, 8, 9, 10, 12, 14, 15, 18, 20, 21\}.$$

We can actually describe this union of lots of sets by taking unions of pairs of sets two at a time. For example, the union $A \cup B \cup C$ could be written as $(A \cup B) \cup C$. Just as in the case of arithmetic, the parentheses have the highest possible precedence. That is, you always perform the operations in parentheses first.

Example B.12.
Suppose $A = \{\alpha, \beta, \gamma\}$, $B = \{\beta, \kappa, \varphi\}$, and $C = \{\gamma, \delta, \varphi\}$. Then to calculate $A \cup B \cup C$, we can just calculate $(A \cup B) \cup C$. We simply do the things in parentheses first to generate a set $A \cup B$, and then union this set with C.

$$\begin{aligned}(A \cup B) \cup C =& (\{\alpha, \beta, \gamma\} \cup \{\beta, \kappa, \varphi\}) \cup \{\gamma, \delta, \varphi\} \\ =& \{\alpha, \beta, \gamma, \kappa, \varphi\} \cup \{\gamma, \delta, \varphi\} \\ =& \{\alpha, \beta, \gamma, \delta, \kappa, \varphi\}\end{aligned}$$

Conveniently, the order we perform unions in doesn't matter. That is, $(A \cup B) \cup C = A \cup (B \cup C)$.

We define intersections of multiple sets in a similar way. If we have $n \geq 2$ sets, labelled A_1, A_2, ..., A_n, then we define their intersection,

$$A_1 \cap A_2 \cap \cdots \cap A_n,$$

as the set containing the elements that are in *each and every one* of the A_k's.

Example B.13.
Consider the three sets below:

$$\begin{aligned}A_1 =& \{1, 2, 3, 4, 5\} \\ A_2 =& \{2, 4, 6, 8, 10\} \\ A_3 =& \{4, 8, 12, 16, 20\}\end{aligned}$$

The intersection $A_1 \cap A_2 \cap A_3$ consists of those elements that are in every set A_1, A_2, and A_3. By inspection, we see that the only element that is in all three sets is 4. That is,

$$A_1 \cap A_2 \cap A_3 = \{4\}.$$

Just as for unions, we can use parentheses to intersect pairs of sets two at a time.

Example B.14.
Consider A_1, A_2, and A_3 as in the last example. We can calculate the intersection $A_1 \cap A_2 \cap A_3$ as follows:

$$\begin{aligned}(A_1 \cap A_2) \cap A_3 =& (\{1,2,3,4,5\} \cap \{2,4,6,8,10\}) \cap \{4,8,12,16,20\} \\ =& \{2,4\} \cap \{4,8,12,16,20\} \\ =& \{4\}\end{aligned}$$

We can also consider unions and intersections of infinitely-many sets. We will not have a need to do that in this course, but it is an operation we can define. Intuitively, the intersection of infinitely-many sets is the set of elements in every one of the sets we are intersecting; the union of infinitely-many sets is the set of elements that are in at least one of the sets we are unioning.

B.4 Cartesian Products

One construction that's particular important for us is the idea of a *Cartesian product*. This is another way to take two sets and combine them get a third set, but Cartesian products are a little bit different from unions and intersections above. What we'll do is take two sets, A and B, and consider the set of all *pairs* (a, b) where $a \in A$ and $b \in B$. This is called the Cartesian product of A and B and is denoted $A \times B$.

Example B.15.
As a simple example, let's suppose that $A = \{1, 2, 3\}$ and $B = \{w, x, y, z\}$. Then the Cartesian product $A \times B$ is the set of all pairs (a, b) where

a is one of 1, 2, or 3, and b is one of w, x, y, or z. That is,

$$A \times B = \{(1,w),(1,x),(1,y),(1,z),\\ (2,w),(2,x),(2,y),(2,z),\\ (3,w),(3,x),(3,y),(3,z)\}$$

This idea is important for us because can use Cartesian products to combine lower-dimensional spaces together to build higher-dimensional spaces.

We can in fact take Cartesian products of more than two sets, but to do this we need to talk about *tuples*. For a whole number $n \geq 2$, we define an *n-tuple* to be a comma separated list of n items. For example, a 2-tuple is just a pair: it's just a list of two elements such as (x, y) or $(1, -3)$. A 3-tuple is sometimes called a triple; examples are (a, b, c) or $(-3, 12, 0)$. A 4-tuple is a list of four items: e.g., $(\alpha, \beta, \gamma, \delta)$ or $(0, 1, -\pi, \sin(e))$.

Remember that *order matters* for tuples, and does not matter for sets. For example, the following sets are equal:

$$\{4, {}^1\!/\!_2, 2\} = \{{}^1\!/\!_2, 2, 4\},$$

but the following 3-tuples are not equal:

$$(4, {}^1\!/\!_2, 2) \neq ({}^1\!/\!_2, 2, 4).$$

Now, if we have n sets (where $n \geq 2$ is a whole number) labelled A_1, A_2, ..., A_n, then we can form their Cartesian product

$$A_1 \times A_2 \times \cdots \times A_n$$

as the collection of all n-tuples of the form $(a_1, a_2, ..., a_n)$ where $a_1 \in A_1$, $a_2 \in A_2$, ..., and $a_n \in A_n$.

Example B.16.
Consider the four sets below:

$$A_1 = \{a, b, c\}\\ A_2 = \{1\}\\ A_3 = \{0, 1\}\\ A_4 = \{♂, ♀\}.$$

The Cartesian product of these three sets is the collection of all 4-tuples of the form (a_1, a_2, a_3, a_4) where a_1 is one of a, b, or c; where a_2 is 1; where a_3 is one of 0 or 1; and where a_4 is either ♂ or ♀. Thus,

$$\begin{aligned} A_1 \times A_2 \times A_3 \times A_4 = \{&(a,1,0,♂),(a,1,0,♀),(a,1,1,♂),(a,1,1,♀) \\ &(b,1,0,♂),(b,1,0,♀),(b,1,1,♂),(b,1,1,♀) \\ &(c,1,0,♂),(c,1,0,♀),(c,1,1,♂),(c,1,1,♀)\} \end{aligned}$$

Example B.17.
We can also take the Cartesian product of a set with itself. For example, if $A = \{x, y, z\}$, then $A \times A$ is the set of all pairs (a, b) where a is one of x, y, or z; and b is also one of x, y, or z:

$$A \times A = \{(x,x),(x,y),(x,z),(y,x),(y,y),(y,z),(z,x),(z,y),(z,z)\}.$$

We could similarly consider $A \times A \times A$, which is the set of all triples (a, b, c) where each of a, b, and c is one of x, y, or z:

$$\begin{aligned} A \times A \times A = \{&(x,x,x),(x,x,y),(x,x,z), \\ &(x,y,x),(x,y,y),(x,y,z), \\ &\vdots \\ &(z,y,x),(z,y,y),(z,y,z), \\ &(z,z,x),(z,z,y),(z,z,z)\}\} \end{aligned}$$

It gets tedious writing out $A \times A$ or $A \times A \times A$ all the time, so we will write A^2 to mean $A \times A$; A^3 denotes $A \times A \times A$; A^4 denotes $A \times A \times A \times A$; and so forth.

B.5 Standard Notations

In this section we describe some standard notation regarding sets. Many times we'll be concerned with sets that contain infinitely-many elements. This presents a little bit of a problem for us in that we obviously can't write out all of the elements in such a set. One convention is to use an ellipsis (that is, the "dot dot dot" symbol: ...) after writing a few elements of the set to mean "continue the pattern."

Example B.18.
Consider the set

$$E = \{2, 4, 6, 8, 10, 12, ...\}.$$

This is the set of all positive even numbers.

Example B.19.
Consider the set

$$P = \{3, 3.1, 3.14, 3.141, 3.1415, 3.14159, ...\}$$

This is a set of numbers approximating π, where the pattern indicated by the "..." means that we continue to append digits of π.

Of course, the "..." notation can be a little bit inconvenient sometimes because it assumes the reader will pick up on the pattern you have in mind. For example, what does the "..." refer to below?

$$\{2, 4, ...\}$$

Does this mean the even numbers, $\{2, 4, 6, 8, 10, ...\}$, or does this mean the powers of two, $\{2, 4, 8, 16, 32, ...\}$?

To avoid ambiguity, we sometimes use an alternative notation called *set builder notation* which gives us a more explicit pattern for determining what's an element of the set. In set builder notation we separate the interior of the curly braces by a vertical bar, $|$. On the left-hand side of the vertical bar we say what the elements of the set look like, and on the right-hand side of the vertical bar we give a condition all of the elements of the set must satisfy. This likely sounds very unclear, but a few examples should help.

Example B.20.
One way to specify the set of all even integers, using set-builder notation, would be to write

$$E = \{2n \mid n \text{ is a positive whole number}\}.$$

The $2n$ on the left-hand side tells us that everything in the set looks like "two times a number n." The right-hand side tells us that the n has to be a positive whole number.

Example B.21.
To specify the set of all the powers of 2, we could use set-builder notation to write

$$T = \{2^n \mid n \text{ is a positive whole number}\}.$$

Example B.22.
Consider all of the points on the line $y = 2x + 3$. This collection of points forms a set. In set-builder notation we could specify this as

$$\{(x, y) | y = 2x + 3\}.$$

That is, we want all of the pairs (x, y) where y is $2x + 3$.

Example B.23.
The set of points on a circle of radius one, centered at the origin, in the plane is specified in set-builder notation as

$$\{(x, y) | x^2 + y^2 = 1\}.$$

This means that we want all of the pairs (x, y) where x and y are related by the equation $x^2 + y^2 = 1$.

Common Sets

Some sets of numbers are so common that mathematicians have agreed on special notations to identify those sets. Here we describe some of these common sets. These sets are usually written in one of two ways. Years ago these sets were identified by a single bold letter in textbooks. However, bold letters are difficult to write on paper or a blackboard. For this reason, people started doubling some of the lines in the letter when writing it on paper or a blackboard. This was used to denote that the letter should be bold. Over time this practice became more and more common, and now many textbooks (but not all) use a special typeface called *blackboard bold* which resembles these doubly lined letters.

Natural numbers, $\mathbb{N}$

The *natural numbers* refer to the positive whole numbers and are denoted by a capital letter N in a blackboard-bold font: $\mathbb{N}$.

$$\mathbb{N} = \{1, 2, 3, 4, 5, ...\}.$$

These are called the natural numbers (or sometimes the *counting numbers*) because they are the numbers that naturally appear when one starts to consider mathematics for the first time: they represent physical, countable quantities.

Integers, $\mathbb{Z}$

The *integers* are simply the natural numbers, plus zero and the negatives, and are denoted by a capital Z in a blackboard-bold font: $\mathbb{Z}$.

$$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}.$$

The reason for the letter *Z* has to do with the many influential German mathematicians, and the fact that *Zahlen* is the German word for *numbers*.

Rational numbers, $\mathbb{Q}$

The *rational numbers* are fractions of integers and are denoted by a capital Q in a blackboard-bold font: $\mathbb{Q}$.

$$\mathbb{Q} = \left\{ \frac{p}{q} \,|\, p, q \in \mathbb{Z}, \text{ and } q \neq 0 \right\}.$$

The reason these numbers are called rational is because they represent *ratios*. The reason for the letter Q is that *quotient* is synonymous with ratio.

Real numbers, $\mathbb{R}$

The real numbers are the numbers we'll typically care about in this class. Giving a rigorous definition of the real numbers is a bit tedious, but these numbers just represent points on a number line. Notice that there are real numbers which are not rational numbers, such as $\sqrt{2}$ or π. (That is, you can not write $\sqrt{2}$ or π as a fraction of two integers.)

The real numbers are represented by $\mathbb{R}$.

B.6 Maps Between Sets

Given two sets, A and B, a *map* from A to B is a way to taking an element of A and associating an element of B to it. You're probably already familiar with maps, but called them *functions* in your previous classes. (The two words, *map* and *function*, and synonymous and so are used interchangeably.)

Example B.24.
Suppose both A and B are the integers: $A = \mathbb{Z}$ and $B = \mathbb{Z}$. One example of a map from A to B is the function that takes a number and squares it. If we call this function f, then we might write $f(x) = x^2$. So, if you give an element of A (say $x = -3$), then I associate to it an element of B (in this case, $f(-3) = 9$).

As the above example indicates, all a map does is take an element in one set, and turns it into an element of another set. There is only one stipulation here: you can't turn the same element of A into two different elements of B. For example, you can not have a map f from A to B which says for some $a \in A$ that $f(a) = b$ and $f(a) = c$. Other than this one rule, maps can do any weird and arbitrary thing you want.

To denote that the inputs of a map f live in a set A, and the outputs live in a set B, we write $f : A \to B$. This means "f is a function which takes elements in A, and turns them into elements of B." We already saw one example of a map $f : \mathbb{Z} \to \mathbb{Z}$ above, $f(x) = x^2$.

Given a function $f : A \to B$, we say that A is the *domain* of f, and that B is the *codomain* of f. The *range* of f is the set of all elements in B (all elements in the codomain) that an element of A can actually be turned into.

Example B.25.
Consider again the function $f : \mathbb{Z} \to \mathbb{Z}$ given by $f(x) = x^2$. The domain of this function is $\mathbb{Z}$, the codomain is also $\mathbb{Z}$, but what is the range? The range is not $\mathbb{Z}$ because we can never square an integer to get a negative integer: anything we square is positive or zero. This means the range of the function is actually $\mathbb{N} \cup \{0\}$ – the natural numbers together with zero.

One notation you might encounter is $x \mapsto f(x)$. This is pronounced "x maps to $f(x)$." For example, the function $f(x) = x^2$ could also be expressed as $x \mapsto x^2$. This just means that if you take an element x in the domain of the function, you get an element in the range by squaring x. (There isn't any new idea here, it's just an alternative notation you should be aware of.)

Graphs of Maps

Given a map $f : A \to B$, the *graph* of f is a subset of the Cartesian product, $A \times B$, which represents the function. Precisely, the graph of f, which we'll denote by $\Gamma(f)$, is the set of pairs (a, b) where a is an element of A, and $b = f(a)$. In set-builder notation,

$$\Gamma(f) = \left\{(a, b) \,\middle|\, a \in A,\; b = f(a)\right\}.$$

Again, note that $\Gamma(f) \subseteq A \times B$.

Example B.26.
Let's suppose $A = \{0, 1, 2, 3\}$ and $B = \{0, 1, 2, ..., 9, 10\}$. Consider the map $f : A \to B$ given by $x \mapsto 2x + 3$. (Written another way, our

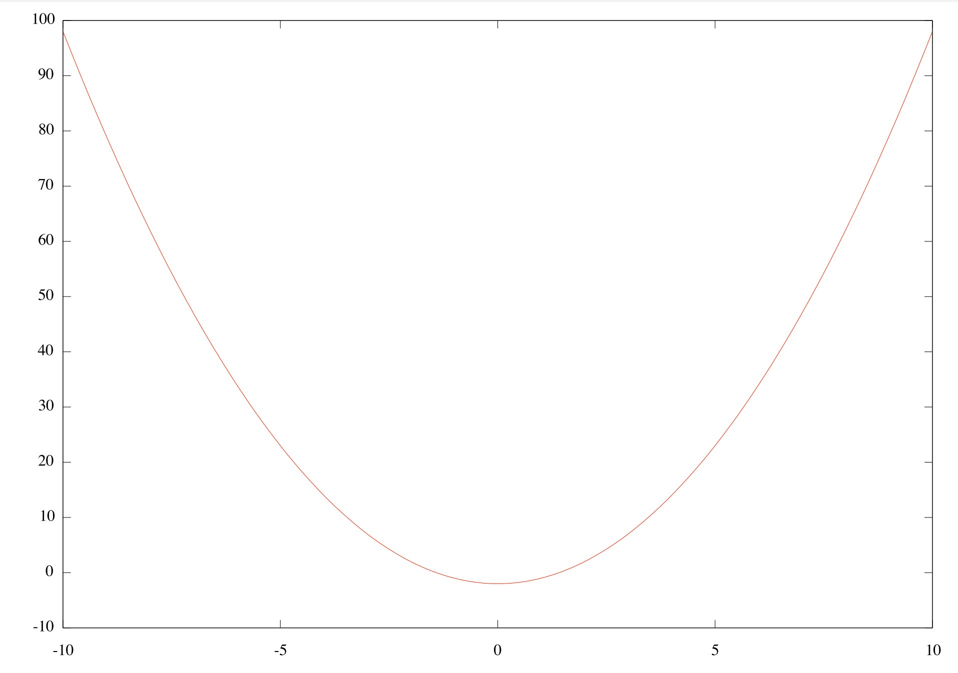
map is $f(x) = 2x + 3$.) The graph of this map is

$$\Gamma(f) = \{(0,3),\ (1,5),\ (2,7),\ (3,9)\}\,.$$

Example B.27.
Of course, a "graph" as defined above is just an abstract version of the graphs you've seen throughout school. For example, the graph of $f(x) = x^2 - 2$ is really means the set

$$\Gamma(f) = \left\{(x,y)\,|\,y = x^2 - 2\right\}$$

which is the set of points marked in figure below.



Again, a "graph" is just an extension of what you think the graph of a function should be. The reason we're introducing a graph in this more abstract way is because later we'll consider graphs of multivariable functions, and it'll be convenient if we can write those graphs down in this more general, set-theoretic language.

B.7 Summary

The idea of a set is ubiquitous in mathematics: even if you've never seen them before, sets have been behind the scenes most of the math you've learned throughout school. The reason is that sets are simple tool for organizing information, and many "obvious" mathematical operations can be rephrased in the language of sets. Having a nice universal language to describe all of mathematics (which is essentially the goal of set theory) is very powerful.

In this appendix we had a very quick introduction to set theory where we learned what a set was, saw some standard notation, learned various operations that could be performed on sets (unions, intersections, and Cartesian products). These ideas will play a role in the calculus that we learn later in the semester, but the main advantage for us learning set theory right now is that it gives us a convenient language for describing certain ideas in mathematics.