Linear Algebra and Multivariable Calculus

Notes from MIT's 18.02 course in fall 2024

Evan Chen

28 April 2025

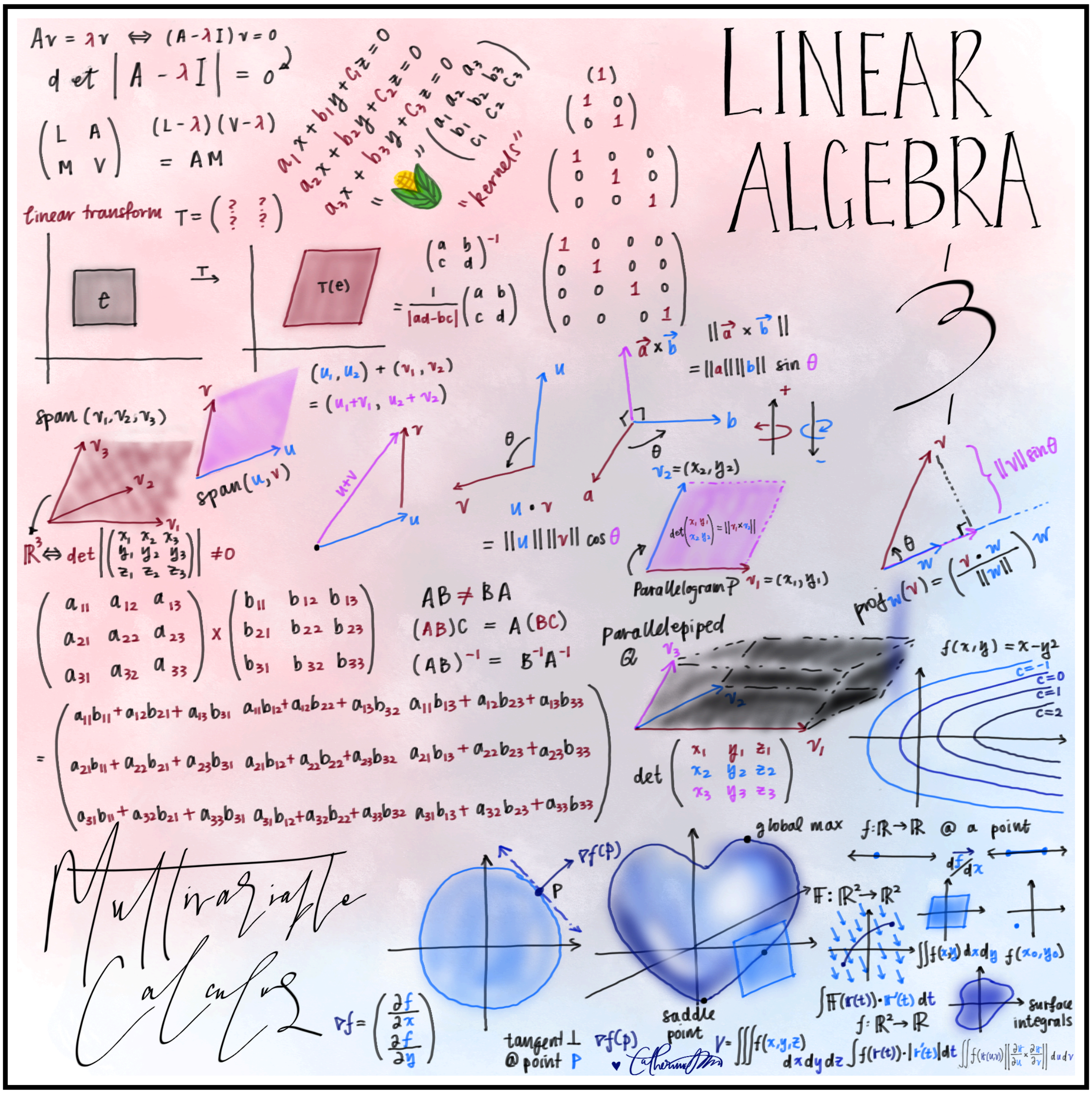

Table of contents

Chapter 1. Preface

At MIT, the course 18.02 (multivariable calculus) is a general institute requirement (GIR); every student must pass this class in order to graduate. These are lecture notes based upon the fall 2024 instance of the course, taught by Davesh Maulik.

§1.1 [TEXT] Goals of this book

These notes have the following lofty goal:

> **Goal**
>
> In theory, an incoming MIT student with a single-variable calculus background should be able to pass the 18.02 final exam by **only** reading these notes and problems, working through several practice final exams, and going to a weekly office-hours[1] to ask questions to a real human.

This is ambitious, and your mileage may vary. Just to be clear, this text is unofficial material and there is no warranty or promise. (Also, if you are actually an MIT student, bear in mind the content of the course will vary by instructor.) But with this goal in mind, here are some parts of the design philosophy of this book.

- **It's practical**. It sticks to the basics and emphasizes giving straight cookbook-like answers to common exam questions.
  - I better say something about memorizing recipes. In principle, if you have perfect memory, you could potentially get a passing score (but not a perfect score) on the final exam by *only* memorizing the recipes.

    I don't recommend this approach; even a vague conceptual understanding of a recipe is at minimum quite helpful for remembering said recipe. But it may be useful to know in principle that the recipe is all you need, and conversely, that you should have the recipes down by heart.
- **It's concrete**. We only work in $\mathbb{R}^n$, and not a generic vector space. We don't use anywhere near the level of abstraction as, say, the Napkin[2]. We don't assume proof experience.
- **It writes things out and has diagrams**. Many lecture notes were meant to go with a in-person lecture rather than replace it. These notes should stand alone.
  - Any sentence that would normally be said out loud is written as text.
  - Any figure that would normally be drawn on the blackboard is actually typeset into the book.
- **It has full solutions to most of its exercises**. I really believe in writing things out. I'd rather have a small number of exercises with properly documented solutions than an enormous pile of mass-produced questions with no corresponding solutions.
- **It tries to explain where formulas come from**. For example, these notes spell out how matrix multiplication corresponds to function composition (in Section 7.3), something that isn't clearly stated in many places. I believe that seeing this context makes it easier to internalize the material.
- **It marks more complicated explanations as "not for exam"**. I hope the digressions are interesting to you (or I wouldn't have written them). But I want to draw a clear boundary between "this explanation is meant for your curiosity or to show where this formula comes from" compared to "this is something you should know by heart to answer exam questions".

[1] You can substitute the office hours for a knowledgeable friend, or similar. The point is that you should have at least some access to live Q/A.

[2] That's the one at `https://web.evanchen.cc/napkin.html`, which *does* assume a proof-based background.

There are two kinds of ways we mark things as not for exam:

- Anything in a gray digression box is not for exam.

> **Digression**
>
> Here's an example of a digression box.

- Anything in an entire section marked **[SIDENOTE]** is not for exam.

- **It's written by Evan Chen**. That's either really good or really bad, depending on your tastes. If you've ever seen me teach a class in person, you know what I mean.

§1.2 [TEXT] Prerequisites

As far as prerequisites go, this text assumes a working knowledge of pre-calculus and calculus as taught in United States high schools.

- **Algebra**: You should be able to work with elementary algebra, so that the following statements make sense

$$x^2 - 7x + 12 = (x-3)(x-4) = 0 \implies x = 3 \quad \text{or} \quad x = 4.$$

  You should also be able to solve two-variable systems of equations, such as

$$\begin{cases} 5x - 2y = 8 \\ 3x + 10y = 16 \end{cases} \implies (x, y) = (2, 1).$$

- **Trigonometry**: You should be know how sin and cos work, in both degrees and radians. So you should know $\sin(30°) = \frac{1}{2}$, and $\cos\left(\frac{7\pi}{6}\right) = -\frac{\sqrt{3}}{2}$.

  You should know a few trig identities; the most important is the double angle formula

$$\begin{aligned} \sin(2\theta) &= 2\sin\theta\cos\theta \\ \cos(2\theta) &= \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta. \end{aligned}$$

- **Precalculus**: You should know some common formulas covered in precalculus for vectors and matrices:
  - You should be able to add and scale vectors, like

$$\begin{pmatrix} 1 \\ 7 \end{pmatrix} + 10\begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \end{pmatrix} + \begin{pmatrix} 30 \\ 50 \end{pmatrix} = \begin{pmatrix} 31 \\ 57 \end{pmatrix}.$$

    (It's really as easy as the equation above makes it look: do everything componentwise.)

  - You should know the rule for matrix multiplication, so that for example you could carry out the calculation

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 1\cdot7 + 2\cdot8 + 3\cdot9 \\ 4\cdot7 + 5\cdot8 + 6\cdot9 \end{pmatrix} = \begin{pmatrix} 50 \\ 122 \end{pmatrix}.$$

    If you haven't seen this before, there are plenty of tutorials online; find any of them. Poonen's notes (mentioned later) do cover this for example; see section 1-2 of `https://math.mit.edu/~poonen/notes02.pdf`.

You are *not* expected to have any idea why the heck the rule is defined this way; an explanation for where this rule comes from is in Section 7.3. So we'll assume you have seen this strange rule before, but don't know what it means.

- We'll assume you know the formula for the determinant of a 2×2 and 3×3 matrix; that is

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

and

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

(The bars are a shorthand for the det symbol; they're not absolute value bars.)

For example, you should be able to verify the correctness of the following equation:

$$\begin{vmatrix} 0 & 1 & 5 \\ 2 & 0 & 13 \\ 1 & 4 & 1 \end{vmatrix} = 51.$$

We won't assume you know where this formula comes from, and in fact we won't be able to explain that within these notes. But if you're curious, you should read Chapter 12 of the Napkin.

- **Calculus**: You should know single variable derivatives and integrals, for example:
  - You should be able to differentiate $x^7 + \sin(x)$ to get $7x^6 + \cos(x)$.
  - You should be able to integrate $\int_0^1 x^2 \, dx$ to get $\frac{1}{3}$.

  This is covered in the course 18.01 at MIT, and also in the AP calculus courses in the United States.

  One note: **by $\log(x)$ we mean the *natural* log with base e.**[3] We will never use a base-2 or base-10 logarithm in these notes.

> **Tip**
>
> If you're not at MIT, you should replace the words "18.01" and "18.02" with the course names corresponding to "single-variable calculus" and "multi-variable calculus" at your home institution.

This book assumes no proof-based background.

§1.3 [TEXT] Topics covered

Here is a brief overview of what happens in these parts.

Alfa and Bravo This part covers **linear algebra** (vectors and matrices). This is intentional, because some working knowledge of linear algebra is important. In fact, if I was designing a serious course in multivariable calculus for math majors, it would come after an entire semester of properly-done linear algebra first.

[3] I considered using the notation $\ln(x)$ to avoid confusion. However, $\ln(x)$ is never used by mathematicians past introductory calculus; see https://math.stackexchange.com/q/293783/229197. I figured I should just get you used to $\log(x)$ being base e now. There's a real chance that if you take an 18.02 exam at MIT, the professor straight-up forgets to remind the students that $\log(x)$ is base e, because they haven't used $\ln(x)$ in a quarter century.

Charlie This short part is review of the **complex numbers** $\mathbb{C}$. I actually don't know why this is part of 18.02, to be honest, but since it happened I included a short chapter on it.

Delta Covers the calculus of functions $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$, which is usually thought of as a **parametric** function $\mathbf{r}(t)$ (a time-indexed trajectory through the vector space $\mathbb{R}^n$). This part turns out to be easy because it's pretty much all 18.01 material. This part is therefore also only a few pages long.

Echo and Foxtrot Cover the **differentiation of multivariable functions** $f : \mathbb{R}^n \to \mathbb{R}$, and the optimization such functions. The star of these two parts in the gradient ∇f, which gets airtime in virtually every kind of question you'll see. This is the first serious multivariable calculus usage.

Golf This part covers **double integrals of functions** $f : \mathbb{R}^2 \to \mathbb{R}$, the first of the parts on integration. We define the double integral and cover techniques for computing them.

Hotel This part covers **integrals of scalar functions in space** $f : \mathbb{R}^3 \to \mathbb{R}$. It introduces the triple integral (which isn't any different from double integral) as well as a side detour on arc length and surface area.

India This part covers **line integrals of vector fields** $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ over a **curve**; that is, work and 2D flux. The famous grad, div, and curl are first mentioned here, together with the generalized Stokes' theorem that ties them all together. This is the iconic part of multivariable calculus (kind of like how France is associated with the Eiffel tower, say).

Juliett This part covers **surface integrals of vector fields** $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ over a **surface**; that is, the flux integral. More versions of Stokes' theorem are given.

Kilo Exercises covering all the earlier parts.

Lima Solutions to exercises from the text.

Mike Appendix of odds and ends such as excessively long digressions.

(The words Alfa, Bravo, Charlie, etc. are from the NATO phonetic alphabet which the author of this book has memorized from overexposure to puzzle hunts.)

§1.4 [TEXT] The structure of this book

You will quickly notice that all the sections are labeled with different headings. Here's an explanation of what they mean.

TEXT Good old prose. An explanation like you might hear in a lecture.
RECIPE Has only the final recipe, as you need it on the exam. As I mentioned before, I don't like the idea of just memorizing recipes, but in theory you might still be able to pass the exams by doing only this.
SIDENOTE An optional extended discussion. You can skip these unless you're interested in them.
RECAP A summary of what happened in the chapter.
EXER Problems to work on. Starred exercises are harder than questions that will appear in the actual MIT course.

You'll also see some colored boxes that mark where certain chunks begin and end. These should be self-explanatory.

§1.5 [TEXT] Other references

The best resource I have for 18.02 in text is definitely Bjorn Poonen's fall 2021 notes, available at

`https://math.mit.edu/~poonen/notes02.pdf`.

Poonen is a really great writer of mathematical exposition in general, and I highly recommend these notes as a result. In fact, I will even tell you, for each part, what the corresponding sections of Poonen are if you decide something I write doesn't make sense and you want to reference the corresponding text. (That said, this text should stand alone.)

There are many other resources on multivariable calculus out there too. For example, MIT OpenCourseWare has some supplementary notes and problems still used by the math department. And so on. You can also find countless final exams from previous years of 18.02 on OCW.

I think the term "treatment" for the way a course is taught is apt, because it reflects a reality about education: like medicine, there is no one treatment that works for everyone. In theory, there might be some people who only read this book and that's all they need. In practice, many of you would benefit from asking friends to explain things differently for the sections of the book that don't work from you, or consulting another text when things here don't suit you. You should not feel under any obligation to treat this book as the one true bible of 18.02. This book is meant to be an aid, not a cage.

§1.6 [SIDENOTE] If you're thinking of becoming a math major

If you're thinking of becoming a math major, there's some advice in Chapter 53.

§1.7 [SIDENOTE] My exercises are harder, so take your time

When setting exercises, I tried to come up with questions that require a bit of thought and understanding, for learning purposes. I'm intentionally trying to stretch you slightly with my exercises while the timer is off — I want to give you a little bit of an opportunity to take your time and think. I think you'll internalize the material better this way and it'll pay off.

But when you actually take an 18.02 midterm in real life, you have no time to think[4]— you have to answer each question in 5-10 minutes. So on the flip side, you will probably be pleasantly surprised when you find that 50%-80% of real midterm questions can be solved by turning *off* your brain and following recipes to the letter. It has to be this way because of the short time limit and the amount of material.

All this is to say to **not be discouraged if you find the exercises in this book harder**. It's by design. The real exam will have many cookie-cutter no-thought questions in return for the short time limit.

(Like most textbooks, the starred exercises are more challenging.)

§1.8 [SIDENOTE] Acknowledgments

- Thank you to the staff and other recitation leaders who made this course possible; particularly Davesh Maulik for leading the instance of the course this year full-heartedly and Karol Bacik for making so much happen behind the scenes. Thanks also to Sefanya Hope for coordinating many other logistics, and particularly for helping me book classrooms on short notice on many occasions.

  I also thank Ting-Wei Chao for his permission to use Exercise 19.3, Exercise 33.1, Exercise 34.3, Exercise 35.1 from his recitation section.

- Thank you to all the students in my recitation session (and those officially enrolled in other sessions, but who came to my session anyway!) who regularly attended my class every Monday and Wednesday at 9am. That's some real early-morning dedication. There's a saying that the enthusiasm of an instructor can be contagious, but I definitely think the enthusiasm of students can be as well.

[4] If you're in India, the JEE exam is even more about speed and tricks than having any real understanding, and I apologize that you have to suffer through it.

- In particular, I got many words of thanks and encouragements from my students this year, which I am grateful for. I certainly wouldn't have had the motivation to type these notes without these kind words.

- I thank Aaryan Vaishya, Alan Cheng, Alexander Wang, Calvin Wang, John Zhou, Nick Zhang, Rémi Geron, Ritwin Narra, Rohan Garg, and Royce Yao for many corrections. (Your name could be here too — find me some typos! If you know how to open a GitHub pull request, the relevant repository is `https://github.com/vEnhance/1802`.)
- Thank you to Catherine Xu for the cover art. You can download a full-resolution copy at `https://web.evanchen.cc/textbooks/lamv-cover-art.png` or find it in the GitHub repository.
- Thanks to OpenAI for gifting me a Plus subscription to ChatGPT. Writing this text gave me an excuse to get a chance to use ChatGPT 4o and ChatGPT o1-preview, to see what kind of things it did well (and what I could still do faster by hand).
  - ChatGPT was helpful at writing full step-by-step solutions to the routine exercises. All the solutions went through much editing from me (in part to make the notation consistent throughout the whole text), to the point where maybe only a third of the output from each solution actually survives editing. Even then, because it's faster to edit or rewrite text[5] than write from scratch, it still saved time.

    I think when humans write solutions they err on the size of laziness in skipping steps that are really routine or obvious to them, because typing is slow. ChatGPT doesn't; in fact, it's actually *too* verbose, and I almost always had to trim down the solution. But it much easier to trim down an overly verbose solution than to flesh out one that's too terse.

    It's also nice to not have to worry much about arithmetic errors anymore. If I had written the solutions by hand, I would certainly drop plenty of factors of 2 or flip signs. ChatGPT actually made fewer errors than me, and when it did it was usually easy for me to spot because it wrote everything out. It turns out that proofreading someone else's work is much, much easier than proofreading your own.
  - It was also fairly good at *generating* new routine exercises that are solved by just applying the formula. It wasn't perfect; some of the exercises it would generate were obviously broken. But again, with some editing, it was still faster than trying to make up uninspired exercises one after another en masse.[6]
  - It's pretty good at "explain things in many words" in a conversational way. For example, the aquatic descriptions of what curl or divergence or work mean were largely generated by ChatGPT.
- This book was not written in LaTeX! It's written in the recently released version of Typst 0.12, which is open source at `https://github.com/typst/typst`. I used NeoVim as an editor, doing everything locally rather than by web app.

  It was really nice being able to write math without having to constantly use the backslash key or the curly brace, and the compiler was much faster, so I was overall quite impressed with my Typst experience. Typst is quite new and you should check them out if you'd like to give them a try!
  - ChatGPT is not good at producing Typst output yet, so I had ChatGPT output everything in LaTeX and would then convert using `pandoc`. This conversion had some undesirable

[5] Vim on top. Fight me.

[6] Which I admit I have too much pride to do either, as a math olympiad kid whose success story was built on solving many non-routine problems growing up, rather than mass-generated ones.

irregularities, so I ended up writing a few Python scripts to handle those irregularities and just fixed the remaining issues by hand.

- You can see the source code for this textbook at `https://github.com/vEnhance/1802`. Note that it relies on an external file called `evan.typ` which you can find in my dotfiles repository. Thanks to the couple packages that were already used in `evan.typ`:
  - The v2 version of `https://github.com/sahasatvik/typst-theorems`.
  - The v1 version of `https://github.com/jomaway/typst-gentle-clues`.

Chapter 2. Type safety

Before we get started with the linear algebra and calculus, I want to talk quickly about *types of objects*. This is an important safeguard for the future in checking your work and auditing your understanding of a topic; a good instructor will point out, in your work, any time you make a type-error.

§2.1 [TEXT] Type errors

In mathematics, statements are usually either true or false. Examples of false statements[7] include

$$\pi = \frac{16}{5} \qquad \text{or} \qquad 2+2=5.$$

However, it's possible to write statements that are not merely false, but not even "grammatically correct", such as the nonsense equations

$$\pi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix} = \cos \begin{pmatrix} 6 \\ 7 \end{pmatrix}, \qquad \det \begin{pmatrix} 5 \\ 11 \end{pmatrix} \neq \sqrt{2}.$$

To call these equations false is misleading. If your friend asked you whether $2+2=5$, you would say "no". But if your friend asked whether π equals the 2×2 identity matrix, the answer is a different kind of "no"; really, it's "your question makes no sense".

These three examples are **type errors**. This term comes from programming: most programming languages have different data types like integer, boolean, string, array, etc., and will usually[8] prevent you from doing anything idiotic like adding a string to an array.

Objects in mathematics work in a really similar way. In the first weeks of 18.02, you will meet real numbers, vectors, and matrices; these are all different types of objects, and certain operations are defined (aka "allowed") or undefined (aka "not allowed") depending on the underlying types. Table Table 1 lists some common examples with vectors you've seen from precalculus.

Operation	Notation	Input 1	Input 2	Output
Add/subtract	$a \pm b$	Scalar	Scalar	Scalar
Add/subtract	$\mathbf{v} \pm \mathbf{w}$	d-dim vector	d-dim vector	d-dim vector
Add/subtract	$M \pm N$	$m \times n$ matrix	$m \times n$ matrix	$m \times n$ matrix
Multiply	$c\mathbf{v}$ or $c \cdot \mathbf{v}$	Scalar	d-dim vector	d-dim vector
Multiply	ab or $a \cdot b$	Scalar	Scalar	Scalar
Multiply	MN or $M \cdot N$	$m \times n$ matrix	$n \times p$ matrix	$m \times p$ matrix
Dot product	$\mathbf{v} \cdot \mathbf{w}$	d-dim vector	d-dim vector	Scalar
Cross product	$\mathbf{v} \times \mathbf{w}$	3-dim vector	3-dim vector	3-dim vector
Length/mag.	$\lvert\mathbf{v}\rvert$	Any vector	*n/a*	Scalar
Determinant	$\det A$	Any square matrix	*n/a*	Scalar

Table 1: Common linear algebra operations. For 18.02, "scalar" and "real number" are synonyms.

[7] Indiana Pi bill and 1984, respectively.

[8] JavaScript is a notable exception. In JavaScript, you may know that `[]` and `{}` are an empty array and an empty object, respectively. Then `[]+[]` is the empty string, `[]+{}` is the string `'[object Object]'`, `{}+[]` is `0`, and `{}+{}` is `NaN` (not a number).

Digression

A common question at this point is how you are supposed to figure out whether a certain operation is allowed or not. For example, many students want to try and multiply two vectors together component-wise; why is

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} \stackrel{?}{=} \begin{pmatrix} 8 \\ 15 \end{pmatrix}$$

not a legal sentence? It seems like it would make sense.

The answer is that you *don't* have to figure out — you are *told*; Table 1 isn't something that you derive. That is, Table 1 consists of the *definitions* which you have been given.

(Or more sarcastically, it's all just a social construct. Well, it's bit more nuanced than that. Definitions aren't judged by "correctness"; that doesn't make sense; you are allowed to make up whatever definitions you want. Instead, definitions are judged by whether they are *useful*. Which is obviously subjective, but it's less subjective than you might guess.)

§2.2 [TEXT] Why you should care

There are two action items to take away from this chapter.

§2.2.1 When learning a new object, examine its types first

What this means is that, every time you encounter a new kind of mathematical object or operation (e.g. partial derivative), **the first thing you should do is ask what types are at play**. This helps give you a sanity check on your understanding of the new concept.

We'll use boxes like this throughout the box to do this:

</> Type signature

This is an example of a type signature box. When we want to make comments about the types of new objects, we'll put them in boxes like this.

§2.2.2 Whenever writing an equation, make sure the types check out

Practically, what's really useful is that if you have a good handle on types, then it **gives you a way to type-check your work**. This is the analog of dimensional analysis from physics, where you know you messed up if some equation has $\text{kg} \cdot \text{meters} \cdot \text{seconds}^{-2}$ on the left but $\text{kg} \cdot \text{meters} \cdot \text{seconds}^{-1}$ on the right.

For example, if you are reading your work and you see something like

$$|\mathbf{v} \times \mathbf{p}| = 9\mathbf{p} \tag{1}$$

then you can immediately tell that there's a mistake, because the two sides are incompatible — the left-hand side is a real number (scalar), but the right-hand side is a vector.

§2.3 [RECAP] Takeaways from type safety

- Throughout this book, every time you meet a new operation, make sure you know what types of objects it takes as input and which it takes as output.

- Whenever you write an equation, make sure it passes a type-check. You can catch a lot of errors like Equation 1 using type safety alone.

§2.4 [EXER] Exercises

Exercise 2.1. Let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ be vectors in $\mathbb{R}^3$. By using Table 1 (or skimming Section 4.1 briefly), determine whether each of the following expressions is a real number, a vector, or nonsense (type-error); there should be one of each.

- $(\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w}$
- $\mathbf{u} \cdot \mathbf{v} + \mathbf{w}$ (here order of operations is $\cdot$ before $+$)
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$

(The symbol $\cdot$ confusingly can refer to three different things: grade-school multiplication, scalar multiplication, or the dot product.)

(The answer to this exercise is written in Chapter 42, and in general Part Lima contains solutions to all the exercises.)

Part Alfa: Linear Algebra of Vectors

For comparison, Part Alfa corresponds roughly to §1, §2, §3.9 of Poonen's notes.

Chapter 3. Review of vectors

§3.1 [TEXT] Notation for scalars, vectors, points

If you haven't seen the symbol $\mathbb{R}$ before, let's introduce it now:

Definition

We denote by $\mathbb{R}$ the real numbers, so $3, \sqrt{2}, -\pi$ are elements of $\mathbb{R}$. Sometimes we'll also refer to a real number as a **scalar**.

The symbol "$\in$", if you haven't seen it before, means "is a member of". So $3 \in \mathbb{R}$ is the statement "3 is a real number". Or $x \in \mathbb{R}$ means that x is a real number.

Unfortunately, right off the bat I have to mention that the notation $\mathbb{R}^n$ could mean two things:

Definition

By $\mathbb{R}^n$ we could mean one of two things, depending on context:

- The vectors of length n, e.g. the vector $\begin{pmatrix} \pi \\ 5 \end{pmatrix}$ is a vector in $\mathbb{R}^2$.
- The points in n-dimensional space, e.g. $(\sqrt{2}, 7)$ is a point in $\mathbb{R}^2$.

To work around the awkwardness of $\mathbb{R}^n$ meaning two possible things, this book will adopt the following conventions for variable names:

Type signature

- Bold lowercase letters like $\mathbf{u}$ and $\mathbf{v}$ will be used for vectors. When we draw pictures of vectors, we always draw them as *arrows*.
- Capital letters like P and Q are used for points. When we draw pictures of points, we always draw them as *dots*.
- Sometimes, if we need to refer to the vector drawn as an arrow which starts at point P and ends at Q, we write $\overrightarrow{PQ}$ for it.

We'll also use different notation for actual elements:

Type signature

- A vector will either be written in column format like $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, or with angle brackets as $\langle 1, 2, 3 \rangle$ if the column format is too tall to fit.
- But a point will always be written with parentheses like $(1, 2, 3)$.

Some vectors in $\mathbb{R}^3$ are special enough to get their own shorthand. (The notation "$:=$" means "is defined as".)

Definition

When working in $\mathbb{R}^2$, we define

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$\mathbf{0} := \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Definition

When working in $\mathbb{R}^3$, we define

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

We also let

$$\mathbf{0} := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

In other places, you'll sometimes see $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ instead, but this book will always use $\mathbf{e}_i$.

§3.2 [TEXT] Length

Definition

The **length** of a vector is denoted by $|\mathbf{v}|$ and corresponds to the length of the arrow drawn. It is given by the Pythagorean theorem.

- In two dimensions:

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} \Longrightarrow |\mathbf{v}| := \sqrt{x^2 + y^2}.$$

- If three dimensions:

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Longrightarrow |\mathbf{v}| := \sqrt{x^2 + y^2 + z^2}.$$

In n dimensions, if $\mathbf{v} = \langle a_1, ..., a_n \rangle$, the length is $|\mathbf{v}| := \sqrt{a_1^2 + ... + a_n^2}$.

Type signature

The length $|\mathbf{v}|$ has type scalar. It is always positive unless $\mathbf{v} = \mathbf{0}$, in which case the length is 0.

§3.3 [TEXT] Directions and unit vectors

Remember that a vector always has

- both a **magnitude**, which is how long the arrow is in the picture, and
- a **direction**, which refers to which way the arrow points.

In other words, the geometric picture of a vector always carries two pieces of information. (Here, I'm imagining we've drawn the vector as an arrow with one endpoint at the origin and pointing some way.)

In a lot of geometry situations we only care about the direction, and we want to ignore the magnitude. When that happens, one convention is to just set the magnitude equal to 1:

Definition

A **unit vector** will be a vector that has magnitude 1.

Thus we use the concept of unit vector to capture direction. So in $\mathbb{R}^2$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is thought of as "due east" and $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ is "due west", while $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is "due north" and $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ is "northeast".

Definition

Given any vector $\mathbf{v}$ in $\mathbb{R}^n$ besides the zero vector, the **direction along** $\mathbf{v}$ is the unit vector

$$\frac{\mathbf{v}}{|\mathbf{v}|}$$

which is the unit vector that points the same way that $\mathbf{v}$ does.

(Depending on what book you're following, more pedantic authors might write "the unit vector in the direction of $\mathbf{v}$" or even "the unit vector in the same direction as $\mathbf{v}$" rather than "direction along $\mathbf{v}$". This is too long to type, so I adopted the shorter phrasing. I think everyone will know what you mean.)

We will avoid referring to the direction of the zero-vector $\mathbf{0}$, which doesn't make sense. (If you try to apply the formula here, you get division by 0, since $\mathbf{0}$ is the only vector with length 0.) If you need it, the convention is that it has *every* direction.

Type signature

If $\mathbf{v}$ is a nonzero vector in $\mathbb{R}^n$, then the direction along $\mathbf{v}$ is a (unit) vector in $\mathbb{R}^n$.

Example

Let's first do examples in $\mathbb{R}^2$ so we can drawn some pictures.

- The direction along the vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 5 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 1337 \\ 0 \end{pmatrix}$ are all $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, thought of as *due east*.
- But the direction along the vectors $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} -9 \\ 0 \end{pmatrix}$ are both $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$, thought of as *due west*.
- The direction along the vectors $\begin{pmatrix} 0 \\ -2 \end{pmatrix}$, $\begin{pmatrix} 0 \\ -17 \end{pmatrix}$ are all $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$, thought of as *due south*.

Example

How about the direction along $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$? We need to first find the length of the vector so we can scale it down. That's given by the Pythagorean theorem, of course:

$$\left| \begin{pmatrix} 3 \\ -4 \end{pmatrix} \right| = \sqrt{3^2 + 4^2} = 5.$$

So the direction along $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$ would be

$$\frac{1}{5} \begin{pmatrix} 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix}.$$

See Figure 1. The direction is somewhere between south and southeast.

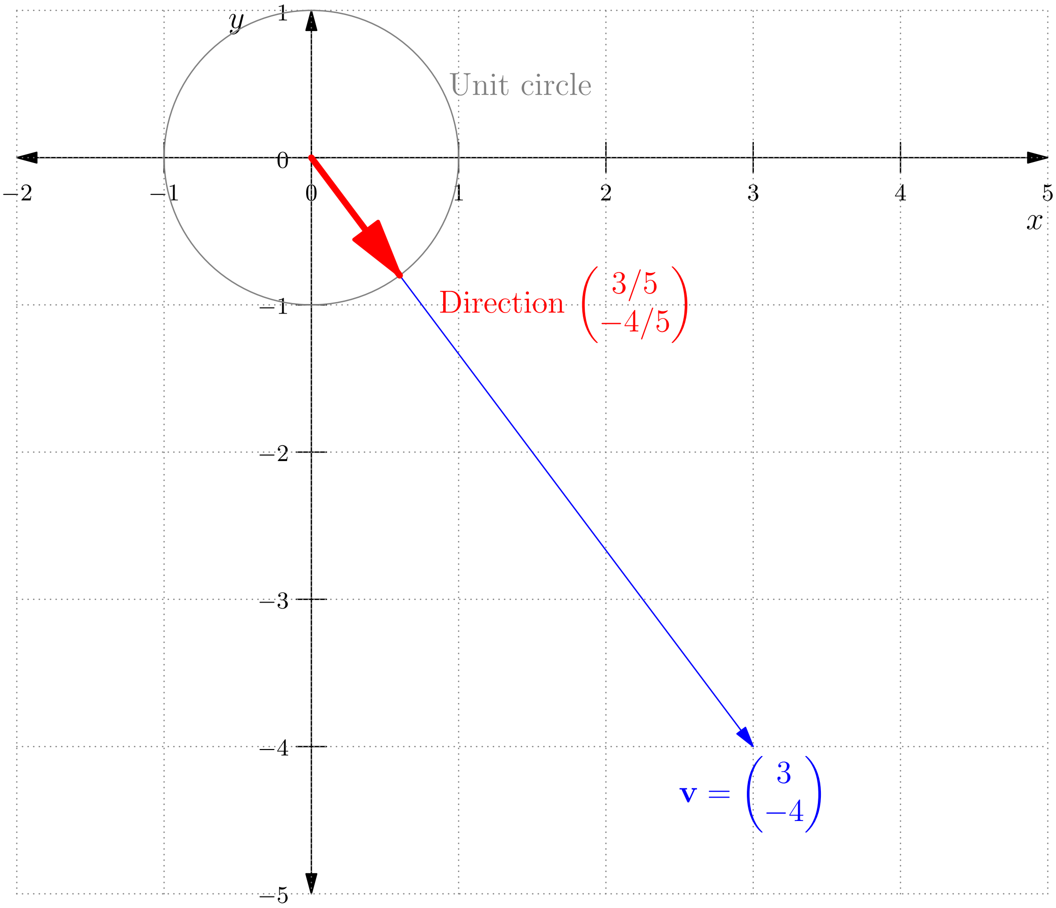

Figure 1: The direction along $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$ (blue) is $\begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix}$ (red). Unit vectors always lie on the grey circle (unit circle) by definition.

When drawn like Figure 1 in the two-dimensional picture $\mathbb{R}^2$, the notion of direction along $\begin{pmatrix} x \\ y \end{pmatrix}$ is *almost* like the notion of slope $\frac{y}{x}$ in high school algebra (so the slope of the blue ray in Figure 1). But there are a few reasons our notion of direction is more versatile than just using the slope of the blue ray.

- The notion of direction can tell the difference between $\begin{pmatrix} 3 \\ -4 \end{pmatrix}$, which goes southeast, and $\begin{pmatrix} -3 \\ 4 \end{pmatrix}$, which goes northwest. Slope cannot; it would assign both of these "slope $-\frac{4}{3}$".

- The due-north and due-south directions $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ would have undefined slope due to division-by-zero, so you always have to worry about this extra edge case. With unit vectors, due-north and due-south don't cause extra headache.
- Our definition of direction works in higher dimension $\mathbb{R}^3$. There isn't a good analog of slope there; a single number cannot usefully capture a notion of direction in $\mathbb{R}^n$ for $n \geq 3$.

Example

The direction along the three-dimensional vector $\begin{pmatrix} 12 \\ -16 \\ 21 \end{pmatrix}$ is

$$\begin{pmatrix} 12/29 \\ -16/29 \\ 21/29 \end{pmatrix}.$$

See if you can figure out where the 29 came from.

§3.4 [RECIPE] Areas and volumes

If $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ are vectors, drawn as arrows with a common starting point, then their sum $\mathbf{v}_1 + \mathbf{v}_2$ makes a parallelogram in the plane with $\mathbf{0}$ as shown in Figure 2.

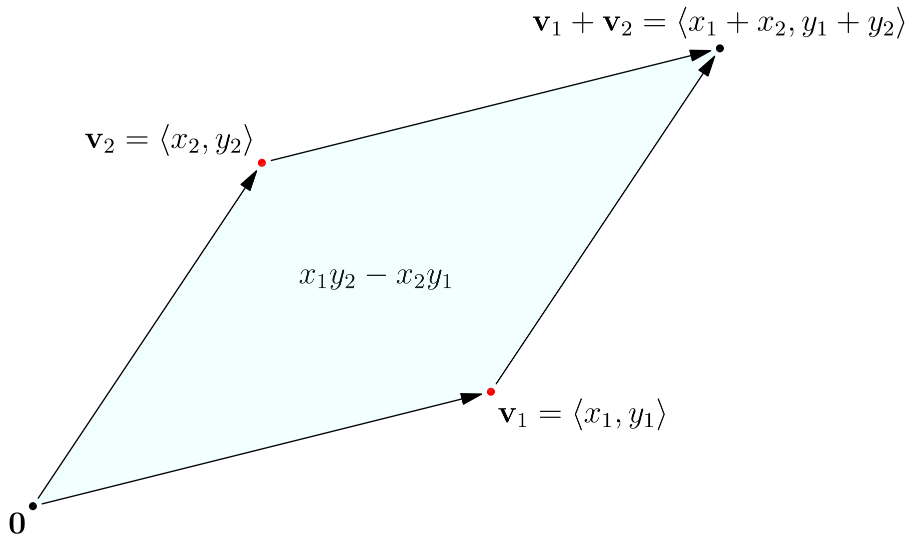

Figure 2: Vector addition in $\mathbb{R}^2$.

The following theorem is true, but we won't be able to prove it in 18.02.

Recipe for area of a parallelogram

The signed area of the parallelogram formed by $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is equal to

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

A similar theorem is true for the parallelepiped[9] with three vectors in $\mathbb{R}^3$; see Figure 3.

[9] I hate trying to spell this word.

Recipe for volume of a parallelepiped

The signed volume of the parallelepiped formed by $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}$ is equal to

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

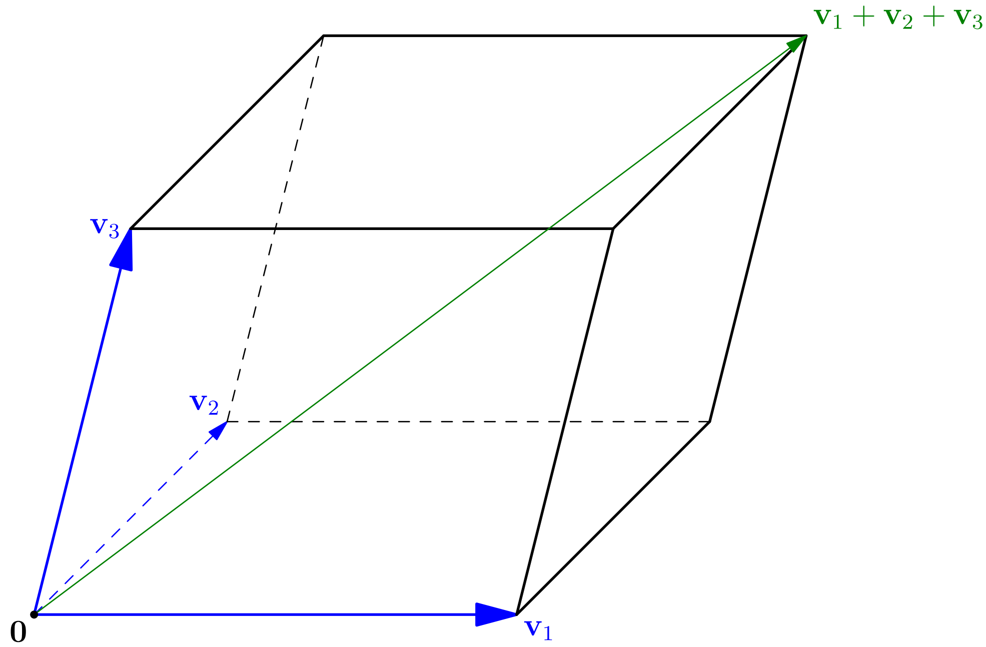

Figure 3: Three vectors in $\mathbb{R}^3$ making a parallelepiped.

You might noticed that the word "signed" has slipped in before "area" and "volume". What does that mean? Well, if you only care about the area of the volume itself, it doesn't matter for you; you should just take the absolute value of the determinant. But the sign carries a bit more information.

- In 2D, consider the angle between $\mathbf{v}_1$ and $\mathbf{v}_2$, between $0°$ and $180°$. Then we consider the sign to be $+$ if the angle goes counterclockwise from $\mathbf{v}_1$ to $\mathbf{v}_2$, (like the example in Figure 2), and negative otherwise. So in Figure 2, we would have

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = +\text{area}, \quad \begin{vmatrix} x_2 & x_1 \\ y_2 & y_1 \end{vmatrix} = -\text{area}.$$

- In 3D, the convention follows the right-hand rule: suppose vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are given it that order. Curl the fingers of your right hand from $\mathbf{v}_1$ to $\mathbf{v}_2$; then the signed volume is positive if your thumb points in the direction of $\mathbf{v}_3$ (as in Figure 3, for example) and negative otherwise.

Digression

If you're interested in the proof of these results and their n-dimensional generalizations, the tool needed is the **wedge product**, which is denoted

$$\bigwedge^k (\mathbb{R}^n).$$

This is well beyond the scope of 18.02, but it's documented in Chapter 12 of my Napkin for those of you that want to read about it.

Alternatively, I think Wikipedia and Axler[10], among others, use a definition of the determinant as the unique multilinear alternating map on n-tuples of column vectors in $\mathbb{R}^n$ that equals 1 for the identity. This definition will work, and will let you derive the formula for determinant, and gives you a reason to believe it should match your concept of area and volume. It's probably also easier to understand than wedge products. However, in the long term I think wedge products are more versatile, even though they take much longer to setup.

§3.5 [EXER] Exercises

Exercise 3.1. Compute the unit vector along the direction of the vector

$$\begin{pmatrix} -0.0008\pi \\ -0.0009\pi \\ -0.0012\pi \end{pmatrix}.$$

Exercise 3.2. If A is a 3×3 matrix with determinant 2, what values could $\det(10A)$ take?

Exercise 3.3. Compute the real number a for which the points $(0,0,0)$, $(1,0,1)$, $(0,1,2)$ and $(1,2,a)$ all lie on one plane.

[10] Who has a paper called Down with Determinants!, which I approve of.

Chapter 4. The dot product

The dot product is the first surprising result you'll see in this class, because it has *two* definitions that look nothing alike, one algebraic and one geometric. Because of that, we'll be able to get a ton of mileage out of it.

This will be a general theme across the course: almost every new concept we define will have some sort "algebraic" side (like the coordinates for vector addition) and some "geometric" side (the parallelogram in Figure 2). This is the bar a concept has to pass for us to study it in this class: in order for us to deem a concept worthy of our attention in 18.02, it must have both an interpretation with algebra and an interpretation in geometry.

§4.1 [TEXT] Two different definitions of the dot product

I promised you two definitions right? So here they are.

Definition

Suppose $\mathbf{v} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ are two vectors in $\mathbb{R}^n$.

The *algebraic definition* is to take the sum of the component-wise products:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := a_1 b_1 + \ldots + a_n b_n.$$

The *geometric definition* is that if θ is the angle between the two vectors when we draw them as arrows with a common starting point, then

$$\mathbf{v} \cdot \mathbf{w} := |\mathbf{v}| \; |\mathbf{w}| \cos \theta.$$

That is, the dot product equals the product of the lengths times the cosine of the included angle.

It's totally not obvious that these two definitions are the same! I know two reasonable proofs, both of which I've typed in the appendix:

- The standard proof uses the law of cosines; it's documented in Section 54.1. It's short but seems somewhat magical.
- I came up with a geometric proof without trigonometry; it's documented in Section 54.2. It's longer but easier to come up with.

I won't dwell on this proof too much in the interest of moving these notes forward.

Type signature

Remember, the dot product takes two vectors *of equal dimensions* as inputs and outputs a *scalar* (i.e. a real number). **It does not output a vector!** This is the mistake every calculus or linear algebra instructor dreads for the first few weeks of class.

Repeat: dot product output type is **number**! Not a vector!

Warning: There are a lot of dots, aren't there?

Confusingly, the multiplication $\cdot$ is also used for normal multiplication (as we saw in Exercise 2.1). This is why you need to always look at the types of objects so you know which $\cdot$ is happening. To spell this out:

- If a and b are two numbers, $a \cdot b = ab$ is *grade-school multiplication*, e.g. $3 \cdot 5 = 15$.
- If a is a number and $\mathbf{v}$ is a vector, $a \cdot \mathbf{v}$ is *scalar multiplication*, e.g. $3 \cdot \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 15 \\ 21 \end{pmatrix}$.
- If $\mathbf{v}$ and $\mathbf{w}$ are vectors, then $\mathbf{v} \cdot \mathbf{w}$ is *dot product*, e.g. $\begin{pmatrix} 5 \\ 7 \end{pmatrix} \cdot \begin{pmatrix} 9 \\ 11 \end{pmatrix} = 5(9) + 7(11) = 122$.

Example

Let's compute the dot product of $\mathbf{v} = \begin{pmatrix} -5 \\ 5\sqrt{3} \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 7\sqrt{3} \\ -7 \end{pmatrix}$, both ways.

- The algebraic definition is easy:
$$\mathbf{v} \cdot \mathbf{w} = -5 \cdot 7\sqrt{3} + 5\sqrt{3} \cdot (-7) = -70\sqrt{3}.$$
- The geometric definition is a bit more work, see Figure 4. In this picture, you can see there are two 30° angles between the axes, and the lengths of the vectors are 10 and 14. Hence, the angle θ between them is $\theta = 90^\circ + (30^\circ + 30^\circ) = 150^\circ$. So the geometric definition gives that
$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\ |\mathbf{w}| \cos\theta = 10 \cdot 14 \cdot \cos(150^\circ) = 140 \cdot -\frac{\sqrt{3}}{2} = -70\sqrt{3}.$$

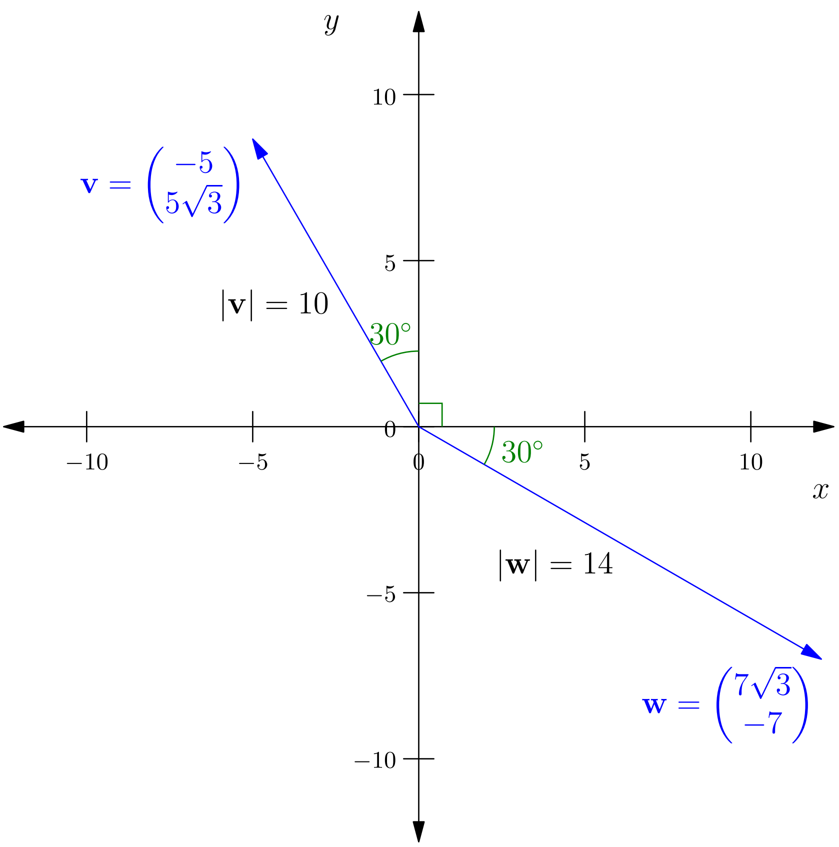

Figure 4: The dot product of $\mathbf{v} = \begin{pmatrix} -5 \\ 5\sqrt{3} \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 7\sqrt{3} \\ -7 \end{pmatrix}$ is $-70\sqrt{3}$.

Tip

You can see from this example that computing the dot product of two given vectors with coordinates is way easier to do with the algebraic definition. This will be true in general throughout this class:

- Use the algebraic definition when you need to do practical calculation.
- Use the geometric definition to interpret the result in some way.

Example

Let's compute the dot product of $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -6 \\ 3 \end{pmatrix}$ both ways. See Figure 5.

- The algebraic definition is easy:

$$\mathbf{v} \cdot \mathbf{w} = 1 \cdot (-6) + 2 \cdot (3) = 0.$$

- In this case the two vectors $\mathbf{v}$ and $\mathbf{w}$ form a 90° angle between them. You should know this from high school, since the two blue rays in Figure 5 have slopes 2 and $-\frac{1}{2}$ respectively. So the cosine of the angle is 0, and the whole dot product is 0. (The lengths are $|\mathbf{v}| = \sqrt{5}$ and $|\mathbf{w}| = 3\sqrt{5}$, but there's no need to actually calculate these.)

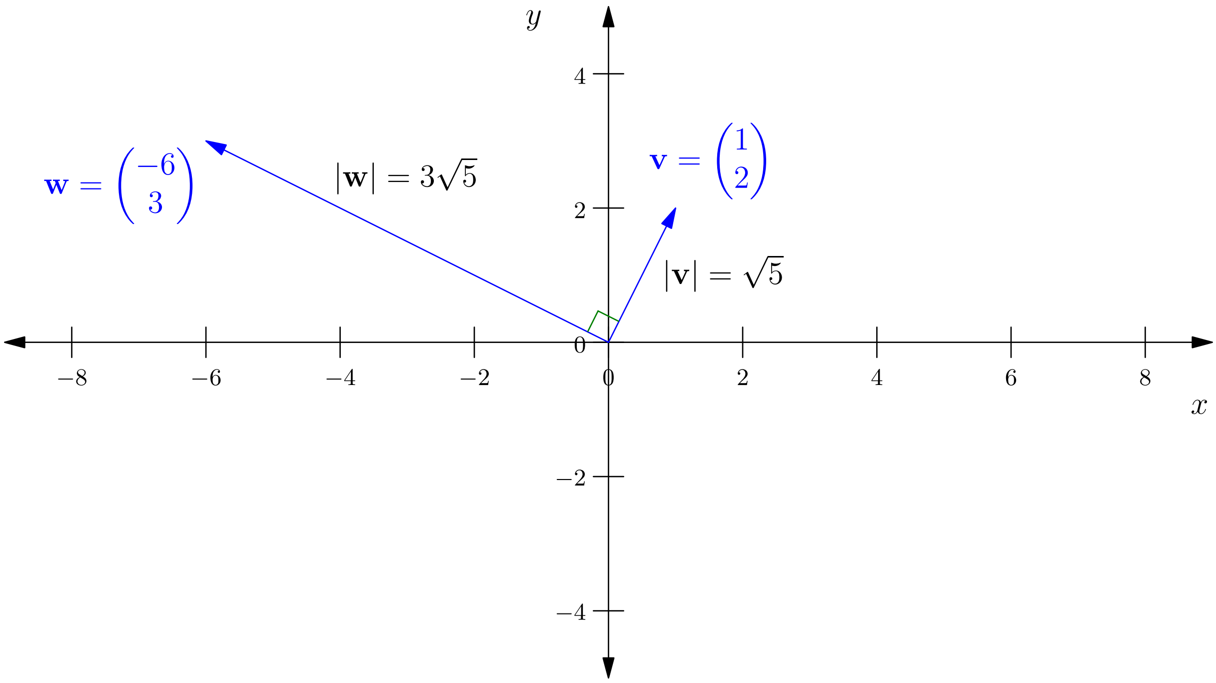

Figure 5: Two perpendicular dot products

This example shows something new:

Memorize

Two nonzero vectors have perpendicular directions if and only if their dot product is 0.

This might seem stupid in two dimensions, because it's doing something you already knew how to do with slope. But in $\mathbb{R}^3$ there isn't a concept of slope, so if you want to see whether two vectors are perpendicular in $\mathbb{R}^3$, you'll want to use the dot product.

Sample Question

Compute the real number t such that $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 5 \\ t \end{pmatrix}$ are perpendicular.

Solution. We need $1 \cdot 4 + 2 \cdot 5 + 3 \cdot t = 0$, so $\boxed{t = -\frac{14}{3}}$. ☐

Example: $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2$

If one takes the dot product of a vector $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ with itself, one gets the **squared length**.

- To see this from the algebraic definition, note that $\mathbf{v} \cdot \mathbf{v} = x_1^2 + \dots + x_n^2 = |\mathbf{v}|^2$.
- To see this from the geometric definition, note that $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|\ |\mathbf{v}| \cos(0) = |\mathbf{v}|^2$.

§4.2 [TEXT] Properties of the dot product

If you look at the algebraic definition, you should be able to see easily that:

- $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$, i.e., dot product is commutative.
- $\mathbf{v} \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v} \cdot \mathbf{w}_1 + \mathbf{v} \cdot \mathbf{w}_2$, i.e., the dot product is distributive.

I point this out briefly so it's on the record, but you'll probably also internalize it automatically as you get more practice with actually computing the dot product.

§4.3 [TEXT] Projection

Suppose $\mathbf{v}$ and $\mathbf{w}$ are two nonzero vectors in $\mathbb{R}^n$. Let θ denote the angle between them. Imagine projecting the vector $\mathbf{v}$ onto the line through $\mathbf{w}$, to get the purple vector shown in Figure 6. This purple vector is typically written $\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$.

Type signature

The vector projection $\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$ is a vector that points in either the same or opposite direction as $\mathbf{w}$.

Let's do an example to see how the dot product lets us compute this.

Example

Suppose $\mathbf{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$, as in Figure 6. How can we find the purple vector $\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$?

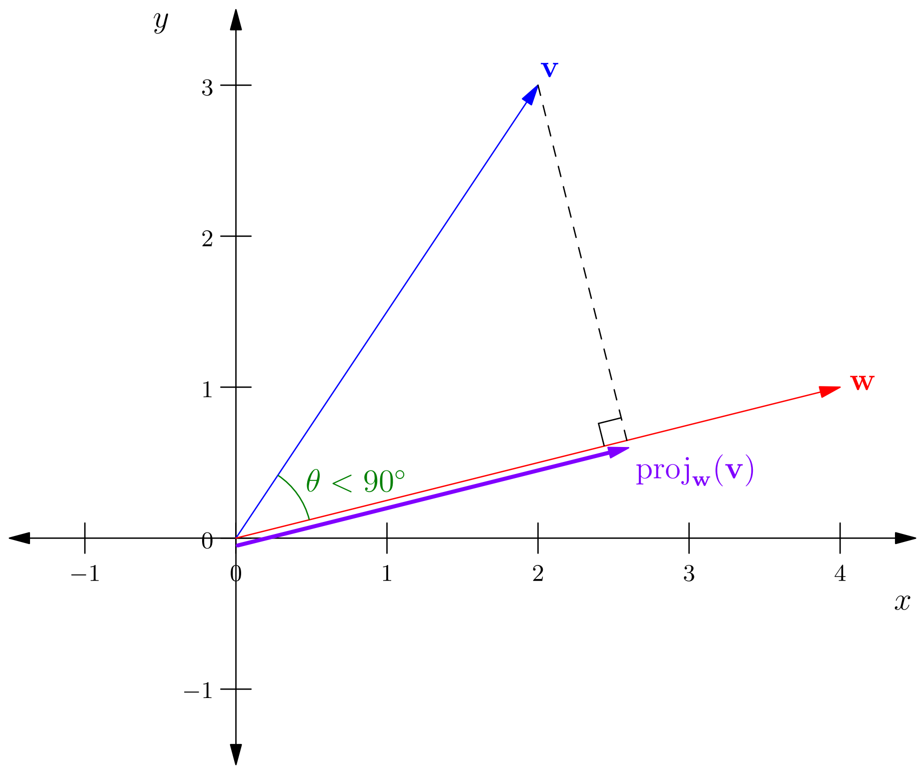

Figure 6: The projection of $\mathbf{v} = \binom{2}{3}$ along $\mathbf{w} = \binom{4}{1}$.

Solution. First, let's figure out the *length* of the purple vector. For trigonometry reasons, we know the *length* of the purple vector is

$$\text{length of purple vector} = |\mathbf{v}| \cos \theta.$$

However, we don't really want to go to the work of figuring out what θ is.

This is where the dot product comes in. It's easy to compute the dot product:

$$11 = 2 \cdot 4 + 3 \cdot 1 = \mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\ |\mathbf{w}| \cos \theta.$$

This is almost what we want, except there's an unneeded $|\mathbf{w}|$ we want to strip out. We know $|\mathbf{w}| = \sqrt{4^2 + 1^2} = \sqrt{17}$, and hence we get

$$\text{length of purple vector} = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|} = \frac{11}{\sqrt{17}}.$$

Now how do we get the purple vector itself? Well, the direction along $\mathbf{w}$ is the unit vector

$$\frac{\mathbf{w}}{|\mathbf{w}|} = \begin{pmatrix} \frac{4}{\sqrt{17}} \\ \frac{1}{\sqrt{17}} \end{pmatrix}$$

and so multiplying by the length gives the desired result:

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = (\text{length of purple vector}) \frac{\mathbf{w}}{|\mathbf{w}|} = \frac{11}{\sqrt{17}} \begin{pmatrix} \frac{4}{\sqrt{17}} \\ \frac{1}{\sqrt{17}} \end{pmatrix} = \begin{pmatrix} \frac{44}{17} \\ \frac{11}{17} \end{pmatrix}.$$ □

> **Remark**
>
> The projection only depends on the direction of $\mathbf{w}$. So if one had re-done the problem above with $\mathbf{w} = \begin{pmatrix} 200 \\ 300 \end{pmatrix}$ instead of $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$, the answer would be the same.

§4.4 [RECIPE] Projection

This procedure last section works in general for any vectors, in any number of dimensions. The only catch is that we have to pay a bit of attention to $\theta < 90°$ and $\theta > 90°$ behaving slightly differently. An example of a situation of that shape is shown in Figure 7.

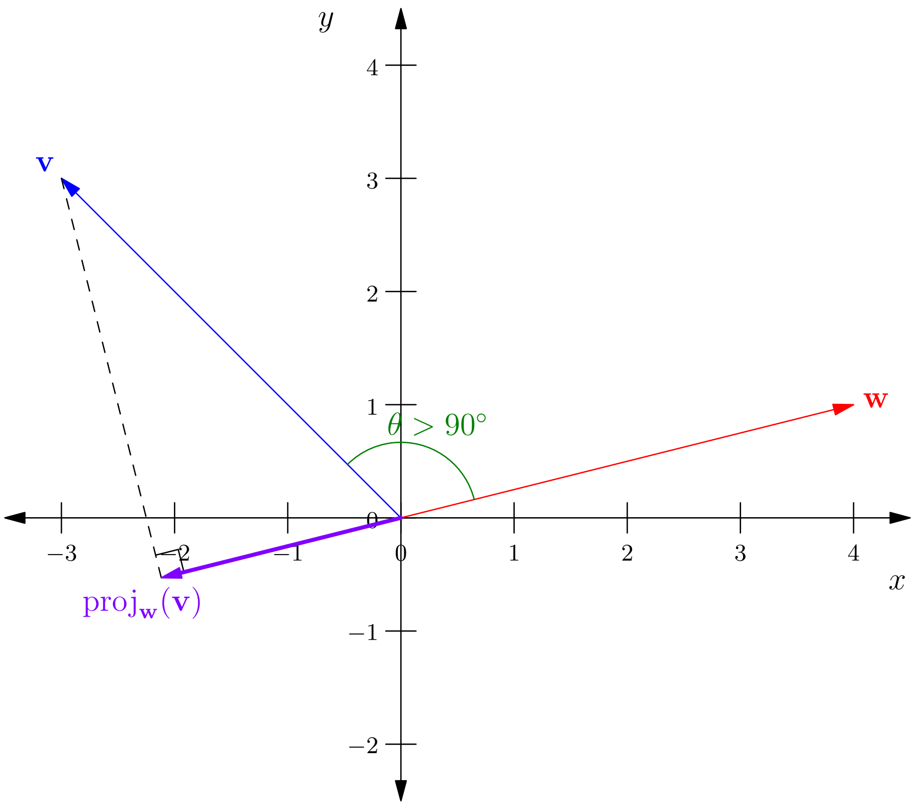

Figure 7: The projection of $\mathbf{v} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$ along $\mathbf{w} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

Here, the purple vector points *away* opposite $\mathbf{w}$.

In the previous example, we used the word "length" and it was fine. In the new figure Figure 7, we would end up taking negative length instead. That works fine, but it's annoying; and so we introduce a new word that works in *both* cases $\theta < 90°$ and $\theta > 90°$:

> **Definition of scalar component**
>
> The **scalar component** of $\mathbf{v}$ in the direction of $\mathbf{w}$ is the number defined by
>
> $$\operatorname{comp}_{\mathbf{w}}(\mathbf{v}) := |\mathbf{v}| \cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|}.$$

This is the analog of purple length from before, but now we allow it to be positive, negative, or zero according to $\theta < 90°$, $\theta > 90°$, and $\theta = 90°$, respectively. But the point is that the cosine can take care of this automatically, and because the cosine is baked into the dot product, life is great: **we just don't have to think about the sign issue at all**. That is, the formula with the new notation

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \operatorname{comp}_{\mathbf{w}}(\mathbf{v}) \frac{\mathbf{w}}{|\mathbf{w}|}$$

is just always true.

Type signature

The scalar component is a number, and can be either positive, negative, or zero.

Recipe for projecting one vector along another

Suppose $\mathbf{v}$ and $\mathbf{w}$ are given vectors in $\mathbb{R}^n$.

1. To compute the **scalar component**, use the formula
$$\operatorname{comp}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|}.$$
2. To compute the **vector projection**, use the formula
$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \operatorname{comp}_{\mathbf{w}}(\mathbf{v}) \frac{\mathbf{w}}{|\mathbf{w}|}.$$

Type signature

Pay attention to type safety when carrying out the recipe to avoid shooting yourself in the foot:
- In the formula $\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|}$, the numerator is a number (it's a dot product), the denominator is a number (it's a length), and we're dividing two numbers.
- The formula $\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|^2}\mathbf{w}$ is more complicated. Focus on just the fraction in front first: the numerator is a number (it's a dot product), and the denominator is a number (it's a squared length), so the entire fraction is a number. This fraction then gets multiplied onto a vector $\mathbf{w}$, so the output type is a vector (actually a multiple of $\mathbf{w}$).

Warning

It's possible to write the projection formula written in other equivalent ways, e.g.

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \operatorname{comp}_{\mathbf{w}}(\mathbf{v}) \frac{\mathbf{w}}{|\mathbf{w}|} = \left(\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|}\right) \frac{\mathbf{w}}{|\mathbf{w}|} = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|^2}\mathbf{w} = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}}\mathbf{w}.$$

I don't like the last few as much because I think they make it harder to see where the formula comes from, but if you know what you're doing, feel free to use them.

To show you the recipe isn't doing anything you haven't seen before, we redo the earlier example using the new notation. You should notice we get the same numbers as before.

Sample Question

Suppose $\mathbf{v} = \binom{2}{3}$ and $\mathbf{w} = \binom{4}{1}$, as in Figure 6. Calculate $\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$.

Solution. First, compute

$$\operatorname{comp}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|} = \frac{2 \cdot 4 + 3 \cdot 1}{\sqrt{4^2 + 1^2}} = \frac{11}{\sqrt{17}}.$$

Then,

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \operatorname{comp}_{\mathbf{w}}(\mathbf{v}) \frac{\mathbf{w}}{|\mathbf{w}|} = \frac{11}{\sqrt{17}} \begin{pmatrix} \frac{4}{\sqrt{17}} \\ \frac{1}{\sqrt{17}} \end{pmatrix} = \begin{pmatrix} \frac{44}{17} \\ \frac{11}{17} \end{pmatrix}. \qquad \square$$

Let's also do the example in Figure 7.

Sample Question

Suppose $\mathbf{v} = \binom{-3}{3}$ and $\mathbf{w} = \binom{4}{1}$, as in Figure 6. Calculate $\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$.

Solution. First, compute

$$\operatorname{comp}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|} = \frac{-3 \cdot 4 + 3 \cdot 1}{\sqrt{4^2 + 1^2}} = \frac{-9}{\sqrt{17}}.$$

Then,

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \operatorname{comp}_{\mathbf{w}}(\mathbf{v}) \frac{\mathbf{w}}{|\mathbf{w}|} = \frac{-9}{\sqrt{17}} \begin{pmatrix} \frac{4}{\sqrt{17}} \\ \frac{1}{\sqrt{17}} \end{pmatrix} = \begin{pmatrix} -\frac{36}{17} \\ -\frac{9}{17} \end{pmatrix}. \qquad \square$$

§4.5 [EXER] Exercises

Exercise 4.1. In four-dimensional space $\mathbb{R}^4$, the vectors $\langle 1, 2, 3, 4 \rangle$ and $\langle 5, 6, 7, t \rangle$ are perpendicular. Compute t.

Exercise 4.2.

- Compute the vector projection of $\langle 123, 456, 789 \rangle$ in the direction of $\mathbf{e}_1$.
- Compute the scalar component and vector projection of $\mathbf{v} = \langle 1, 2, 3 \rangle$ along the direction of $\mathbf{w} = \langle -3000, -4000, 0 \rangle$.

Exercise 4.3. Let $\mathbf{w} = \langle 3, 4 \rangle$. Compute all unit vectors $\mathbf{v}$ in $\mathbb{R}^2$ for which $\mathbf{v} \cdot \mathbf{w} = 3$.

Exercise 4.4 (*). Determine all possible values of $ax + by + cz$ over real numbers a, b, c, x, y, z satisfying $a^2 + b^2 + c^2 = 2$ and $x^2 + y^2 + z^2 = 5$.

Chapter 5. Planes and their normal vectors

In general, the equation of a plane in $\mathbb{R}^3$ takes the shape

$$ax + by + cz = d$$

where a, b, c, d are real numbers (and a, b, c are not all zero).

§5.1 [TEXT] Normal vectors to planes in $\mathbb{R}^3$

In 18.01, we had lines in $\mathbb{R}^2$, and we used the notion of slope of the line often. For 18.02, planes don't have a "slope"; not a single number, anyway. So the thing I want to communicate is:

Idea

Everything we used slope for in 18.01, we should rephrase in terms of normal vectors for 18.02.

So what's that? A **normal vector** $\mathbf{v}$ to a plane is a vector such that, if you pick any point P on the plane, then the arrow joining P to $P + \mathbf{v}$ — that is, the arrow $\mathbf{v}$ when you draw the starting point as P — is perpendicular to that plane.

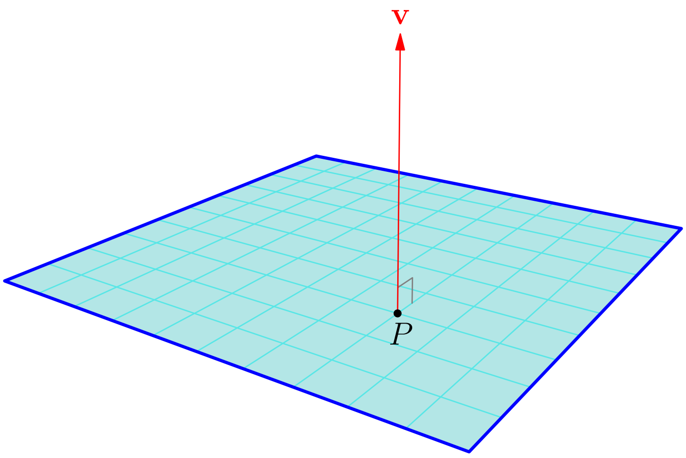

Figure 8: Normal vector to a plane.

(Note that it doesn't matter which point P you pick. You could equally well even ignore P together, imagine drawing $\mathbf{v}$ as an arrow starting from some random point not necessarily on the plane — like the origin — and requiring that $\mathbf{v}$ punctures the plane at a right angle.)

The main goal of this chapter is to prove the following result:

Memorize: Normal vectors of plane

Given a plane $ax + by + cz = d$, the normal vectors to it are the multiples of $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

Keep in mind that normal vectors only matter up to scaling: if $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is a normal vector, then so are $\begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$, $\begin{pmatrix} -100 \\ -200 \\ -300 \end{pmatrix}$, etc.

§5.2 [TEXT] Normal vectors to lines in $\mathbb{R}^2$

Before explaining why this is true in $\mathbb{R}^3$, I want to do everything in $\mathbb{R}^2$ first for comparison, where pictures are easier to draw and you have intuition from eighth or ninth grade algebra.

Here's a question: which vectors in $\mathbb{R}^2$ are perpendicular to $\binom{1}{2}$? They're the vectors lying on a line of slope $-\frac{1}{2}$ through the origin, namely

$$0 = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} \iff 0 = x + 2y.$$

See Figure 9.

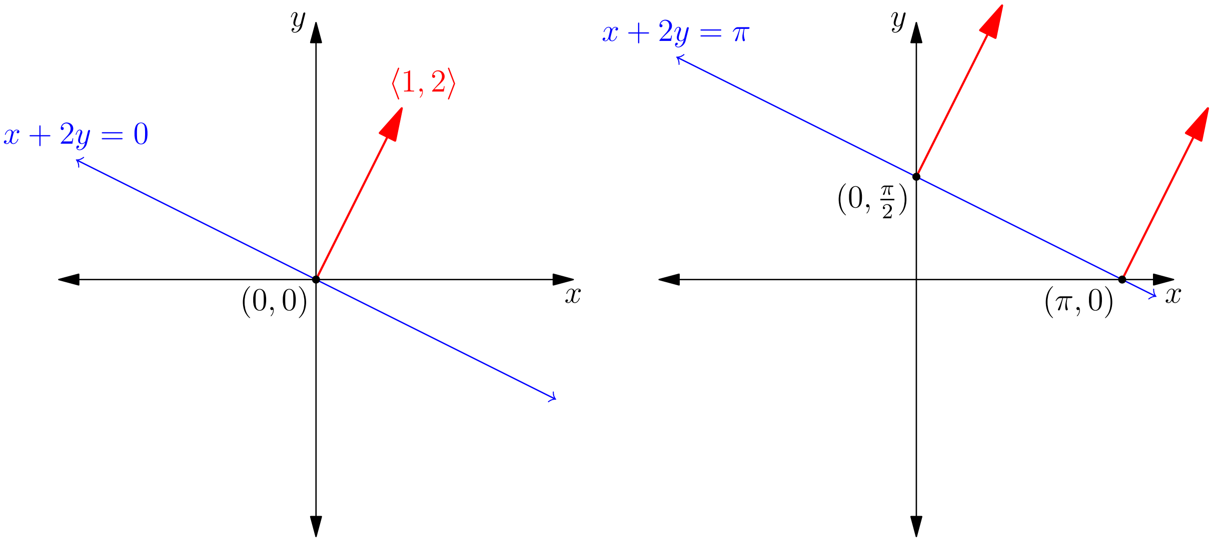

Figure 9: Plots of $x + 2y = 0$ and $x + 2y = \pi$.

Okay, in that case what does the line

$$x + 2y = \pi$$

look like? Well, it's a parallel line, the slope is still the same.

Equivalently, you could also imagine it as the vectors $\binom{x}{y}$ such that

$$\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \pi \\ 0 \end{pmatrix} \quad \text{is perpendicular to} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

or do the same thing for any point on the line, like

$$\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 0 \\ \pi/2 \end{pmatrix} \quad \text{is perpendicular to} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

or even

$$\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 0.218\pi \\ 0.564\pi \end{pmatrix} \quad \text{is perpendicular to} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

But that's silly. Most of the time you don't care about base points. All you care is the line has slope $-\frac{1}{2}$, and for that the LHS just needs to be $x + 2y$ (or even $100x + 200y$). The RHS can be whatever you want.

§5.3 [TEXT] Normal vectors to a plane

In $\mathbb{R}^3$, the exact same thing is true for the expression $ax + by + cz = d$. The only difference is that the word “slope” is banned. Nevertheless, even if we can’t talk about slope, we can still talk about parallel planes, and now the whole discussion carries over wholesale.

For example, what’s the set of vectors perpendicular to $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$? That’s the same as requiring

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

So the plane $x + 2y + 3z = 0$ passes through the origin and has normal vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

What about something like $x + 2y + 3z = 6$? Analogous to last section different ways to write it are:

- Rewriting the equation as $1(x - 6) + 2(y - 0) + 3(z - 0) = 0$, the plane can be thought of as the points (x, y, z) such that
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} \text{ is perpendicular to } \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$
- Rewriting the equation as $1(x - 0) + 2(y - 3) + 3(z - 0) = 0$, the plane can be thought of as the points (x, y, z) such that
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} \text{ is perpendicular to } \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$
- Rewriting the equation as $1(x - 0) + 2(y - 0) + 3(z - 2) = 0$, the plane can be thought of as the points (x, y, z) such that
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \text{ is perpendicular to } \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$
- Rewriting the equation as $1(x - 0.753) + 2(y - 0.618) + 3(z - 1.337) = 0$, the plane can be thought of as the points (x, y, z) such that
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0.753 \\ 0.618 \\ 1.337 \end{pmatrix} \text{ is perpendicular to } \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

But again, like last time, the base point doesn’t really matter. The end story is the same: the coefficients control the direction of the plane via the normal vector (see Figure 10).

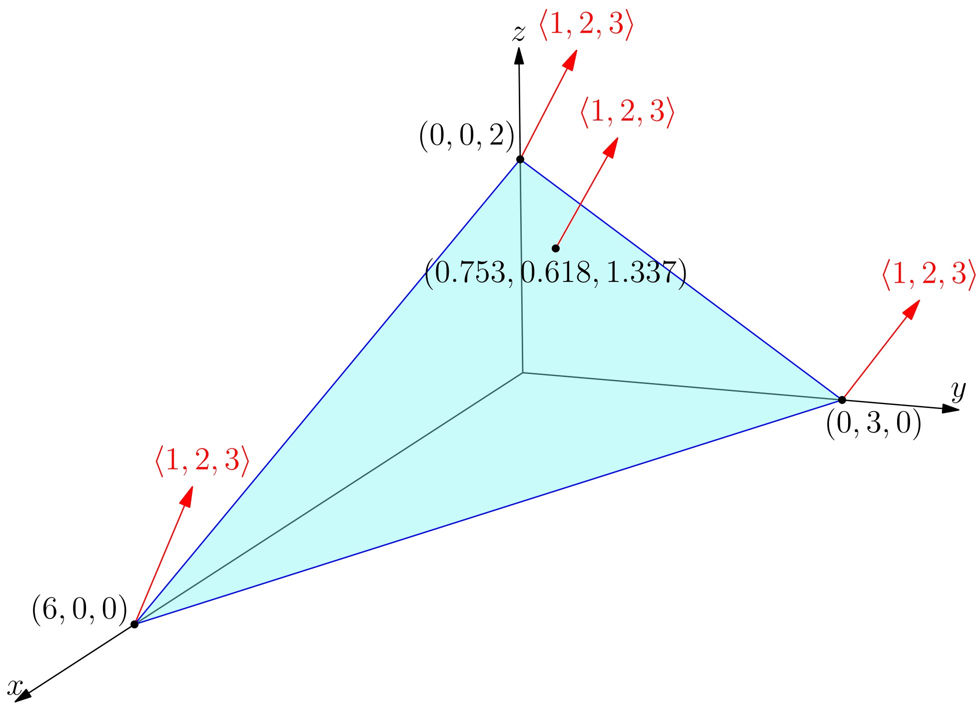

Figure 10: Normal vectors to the plane $x + 2y + 3z = 6$.

§5.4 [RECIPE] Finding a plane through a point with a direction

Sometimes you know the direction the plane goes, and you need to get one point to lie on it. This just means you have to pick the number d to match:

Recipe for finding a plane given a normal vector and a point on it

Suppose the given normal vector is $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, and $P \in \mathbb{R}^3$ is a given point.

1. Write $ax + by + cz$ for the left-hand side.
2. Evaluate the left-hand side at P to get a number d.
3. Output $ax + by + cz = d$.

Sample Question

Compute the equation of the plane parallel to $x + 2y + 3z = 100$ which passes through the point $(1,4,9)$.

Solution

Planes are parallel when they have normal vectors in the same direction, so we use the normal vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ for both. Hence the answer should take the form

$$x + 2y + 3z = d$$

for some d. In order to pass through $(1,4,9)$ we should choose $d = 1 + 2 \cdot 4 + 3 \cdot 9 = 36$. So output $x + 2y + 3z = 36$.

§5.5 [TEXT] Calculating the distance from a point to a plane

There's a classical exercise that's used to test understanding of normal vectors to plane and projections, which is to find the distance from a point to a plane.

Sample Question

Compute the distance from the point $(7, 8, 5)$ to the plane $x + 2y + 3z = 0$.

Solution. The plane $x + 2y + 3z = 0$ has a normal vector $\mathbf{n}$ given by the coefficients of x, y, and z:

$$\mathbf{n} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Now consider the vector $\mathbf{v} = \begin{pmatrix} 7 \\ 8 \\ 5 \end{pmatrix}$ pointing from the origin (which lies on the plane) to the given point $(7, 8, 5)$. The main insight is that the scalar component of $\mathbf{v}$ to the vector $\mathbf{n}$ coincides with the distance we're trying to compute; look at the figure Figure 11 to see why this is true. The point is that there's a rectangle formed by the origin, the endpoint of $\mathbf{v}$, and the projections of $\mathbf{v}$ onto $\mathbf{n}$ and the plane, respectively.

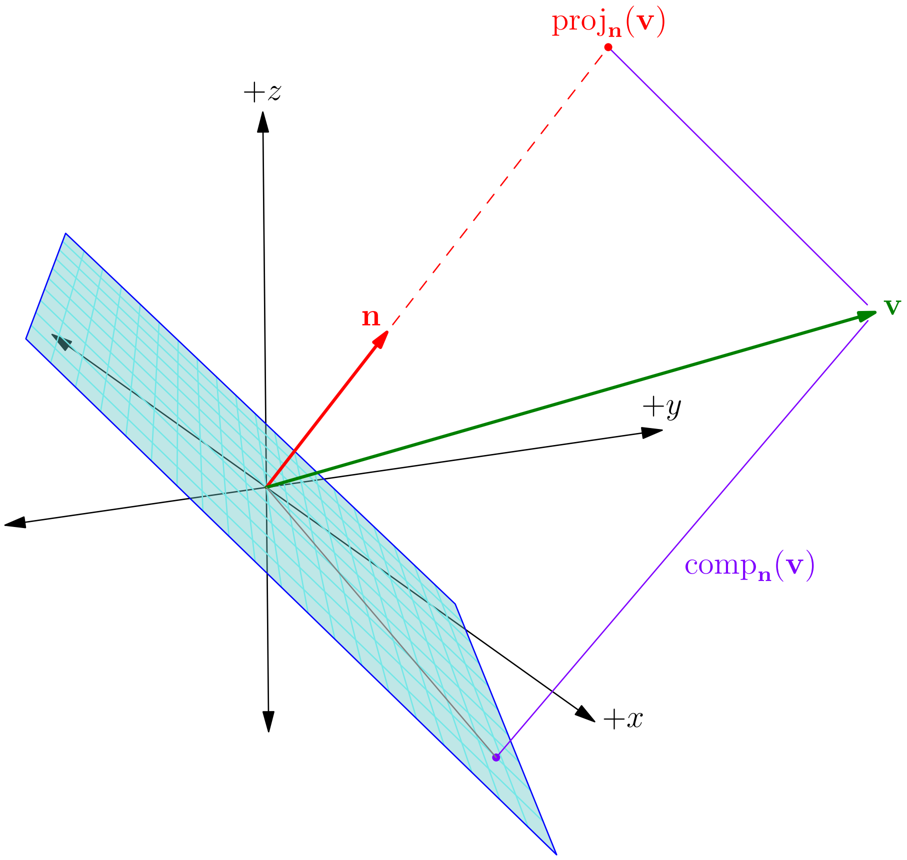

Figure 11: The two projections from $\mathbf{v}$ onto $\mathbf{n}$ and the plane (in purple) form a rectangle, so that the distance from $\mathbf{v}$ to the plane is given exactly by $\mathrm{comp}_{\mathbf{n}}(\mathbf{v})$.

Calculate the dot product:

$$\mathbf{v} \cdot \mathbf{n} = (7)(1) + (8)(2) + (5)(3) = 7 + 16 + 15 = 38.$$

Calculate the magnitude:

$$|\mathbf{n}| = \sqrt{(1)^2 + (2)^2 + (3)^2} = \sqrt{1+4+9} = \sqrt{14}.$$

Hence, the scalar component is:

$$\operatorname{comp}_{\mathbf{n}} \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{n}}{|\mathbf{n}|} = \frac{38}{\sqrt{14}}.$$

This is the answer. □

Here's the same exercise with one change: we change to $x + 2y + 3z = 60$. This means we'll have to pick a point on the plane besides the origin.

Sample Question

Compute the distance from the point $(7, 8, 9)$ to the plane $x + 2y + 3z = 60$.

Solution. As before, the plane $x + 2y + 3z = 60$ has a normal vector $\mathbf{n}$ given by the coefficients of x, y, and z:

$$\mathbf{n} = \langle 1, 2, 3 \rangle.$$

Now we can't use the origin $(0, 0, 0)$ this time, but we can pick any other point on the plane; we'll chose $(0, 0, 20)$. (You could do the problem with $(60, 0, 0)$ or $(0, 30, 0)$ or even $(-77, 13, 37)$ if you prefer; they all give the same answer.)

The vector $\mathbf{v}$ from $(0, 0, 20)$ to $(7, 8, 5)$ is:

$$\mathbf{v} = \langle 7-0, 8-0, 5-20 \rangle = \langle 7, 8, -15 \rangle.$$

Now we can just repeat the steps from before, where

$$\begin{aligned} \mathbf{v} \cdot \mathbf{n} &= (7)(1) + (8)(2) + (-15)(3) = -22 \\ |\mathbf{n}| &= \sqrt{(1)^2 + (2)^2 + (3)^2} = \sqrt{1+4+9} = \sqrt{14}. \end{aligned}$$

Hence

$$\operatorname{comp}_{\mathbf{n}} \mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{n}}{|\mathbf{n}|} = \frac{-22}{\sqrt{14}}.$$

The distance is the absolute value $\frac{22}{\sqrt{14}}$. □

Remark

It's fine that you get a negative number for the scalar component. This corresponds to the fact that the point $(7, 8, 9)$ is sandwiched between the two planes $x + 2y + 3z = 0$ and $x + 2y + 3z = 60$. Depending on which way you choose to point $\mathbf{n}$, one of the components will be positive and the other negative. See Figure 12.

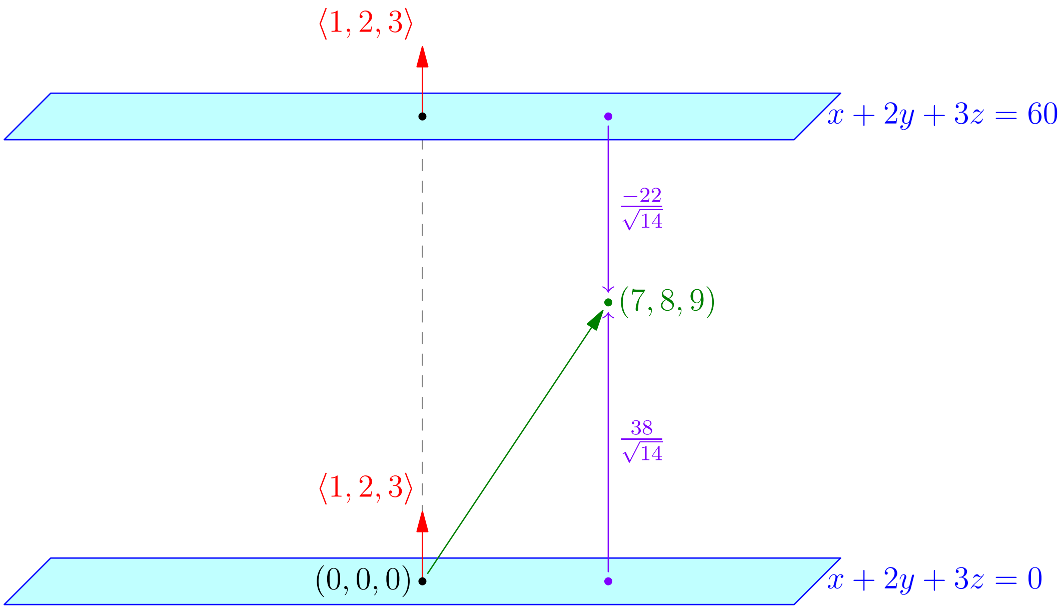

Figure 12: A 2D cartoon of the point $(7,8,9)$ sandwiched between the planes $x+2y+3z=0$ and $x+2y+3z=60$. If we choose $\mathbf{n}=\langle 1,2,3\rangle$ then we get $+\frac{38}{\sqrt{14}}$ and $-\frac{22}{\sqrt{14}}$ for the scalar components as shown in purple.

The thing about this exercise is that you can just do it with symbols instead of numbers and get a general formula, which means that doing it with specific numbers is sort of a fool's errand. Let's just do them all at once.

Sample Question

Compute the distance from a point (x_0, y_0, z_0) to the plane defined by the equation

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

Give the answer in terms of a, b, c, d, x_0, y_0, z_0.

Solution. The normal vector $\mathbf{n}$ to the plane is given by the coefficients of x, y, and z in the plane equation:

$$\mathbf{n}=\langle a,b,c\rangle.$$

Now we select any base point (x_1, y_1, z_1) that lies on the plane $\mathcal{P}$. We'll do the case $c\neq 0$ and use

$$(x_1,y_1,z_1)=\left(0,0,\frac{d}{c}\right)$$

but the other cases $a\neq 0$ and $b\neq 0$ are done in the same way. (In fact, you don't really need to pick the base point either, it just makes the calculation a bit easier to think about in what follows.)

The vector $\mathbf{v}$ from (x_1, y_1, z_1) to (x_0, y_0, z_0) is:

$$\mathbf{v}=\langle x_0-x_1, y_0-y_1, z_0-z_1\rangle=\langle x_0, y_0, z_0-\frac{d}{c}\rangle.$$

Now, the scalar component of $\mathbf{v}$ along $\mathbf{n}$ is given by

$$\operatorname{comp}_{\mathbf{n}}\mathbf{v}=\frac{\mathbf{v}\cdot\mathbf{n}}{|\mathbf{n}|}.$$

We compute it. Compute the dot product $\mathbf{v} \cdot \mathbf{n}$:

$$\mathbf{v} \cdot \mathbf{n} = ax_0 + by_0 + c\left(z_0 - \frac{d}{c}\right) = ax_0 + by_0 + cz_0 - d.$$

Compute the magnitude of $\mathbf{n}$:

$$|\mathbf{n}| = \sqrt{a^2 + b^2 + c^2}.$$

Therefore, the scalar component is:

$$\operatorname{comp}_{\mathbf{n}} \mathbf{v} = \frac{ax_0 + by_0 + cz_0 - d}{\sqrt{a^2 + b^2 + c^2}}.$$

Finally, the distance from the point to the plane is the absolute value of the scalar component:

$$|\operatorname{comp}_{\mathbf{n}} \mathbf{v}| = \left|\frac{ax_0 + by_0 + cz_0 - d}{\sqrt{a^2 + b^2 + c^2}}\right|.$$

□

§5.6 [RECIPE] Distance to a plane

If you just want to memorize the final result, here it is:

Recipe for distance from point to plane

If asked to find the distance from a point (x_0, y_0, z_0) to the plane defined by the equation $ax + by + cz = d$, output the answer

$$\frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

(We've moved the absolute value to the numerator, since the square root is always positive.)

§5.7 [EXER] Exercises

Exercise 5.1. A cube is drawn somewhere in $\mathbb{R}^3$ (its faces are not parallel to the coordinate axes). Two of the faces of the cube are contained in the planes $x + 2y + 3z = 4$ and $5x + 6y + kz = 7$, respectively, for some real number k. Given this information, compute k.

Exercise 5.2. The distance from a certain point P to the plane $3x + 4y + 12z = -1$ is 42. What are the possible distances from P to the plane $3x + 4y + 12z = 1000$?

Chapter 6. The cross product

The cross product is the last major linear algebra tool we'll need to introduce (together with determinants and the dot product). Like the dot product, the cross product also has two definitions, one algebraic and one geometric.

However, unlike the dot product, the cross product is more stilted and unnatural, and not used as much — in fact it won't show up again until Chapter 29. (More on that in Section 6.4.) I'll try to keep this chapter brief.

§6.1 [TEXT] The two definitions of the cross product

This definition is terrible, so bear with me.

Definition

Suppose $\mathbf{v} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ are two vectors in $\mathbb{R}^3$.

The *algebraic definition* of the cross product is:

$$\mathbf{v} \times \mathbf{w} := \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{e}_3 = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \tag{2}$$

(See the tip below for a way to remember this formula more easily.)

The *geometric definition* of the cross product is based on specifying both the direction and magnitude.

- The **magnitude** of $\mathbf{v} \times \mathbf{w}$ is equal to the area of the parallelogram formed by $\mathbf{v}$ and $\mathbf{w}$. In trigonometry turns, if θ is the included angle, this equals $|\mathbf{v}|\ |\mathbf{w}| \sin\theta$.
- The **direction** is given by requiring $\mathbf{v} \times \mathbf{w}$ to be perpendicular to *both* $\mathbf{v}$ and $\mathbf{w}$, and also satisfying the **right-hand rule** (we'll say more about it in a moment).

The geometric definition is illustrated in Figure 13 below.

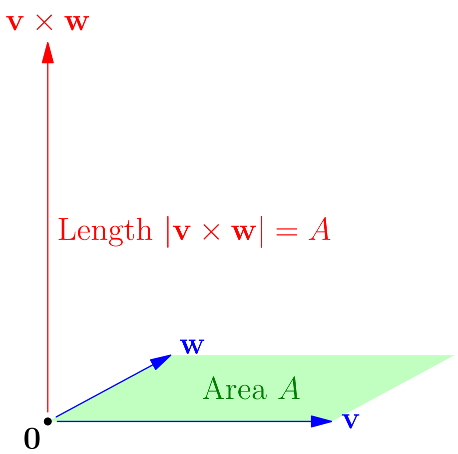

Figure 13: Illustration of the geometric definition of the cross product.

Like with the dot product, it's not obvious at all why these definitions are compatible! Equation 2 is probably also really mysterious and seems to come from nowhere. In this case, I think the idea is that you should start with the geometric definition, then grind through some calculation to get a system of equations. If you solve the system of equations, you wind up with Equation 2 as the result. We do this in the next section as an optional sidenote.

</> **Type signature**

The cross product **only** accepts two vectors **both in** $\mathbb{R}^3$. And it outputs a single **vector in** $\mathbb{R}^3$.

i **Remark: The right-hand rule**

The hack with the right-hand rule is necessary because if I tell you only the length of a vector in $\mathbb{R}^3$ and that it is normal to two other vectors in $\mathbb{R}^3$, there are actually two vectors that work. (For example, there are two vectors of length 5 perpendicular to $\mathbf{e}_1$ and $\mathbf{e}_2$: namely $\pm 5\mathbf{e}_3$.)

So we need to pick one, and the right-hand rule says that if you point your right index finger along $\mathbf{v}$ and right middle finger along $\mathbf{w}$ closer to your palm, and stick out your right thumb, then $\mathbf{v} \times \mathbf{w}$ points along your thumb.

Another way to describe the right-hand rule is to require the following table to be true:

$$\begin{aligned} \mathbf{e}_1 \times \mathbf{e}_2 &= \mathbf{e}_3 = -\mathbf{e}_2 \times \mathbf{e}_1 \\ \mathbf{e}_2 \times \mathbf{e}_3 &= \mathbf{e}_1 = -\mathbf{e}_3 \times \mathbf{e}_2 \\ \mathbf{e}_3 \times \mathbf{e}_1 &= \mathbf{e}_2 = -\mathbf{e}_1 \times \mathbf{e}_3. \end{aligned}$$

It may not be that easy to remember Equation 2. In practice, I think almost everyone uses the following mnemonic for it.

🔥 **Tip: How to remember the algebraic cross product definition**

The algebraic definition is usually remembered using the following mnemonic:

$$\mathbf{v} \times \mathbf{w} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \tag{3}$$

Mathematically speaking, the right-hand side doesn't make sense and is a type-error, because one can't have a matrix where some things in it are numbers and other things in it are vectors. However, if you ignore that and multiply anyway, you'll get the algebraic definition above.

In these notes **I will always use Equation 3 rather than Equation 2** because that's what people actually do in practice. (I do so quite grudgingly, because Equation 3 is officially a type-error, and in theory it is nonsense. But in the words of Linus Torvalds: "Theory and practice sometimes clash. And when that happens, theory loses. Every single time.")

Warning: Cross product is anti-commutative

From either definition, you should be able to see that

$$\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$$

in contrast to the dot product. Note the minus sign. (The right hand rule means that you can't swap your index and middle finger.)

Also, note that $\mathbf{v} \times \mathbf{v} = \mathbf{0}$ (or indeed $\mathbf{v} \times \mathbf{w} = \mathbf{0}$ whenever $\mathbf{v}$ and $\mathbf{w}$ are parallel).

I really want to get this section over with so I'll just give you one example with numbers and not even talk about the corresponding geometry.

Sample Question

Compute the cross product of $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.

Solution. Write

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &:= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} \\ &= \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \mathbf{e}_3 \\ &= -3\mathbf{e}_1 + 6\mathbf{e}_2 - 3\mathbf{e}_3 = \begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix}. \end{aligned}$$

□

As a sanity check for the geometry definition, you can verify that indeed this vector is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$ using the dot product:

$$\begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (-3)(1) + (6)(2) + (-3)(3) = 0$$
$$\begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = (-3)(4) + (6)(5) + (-3)(6) = 0.$$

§6.2 [SIDENOTE] Outline of derivation of the cross product formula

In this sidenote, we outline how the algebraic definition of the cross product can be derived from the geometric one. This proof is too hard to be on an 18.02 exam, but I decided to include it here instead of the appendix because it is actually good practice with using dot products and determinants. For simplicity, I won't worry at all about division-by-zero issues.

So let's say we have two given vectors

$$\begin{aligned} \mathbf{v} &= a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \\ \mathbf{w} &= b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3. \end{aligned}$$

Let's denote the cross product by

$$\mathbf{v} \times \mathbf{w} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3.$$

We need to solve for x, y, z in terms of a_1, a_2, a_3, b_1, b_2, b_3.

So what are the givens? Well, first we have the requirement that $\mathbf{v} \times \mathbf{w}$ should be perpendicular to $\mathbf{v}$ and $\mathbf{w}$. So we get two equations from the dot product:

$$\begin{aligned} 0 &= \mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) = a_1 x + a_2 y + a_3 z \\ 0 &= \mathbf{w} \cdot (\mathbf{v} \times \mathbf{w}) = b_1 x + b_2 y + b_3 z. \end{aligned}$$

This is a system of two equations in three variables, but it's good enough to get the ratio $x : y : z$. For example, if you multiply the first equation by b_3 and the second by a_3, then subtract, you will get that

$$0 = b_3(a_1 x + a_2 y) - a_3(b_1 x + b_2 y) \implies \frac{x}{y} = \frac{a_2 b_3 - a_3 b_2}{a_3 b_1 - a_1 b_3}.$$

In a similar way, you can get the ratio $\frac{y}{z}$ as

$$\frac{y}{z} = \frac{a_3 b_1 - a_1 b_3}{a_1 b_2 - a_2 b_1}.$$

Geometrically, this means we've already recovered the *direction* of $\mathbf{v} \times \mathbf{w}$. Algebraically, it means we know the ratio $x : y : z$; there should be some scalar constant k such that

$$\begin{aligned} x &= k(a_2 b_3 - a_3 b_2) \\ y &= k(a_3 b_1 - a_1 b_3) \\ z &= k(a_1 b_2 - a_2 b_1). \end{aligned}$$

So all we need to do now is find k. For brevity, we'll let

$$\begin{aligned} x_0 &:= a_2 b_3 - a_3 b_2 \\ y_0 &:= a_3 b_1 - a_1 b_3 \\ z_0 &:= a_1 b_2 - a_2 b_1 \end{aligned}$$

so $x = kx_0$, $y = ky_0$, $z = kz_0$, and we need to solve for k.

We have one more condition to use, which is that we want the magnitude $\sqrt{x^2 + y^2 + z^2}$ to be equal to to the area A of the parallelogram spanned by $\mathbf{v}$ and $\mathbf{w}$, and also point in the way specified by the right-hand rule. How can we encode that in an equation? At face value, none of the weapons in our toolkit so far let you access the area of a parallelogram in 3D space. We only have a determinant for parallelograms in 2D space.

But here's a clever trick to get around it. The trick to get the area is to consider the parallelepiped formed by three vectors: the two given vectors $\mathbf{v}$ and $\mathbf{w}$ and the *unit* vector in the direction of $\mathbf{v} \times \mathbf{w}$. That is, consider the vector

$$\mathbf{n} = \frac{\langle x_0, y_0, z_0 \rangle}{\sqrt{x_0^2 + y_0^2 + z_0^2}}.$$

We've just seen $\mathbf{n}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$, and we've scaled $\mathbf{n}$ so that $|\mathbf{n}| = 1$. See Figure 14.

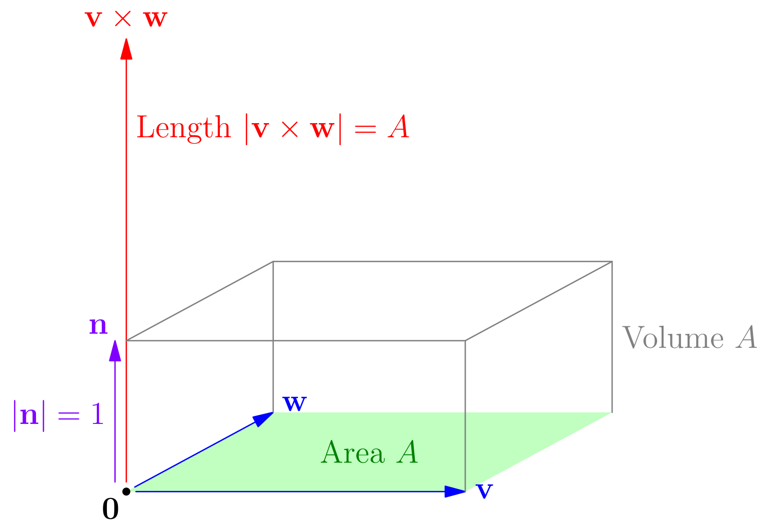

Figure 14: The additional vector $\mathbf{n}$ introduced. It's a unit vector in the direction we want, and it doesn't depend on k.

For simplicity let's assume that $\mathbf{n}$ points the correct way for the right-hand rule. (If $\mathbf{n}$ points the other way, the calculation below will need a bunch of extra minus signs; we won't dwell on it here, again to keep things simple.) So if we consider the parallelepiped formed by $\mathbf{n}$, $\mathbf{v}$, $\mathbf{w}$, its volume will just be the height times the base parallelogram, i.e. it is $1 \cdot A = A$. And the volume of a parallelepiped *is* something we can access: it's a 3×3 determinant. So our trick has managed to let us get our hands on A:

$$A = \begin{vmatrix} \frac{x_0}{\sqrt{x_0^2+y_0^2+z_0^2}} & \frac{y_0}{\sqrt{x_0^2+y_0^2+z_0^2}} & \frac{z_0}{\sqrt{x_0^2+y_0^2+z_0^2}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Here, the determinant is indeed $+A$ (rather than $-A$) because $\mathbf{n}$ was assumed to be pointing the correct direction, so the determinant yields a plus sign. This trick of introducing $\mathbf{n}$ is what lets us get a formula for A which would otherwise be inaccessible with only the determinant at face value.

Now we solve for k. Because we're in the case that $\mathbf{n}$ is pointing the right way, we know we should have $k > 0$. Now we have to set the magnitude of

$$|\mathbf{v} \times \mathbf{w}| = k\sqrt{x_0^2 + y_0^2 + z_0^2}$$

equal to A above; that is, we get the equation

$$k\sqrt{x_0^2 + y_0^2 + z_0^2} = A = \begin{vmatrix} \frac{x_0}{\sqrt{x_0^2+y_0^2+z_0^2}} & \frac{y_0}{\sqrt{x_0^2+y_0^2+z_0^2}} & \frac{z_0}{\sqrt{x_0^2+y_0^2+z_0^2}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Multiplying both sides by $\sqrt{x_0^2 + y_0^2 + z_0^2}$ and expanding the determinant gives

$$k(x_0^2 + y_0^2 + z_0^2) = x_0 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - y_0 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + z_0 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

If you now look at the definition of x_0, y_0, z_0, you will see that the determinants on the right-hand side are conveniently equal to x_0, $-y_0$, and z_0, respectively. So suddenly the whole thing cancels and we just get $k = 1$. So $x = x_0$, $y = y_0$, $z = z_0$ and this gives the algebraic formula we wanted.

§6.3 [RECIPE] What to use the cross product for

Unlike the dot product, which is just a number, the cross product is a vector. So it has more information in it — both a direction and a magnitude.

- The direction of $\mathbf{v} \times \mathbf{w}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$.
- The magnitude is the area of the parallelogram.

However in practice, when we use the cross product, we'll often *only use one piece of information.* (It's not until Chapter 38 that we really start using both parts at once.)

Hence the following two recipes below.

Recipe for normal vectors

To find a vector perpendicular to both $\mathbf{v}$ and $\mathbf{w}$ at once:

1. Output any nonzero multiple of $\mathbf{v} \times \mathbf{w}$.

Recipe for area

To find the area of the parallelogram formed by $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$:

1. Output the magnitude of $\mathbf{v} \times \mathbf{w}$.

Notice in the first recipe, we ignore the magnitude; in the second recipe, we ignore the direction.

Sample Question

Consider the three points $A = (1, 0, 0)$, $B = (0, 2, 0)$, $C = (0, 0, 3)$.

- Find a normal vector to the plane through A, B, C.
- Compute the equation of the plane.
- Compute the area of triangle ABC.

Solution. First, let's find a normal vector to the plane through A, B, and C. The idea is to compute two vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$:

$$\overrightarrow{AB} = \begin{pmatrix} 0-1 \\ 2-0 \\ 0-0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$
$$\overrightarrow{AC} = \begin{pmatrix} 0-1 \\ 0-0 \\ 3-0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}.$$

These two vectors can be drawn as arrows contained in the plane through them. So if we compute the cross product, we'll get a normal vector we wanted! That is,

$$\begin{aligned}\overrightarrow{AB} \times \overrightarrow{AC} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -1 & 2 & 0 \\ -1 & 0 & 3 \end{vmatrix} \\ &= (2 \cdot 3 - 0 \cdot 0)\mathbf{e}_1 - (-1 \cdot 3 - 0 \cdot -1)\mathbf{e}_2 + (-1 \cdot 0 - 2 \cdot -1)\mathbf{e}_3 \\ &= (6 - 0)\mathbf{e}_1 - (-3 - 0)\mathbf{e}_2 + (0 - (-2))\mathbf{e}_3 = \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix}.\end{aligned}$$

That's the normal vector. To find the equation of the plane, we know that we should have

$$6x + 3y + 2z = d$$

for some constant d. Plugging in any of the three points A, B, C gives $d = 6$ (the redundancy here gives us a way to check our arithmetic, too). So the plane is

$$6x + 3y + 2z = 6.$$

Finally, the area of $\triangle ABC$ is half the area of the parallelogram formed by $\overrightarrow{AB}$ and $\overrightarrow{AC}$, so that

$$\text{Area}(\triangle ABC) = \frac{1}{2}\left|\overrightarrow{AB} \times \overrightarrow{AC}\right| = \frac{1}{2}\sqrt{6^2 + 3^2 + 2^2} = \frac{7}{2}.$$ □

> **Remark**
>
> This shape of question is worth remembering: the cross product often gives you a way to find a normal vector to some plane, because it's so good at making right angles. Then once you have the normal vector, you can find the equation of the plane using the recipe from Section 5.4.

§6.4 [SIDENOTE] The cross product sucks

Compared to dot products and determinants, the cross product might feel the most unnatural, for good reason — it's used much less frequently by serious mathematicians than the other tools you see. See Figure 15.

The reason that the cross product isn't popular with mathematicians is the definition of the cross product is **really quite brittle**. For example, the cross product can't be defined for any number of dimensions,[11] and you have to remember this weird right-hand rule that adds one more arbitrary convention. So the definition is pretty unsatisfying.

To replace the cross product, mathematicians use a different kind of object called a *bivector*, an element of a space called $\bigwedge^2(\mathbb{R}^n)$. (They might even claim that bivectors do everything cross products can do, but better.) Again, this new kind of object is well beyond the scope of 18.02 but it's documented in Chapter 12 of my Napkin if you do want to see it.

I'll give you a bit of a teaser though. In general, for any n, bivectors in $\mathbb{R}^n$ are specified by $\frac{n(n-1)}{2}$ coordinates. So for $n = 3$ you *could* translate every bivector in $\mathbb{R}^3$ into a vector in $\mathbb{R}^3$ by just reading the coordinates (although you end up with the right-hand rule as an artifact of the translation), and the cross product is exactly what you get. But for $n = 4$, a bivector in $\mathbb{R}^4$ has six numbers, which is too much information to store in a vector in $\mathbb{R}^4$. Similarly, for $n > 4$, this translation can't be done. That's why the cross product is so brittle and can't work past $\mathbb{R}^3$.

[11] Just kidding, apparently there's a seven dimensional cross product? Today I learned. Except that there are apparently 480 different ways to define it in seven dimensions, so, like, probably not a great thing.

Figure 15: How to think of cross products.

§6.5 [RECAP] Recap of vector stuff up to here

A brief summary of the last few chapters.

- The dot and cross products have algebraic formulas and geometric properties that make them useful in a lot of 3D geometry applications.
- The dot product lets you detect perpendicularity and projections.
  - Two vectors are perpendicular if and only if their dot product is zero.
- The cross products generates perpendicularities and lets you compute area.
- Both are used in the theory of planes:
  - We use the dot product to show that the normal vector to the plane $ax + by + cz = d$ was the vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$.
  - We use the projection from the dot product to find the distance from a point to a plane.
  - Given three points on a plane, the cross product let us find the normal vector.

See also Table 2, which summarizes some of the vectors we've seen in applications.

Vector	Direction	Magnitude
Normal vector $\mathbf{n}$ to plane	Perpendicular to plane	*Irrelevant!*
$\operatorname{proj}_{\mathbf{w}}(\mathbf{v})$	Same as $\mathbf{w}$	Scalar component
Cross product $\mathbf{v} \times \mathbf{w}$	Perpendicular to both $\mathbf{a}$ and $\mathbf{b}$	Area of parallelogram

Table 2: Some commonly used kinds of vectors we've met so far.

§6.6 [EXER] Exercises

Exercise 6.1. Suppose real numbers a and b satisfy

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 100 \\ a \\ b \end{pmatrix} = \mathbf{0}.$$

Compute a and b.

Exercise 6.2. Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^3$ for which $\mathbf{v} \times \mathbf{w} = \left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right)$. Compute $5\mathbf{w} \times 4\mathbf{v}$.

Exercise 6.3. Let $\mathbf{v}$ and $\mathbf{w}$ be unit vectors in $\mathbb{R}^3$. Compute all possible values of

$$|\mathbf{v} \times \mathbf{w}|^2 + (\mathbf{v} \cdot \mathbf{w})^2.$$

Exercise 6.4. Suppose $\mathbf{v}$ is a vector in $\mathbb{R}^3$ and k is a real number such that

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ k \end{pmatrix}.$$

Compute k.

Part Bravo: Linear Algebra of Matrices

For comparison, Part Bravo corresponds roughly to §3, §4, §6 of Poonen's notes.

Chapter 7. Linear transformations and matrices

The goal of this chapter is to tell you how to take a linear transformation and encode it as a matrix. In other words, there is only one recipe covered. However, one upshot of this presentation is that I'll finally be able to explain why matrix multiplication is defined in this way you learned.

This chapter will be presented a bit differently than you'll see in many other places; I talk about linear transformations first, and then talk about matrices as an encoding of linear transformations. I feel quite strongly that this way is better, but if you are in an actual course, their presentation is likely to be different (and worse).

§7.1 [TEXT] Linear transformation

The definition I'm about to give is the 18.700/18.701 definition of linear transform, but the hill I will die on is that this definition is better than the one in 18.02.

Definition of linear transformation

A *linear transform* $T : \mathbb{R}^n \to \mathbb{R}^m$ is *any* map obeying the two axioms $T(c\mathbf{v}) = cT(\mathbf{v})$ and $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$.

So it's a chonky boy: for every $\mathbf{v} \in \mathbb{R}^n$, there's an output value $T(\mathbf{v}) \in \mathbb{R}^m$. I wouldn't worry too much about the axioms until later; for now, read the examples.

Examples of linear transformations

The following are all linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$:

- The constant function where $T(\mathbf{v}) = \mathbf{0}$ for every vector v
- Projection onto the x-axis: $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ 0 \end{pmatrix}$.
- Rotation by an angle
- Reflection across a line
- Projection onto the line $y = x$.
- Multiplication by any 2×2 matrix, e.g. the formula

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix}$$

  is a linear transformation too.

Tip

Note that $T(\mathbf{0}) = \mathbf{0}$ in any linear transformation.

The important principle to understand is that if you know the values of a transformation T at enough points, you can recover the rest.

Here's an easy example to start:

Sample Question

If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

what are the vectors for $T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 203 \\ 204 \end{pmatrix}\right)$?

Solution.

$$\begin{aligned} T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right) &= \begin{pmatrix} \pi \\ 9 \end{pmatrix} + \begin{pmatrix} 0 \\ 12 \end{pmatrix} = \begin{pmatrix} \pi \\ 21 \end{pmatrix} \\ T\left(\begin{pmatrix} 203 \\ 204 \end{pmatrix}\right) &= \begin{pmatrix} \pi \\ 9 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 12 \end{pmatrix} = \begin{pmatrix} \pi \\ 33 \end{pmatrix}. \end{aligned}$$

□

Here's another example.

Sample Question

If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

what is $T\left(\begin{pmatrix} 50 \\ 70 \end{pmatrix}\right)$?

Solution.

$$T\left(\begin{pmatrix} 50 \\ 70 \end{pmatrix}\right) = 50\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 70\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 190 \\ 430 \end{pmatrix}.$$

□

More generally, the second question shows that if you know $T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$ you ought to be able to *calculate* the output of T at any other vector like $\begin{pmatrix} 50 \\ 70 \end{pmatrix}$. To expand on this:

$$T\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = aT\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + bT\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right). \tag{4}$$

More generally, from understanding the solution to the above two questions, you should understand the following important statement that we'll use over and over.

Memorize

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. If you know the outputs T on a basis, then you can deduce the value of T at any other input.

For now "basis" refers to just the n vectors $\mathbf{e}_1$, ..., $\mathbf{e}_n$. But later on we will generalize this notion to some other settings too.

§7.2 [RECIPE] Matrix encoding

A *matrix* is a way of *encoding* the *outputs* of T using as few numbers as possible. That is:

Definition

A matrix **encodes all outputs** of a linear transformation T by **writing the outputs** of $T(\mathbf{e}_1)$, ..., $T(\mathbf{e}_n)$ as a list of **column vectors**.

For example, if you had $T : \mathbb{R}^2 \to \mathbb{R}^2$ with

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}2\\4\end{pmatrix} \iff T \quad \text{encoded as} \quad \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}.$$

To put this into recipe form:

Recipe for encoding a transformation

Given a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, to encode it as a matrix:

1. Compute $T(\mathbf{e}_1)$ through $T(\mathbf{e}_n)$ and write them as column vectors..
2. Glue them together to get an $m \times n$ array of numbers.

Here's more examples.

Sample Question

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the x-axis. Write T as a 2×2 matrix.

Solution. Note that

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix} \qquad T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}0\\0\end{pmatrix}.$$

Glue these together and output T as the matrix

$$T = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}.$$

□

Remark

You might note that indeed multiplication by the encoded matrix

$$\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\0\end{pmatrix}.$$

matches what you expect: $\begin{pmatrix}x\\0\end{pmatrix}$ is indeed the projection of $\begin{pmatrix}x\\y\end{pmatrix}$ onto the x-axis! And this works for every linear transformation. This is so important I'll say it again next section, just mentioning it here first.

Sample Question

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection around the line $y = x$. Write T as a 2×2 matrix.

Solution. Note that

$$T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}0\\1\end{pmatrix} \qquad T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix}.$$

Glue these together and output T as the matrix

$$T = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}.$$

□

Sample Question

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation around the origin by $30°$. Write T as a 2×2 matrix.

Solution. See Figure 16. By looking at the unit circle, we see that

$$T(\mathbf{e_1}) = \begin{pmatrix}\cos 30°\\ \sin 30°\end{pmatrix} = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right).$$

The vector $\mathbf{e}_2$ is $90°$ further along

$$T(\mathbf{e_2}) = \begin{pmatrix}\cos 120°\\ \sin 120°\end{pmatrix} = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right).$$

Glue these together and output T as the matrix

$$T = \begin{pmatrix}\frac{\sqrt{3}}{2} & -\frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}.$$

□

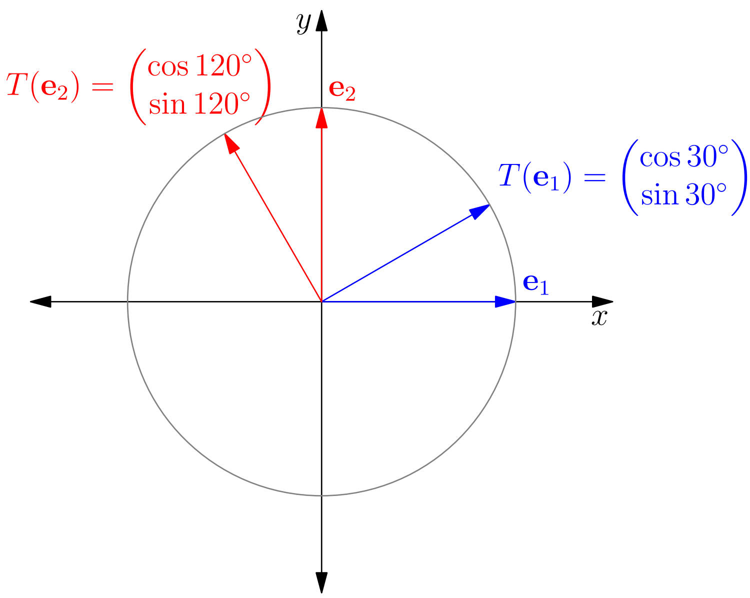

Figure 16: Rotation by 30 degrees.

Remark: This is where the rotation matrix comes from

If you redo this question with 30° replaced by any angle θ, you get the answer

$$T = \begin{pmatrix} \cos\theta & \cos(\theta + 90^\circ) \\ \sin\theta & \sin(\theta + 90^\circ) \end{pmatrix}.$$

So this is the matrix that corresponds to rotation. However, in the literature you will often see this rewritten as

$$T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

to get rid of the $+90^\circ$ offsets. That's fine, but I think it kind of hides where the formula for rotation matrix comes from, personally.

Another example is the identity function:

Example: The identity matrix deserves its name

Let $I : \mathbb{R}^3 \to \mathbb{R}^3$ denote the 3D identity function, meaning $I(\mathbf{v}) = \mathbf{v}$. To encode it, we look at its values at $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$I(\mathbf{e}_1) = \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad I(\mathbf{e}_2) = \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad I(\mathbf{e}_3) = \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

We encode it as a matrix by writing the columns side by side, getting what you expect:

$$I \quad \text{encoded as} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This gives a more natural reason why the identity matrix is the one with 1's on the diagonal and 0's elsewhere (compared to the "well try multiplying by it" you learned in high school).

§7.3 [SIDENOTE] Matrix multiplication

In the prerequisites, I said that you were supposed to know the rule for multiplying matrices, so you should already know for example that

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

The goal of this side note is to now explain why matrix multiplication is defined in this funny way. We will see two results:

- Multiplication of the matrix for T by a column vector $\mathbf{v}$ corresponds to evaluation $T(\mathbf{v})$.
- Multiplication of the matrices for S and T gives the matrix for the composed function $S \circ T$.[12]

[12]The $\circ$ symbol means the function where you apply T first then S first. So for example, if $f(x) = x^2$ and $g(x) = x + 5$, then $(f \circ g)(x) = f(g(x)) = (x + 5)^2$. We mostly use that circle symbol if we want to refer to $f \circ g$ itself without the x, since it would look really bad if you wrote "$f(g$" or something.

§7.3.1 One matrix

Recall from the example in Section 7.1 that if T was the linear transformation for which

$$T(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

then

$$T\left(\begin{pmatrix} 50 \\ 70 \end{pmatrix}\right) = \begin{pmatrix} 190 \\ 430 \end{pmatrix}.$$

We just now also saw that to encode T as a matrix, we have

$$T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Now, what do you think happens if you compute

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 50 \\ 70 \end{pmatrix}$$

as you were taught in high school? Surprise: you get $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \left(\begin{smallmatrix} 50 \\ 70 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \cdot 50 + 2 \cdot 70 \\ 3 \cdot 50 + 4 \cdot 70 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 190 \\ 340 \end{smallmatrix}\right)$ which is not just the same answer, but also the same intermediate calculations. In other words,

> **Idea**
>
> If one multiplies a matrix M by a column vector $\mathbf{v}$, this corresponds to applying the linear transformation T encoded by M to $\mathbf{v}$.

§7.3.2 Two matrices

Now, any time we have functions in math, we can *compose* them. So let's play the same game with a pair of functions S and T, and think about their composition $S \circ T$. Imagine we got asked the following question:

Question 7.1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transform such that

$$T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

Then let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transform such that

$$S\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ 7 \end{pmatrix} \quad \text{and} \quad S\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 6 \\ 8 \end{pmatrix}.$$

Evaluate $S\left(T\left(\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\right)\right)$ and $S\left(T\left(\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\right)\right)$.

Solution.

$$S\left(T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)\right) = S\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}\right) = 1\begin{pmatrix} 5 \\ 7 \end{pmatrix} + 3\begin{pmatrix} 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 23 \\ 31 \end{pmatrix}$$

$$S\left(T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right) = S\left(\begin{pmatrix} 2 \\ 4 \end{pmatrix}\right) = 2\begin{pmatrix} 5 \\ 7 \end{pmatrix} + 4\begin{pmatrix} 6 \\ 8 \end{pmatrix} = \begin{pmatrix} 34 \\ 46 \end{pmatrix}.$$

□

Now, $S \circ T$ is *itself* a function, so it makes sense to encode $S \circ T$ as a matrix too, using the answer to Question 7.1:

$$S\left(T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)\right) = \begin{pmatrix} 23 \\ 31 \end{pmatrix} \text{ and } S\left(T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right) = \begin{pmatrix} 34 \\ 46 \end{pmatrix} \iff S \circ T \text{ encoded as } \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

The matrix multiplication rule is then rigged to give the same answer through the same calculation again:

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

In other words:

Idea

If one multiplies two matrices M and N, this corresponds to composing the linear transformations that M and N encode.

This shows why the 18.700/18.701 definitions are better than the 18.02 ones. In 18.02, the recipe for matrix multiplication is a *definition*: "here is this contrived rule about taking products of columns and rows, trust me bro". But in 18.700/18.701, the matrix multiplication recipe is a *theorem*; it's what happens if you generalize Question 7.1 to eight variables (or $n^2 + n^2 = 2n^2$ variables for $n \times n$ matrices).

Digression

As an aside, this should explain why matrix multiplication is associative but not commutative:

- Because function composition is associative, so is matrix multiplication.
- Because function composition is *not* commutative in general, matrix multiplication isn't either.

§7.4 [EXER] Exercises

Exercise 7.2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map that rotates each vector in $\mathbb{R}^2$ by 30° counterclockwise about the origin, then reflects around the line $y = x$. Write T as a 2×2 matrix.

Chapter 8. Linear combinations of vectors

§8.1 [TEXT] Definition of a basis

Recall in Section 7.1 the following two questions I posed:

Q1. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

what are the vectors for $T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 203 \\ 204 \end{pmatrix}\right)$?

Q2. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transform and it's given that

$$T\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad \text{and} \quad T\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

what is $T\left(\begin{pmatrix} 50 \\ 70 \end{pmatrix}\right)$?

Now if I wanted to make life harder for you, I could have asked:

Q3. Given $T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix}$ and $T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$, what is $T\left(\begin{pmatrix} 13 \\ 37 \end{pmatrix}\right)$?

This is a lot harder than Q1; however, the question is still solvable. If I locked you in a room for an hour and told you to work on it, I think many of you would eventually get the answer:

$$\begin{pmatrix} 13 \\ 37 \end{pmatrix} = 24\begin{pmatrix} 3 \\ 4 \end{pmatrix} - 0.59\begin{pmatrix} 100 \\ 100 \end{pmatrix} \implies T\left(\begin{pmatrix} 13 \\ 37 \end{pmatrix}\right) = 24\begin{pmatrix} \pi \\ 9 \end{pmatrix} - 0.59\begin{pmatrix} 0 \\ 12 \end{pmatrix} = \begin{pmatrix} 24\pi \\ 208.92 \end{pmatrix}.$$

So this shows you something interesting: actually, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is only special insomuch as it makes arithmetic easy. But if you know the outputs of $T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right)$ and $T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right)$, you can *still* find all the outputs of T you want. So really $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ aren't *that* special.

Hence we say $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 100 \end{pmatrix}$ are a *basis* of $\mathbb{R}^2$ (we'll give the definition in just a moment). Every $\mathbf{v}$ like $\begin{pmatrix} 13 \\ 37 \end{pmatrix}$ can be written as $c_1\begin{pmatrix} 3 \\ 4 \end{pmatrix} + c_2\begin{pmatrix} 100 \\ 100 \end{pmatrix}$ for some c_1 and c_2, and hence you can solve Q3.

However, let's consider more variations of the question Q3.

Q4. Given $T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix}$, what is $T\left(\begin{pmatrix} 13 \\ 37 \end{pmatrix}\right)$?

Q5. Given $T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix}$, $T\left(\begin{pmatrix} 100 \\ 100 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$, and $T\left(\begin{pmatrix} 103 \\ 104 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 21 \end{pmatrix}$ what is $T\left(\begin{pmatrix} 13 \\ 37 \end{pmatrix}\right)$?

Q6. Given $T\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) = \begin{pmatrix} \pi \\ 9 \end{pmatrix}$ and $T\left(\begin{pmatrix} 300 \\ 400 \end{pmatrix}\right) = \begin{pmatrix} 100\pi \\ 900 \end{pmatrix}$, what is $T\left(\begin{pmatrix} 13 \\ 37 \end{pmatrix}\right)$?

The variants Q4, Q5, Q6 are all strange in some way and should make you squint.

- Q4 is obviously impossible; *not enough information* to find an answer.
- Q5 is solvable, but has *redundant information*. You can delete any one hypothesis.
- Q6 suffers from both defects. Even though there are two givens, they are redundant.

With the examples illustrated here, let me give the relevant terms:

Definition of a basis

A **linear combination** of a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is a sum of the form $c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k$.

Then a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is:

- **linearly independent** if every linear combination gives a different vector; equivalently, there's no nontrivial linear combination giving $\mathbf{0}$ other than $c_1 = \dots = c_k = 0$;
- **spanning in** $\mathbb{R}^n$ if every other vector $\mathbf{w}$ in $\mathbb{R}^n$ can be written as some linear combination;
- a **basis of** $\mathbb{R}^n$ if both of the above are true; in other words, every vector in $\mathbb{R}^n$ can be made in *exactly* one way.

The punch line is that these concepts correspond to the behaviors in the questions above; Q3 has all the "good" properties above, and each of Q4, Q5, Q6 are missing something.

Example

- In Q3, $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 100 \end{pmatrix}$ are a basis of $\mathbb{R}^2$. Hence, the question Q3 makes a well-formed question.
- In Q4, the vector $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ by itself is not spanning (though it is linearly independent). That's what makes Q4 impossible to answer: you can't make $\begin{pmatrix} 13 \\ 37 \end{pmatrix}$ out of just $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.
- In Q5, $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 100 \\ 100 \end{pmatrix}$, $\begin{pmatrix} 103 \\ 104 \end{pmatrix}$ are not linearly independent (though they are spanning), because there's a dependence
$$\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 100 \\ 100 \end{pmatrix} = \begin{pmatrix} 103 \\ 104 \end{pmatrix}$$
between them.
- In Q6, the two vectors $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 300 \\ 400 \end{pmatrix}$ is missing *both* good properties. The two vectors give redundant information; because $\begin{pmatrix} 300 \\ 400 \end{pmatrix} = 100\begin{pmatrix} 3 \\ 4 \end{pmatrix}$, there is a dependence between the vectors. And the two vectors are not spanning: you can't make $\begin{pmatrix} 13 \\ 37 \end{pmatrix}$ out of $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 300 \\ 400 \end{pmatrix}$.

So that's what the example showing what the definition is trying to communicate.

You might have a sense already that there's no way to write a "good" question like Q3 for $\mathbb{R}^2$ that uses any number of vectors other than two, and you'd be right. It turns out a basis for $\mathbb{R}^n$ always has exactly n vectors; that's what it means that "$\mathbb{R}^n$ is n-dimensional". More examples of correct hunches that we'll explain momentarily:

- In $\mathbb{R}^n$, at least $n+1$ vectors are never linearly independent.
- In $\mathbb{R}^n$, at most $n-1$ vectors are never spanning. (So a basis is always n vectors exactly.)
- Also, if you have exactly n vectors in $\mathbb{R}^n$, and they're linearly independent, then they're a basis.
- Also, if you have exactly n vectors in $\mathbb{R}^n$, and they're spanning, then they're a basis.

§8.2 [RECIPE] How to detect a basis

This begs the question: does there exist a way to easily tell whether some vectors are a basis? For $\mathbb{R}^2$, you can probably tell by looking. But for $\mathbb{R}^n$ for $n \geq 3$, it might be trickier.

In fact, the following theorem is true, though we won't prove it.

Memorize: Basis for $\mathbb{R}^n$, "buy two get one free"

Suppose you have a bunch of vectors in $\mathbb{R}^n$. Any two of the following imply the third:

1. There are exactly n vectors.
2. The vectors are linearly independent.
3. The vectors span all of $\mathbb{R}^n$.

Moreover, if item 1 is true, the following fourth item works too:

4. The determinant of the $n \times n$ matrix with the vectors as column vectors is nonzero.

We can't prove this result in this class, but you might have the instinct that 1-3 should all be true for a basis. The determinant might be more surprising, so here's an explanation why.

Digression on why determinant does the right thing

Let $n = 3$ and consider three vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in $\mathbb{R}^3$. We give an informal explanation for why the determinant lets you tell whether these three vectors are a basis or not.

The matrix in question is formed by taking $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ as the columns of the matrix:

$$A = \begin{pmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{pmatrix}$$

Recall the determinant $\det(A)$ geometrically represents the signed volume of the parallelepiped spanned by these vectors.

Suppose $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are not spanning (i.e. not a basis), meaning they fail to occupy the full three-dimensional space. Then these vectors are coplanar: you can only get "some" vectors out of them. So geometrically, the parallelepiped they form lacks height in the direction perpendicular to the plane; resulting in zero volume, that is, $\det(A) = 0$.

Conversely, if $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ were a basis, then pictorially this means they span the entire space, so the parallelepiped had better be nondegenerate. The nondegeneracy corresponds to having nonzero volume, that is, $\det(A) \neq 0$. (The sign of the determinant tells you something about the orientation, but we don't care about this sign, just the nonzero-ness.)

Anyway, in practice, if you have an explicit set of vectors, the recipe is simple now:

Recipe for checking if vectors form a basis

Suppose you're given a list of vectors in $\mathbb{R}^n$ and want to know if they're a basis.

1. If the *number* of vectors is not n, output "no" and stop.
2. Otherwise, form the matrix with the n vectors as columns, and compute its determinant. Output "yes" if and only if the determinant is nonzero.

The determinant thing matters: the determinant is doing a *lot* of work for you. When $n = 2$ the determinant is unnecessary, because you can just use "slope": it's obvious that $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 200 \end{pmatrix}$ have a dependence. But for $n \geq 3$ **you can't eyeball it**[13]. For example, the three vectors

[13]Though if you have a set of *exactly two* vectors, they're dependent if and only if they're multiples, even in $\mathbb{R}^n$. Which you *can* eyeball; so if you're trying to tell whether a span of two vectors in $\mathbb{R}^3$ is a line or plane, that's easy. (Even more stupidly, a single vector is linearly dependent only when it's the zero vector.)

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix}$$

might look like unrelated small numbers, but surprisingly it turns out that

$$109 \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} - 37 \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix} - 29 \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix} = \mathbf{0}. \tag{5}$$

Without "slope", you cannot notice these dependences by sight for $n \geq 3$, so use the determinant.

Sample Question

Is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 100 \\ 200 \end{pmatrix}$ a basis for $\mathbb{R}^2$?

Solution. No because $\begin{vmatrix} 1 & 2 \\ 100 & 200 \end{vmatrix} = 0$. In this case you can also see directly that $100\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 100 \\ 200 \end{pmatrix}$, so the vectors are not linearly independent. $\square$

Sample Question

Is $\begin{pmatrix} 13 \\ 37 \end{pmatrix}, \begin{pmatrix} 42 \\ 88 \end{pmatrix}$ a basis for $\mathbb{R}^2$?

Solution. Yes, because $\begin{vmatrix} 13 & 37 \\ 42 & 88 \end{vmatrix} = 13 \cdot 88 - 37 \cdot 42 \neq 0$. $\square$

Sample Question

Is $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix}$ a basis for $\mathbb{R}^3$?

Solution. No, because

$$\begin{aligned} \begin{vmatrix} 1 & 3 & 4 \\ 10 & 1 & 11 \\ -9 & 10 & 1 \end{vmatrix} &= 1 \begin{vmatrix} 1 & 11 \\ 10 & 1 \end{vmatrix} - 3 \begin{vmatrix} 10 & 11 \\ -9 & 1 \end{vmatrix} + 4 \begin{vmatrix} 10 & 1 \\ -9 & 10 \end{vmatrix} \\ &= 1(10 - 11) - 3(10 + 99) + 4(100 + 9) = 0. \end{aligned}$$

$\square$

Sample Question

Is $\begin{pmatrix} 3 \\ 42 \\ 18 \end{pmatrix}, \begin{pmatrix} 1 \\ 53 \\ 17 \end{pmatrix}, \begin{pmatrix} 71 \\ 91 \\ 13 \end{pmatrix}$ a basis for $\mathbb{R}^3$?

Solution. Yes, because

$$\begin{vmatrix} 3 & 1 & 71 \\ 42 & 53 & 91 \\ 18 & 17 & 13 \end{vmatrix} = \text{ugly arithmetic} = -18522 \neq 0.$$

$\square$

Remark: If you randomly generate numbers, you'll get a basis

The above sample was generated randomly when I gave this lecture at MIT. The way I presented this was I went up to the board and wrote:

"Is $\begin{pmatrix} ? \\ ? \\ ? \end{pmatrix}, \begin{pmatrix} ? \\ ? \\ ? \end{pmatrix}, \begin{pmatrix} ? \\ ? \\ ? \end{pmatrix}$ *a basis for* $\mathbb{R}^3$*? Answer: no."*

Then I asked my students to make up nine numbers to fill in the question marks. Of course, they picked big numbers, and I got to show off my amazing five-digit multiplication skills.

I did this stunt of writing the answer before I even knew the question to make a point: if you pick large truly random numbers, the determinant will be some large random number too, so there's no chance you'll just get 0. So you should always expect n "random" vectors in $\mathbb{R}^n$ to be a basis, and it's only if you cherry-pick them you'll get a non-basis.

Sample Question

Is $\begin{pmatrix} 1 \\ 4 \\ 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \\ 11 \\ 6 \end{pmatrix}, \begin{pmatrix} 6 \\ 19 \\ 10 \\ 2 \end{pmatrix}$ a basis for $\mathbb{R}^4$?

Solution. No, there are three vectors, and $\mathbb{R}^4$ needs to have exactly four vectors in every basis. □

§8.3 [TEXT] Spans

The basis is the "best-case scenario", because if you have a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of $\mathbb{R}^n$ then it means every vector of $\mathbb{R}^n$ can be made out of $\mathbf{v}_i$ in exactly one way. We won't always be so lucky, so we have a word that means "what you can make out of $\mathbf{v}_i$".

Definition of span

The **span** of a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ refers to the vectors that you can make out of $\mathbf{v}_i$ (i.e. can be written as a linear combination of $\mathbf{v}_i$).

Example of spans in $\mathbb{R}^2$

Consider vectors in $\mathbb{R}^2$.

- The span of any basis, like $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 7 \\ 11 \end{pmatrix}$, is all of $\mathbb{R}^2$, by definition.
- Moreover, the span of any set *containing* a basis is also all of $\mathbb{R}^2$. For example, the span of $\begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 8 \\ 10 \end{pmatrix}, \begin{pmatrix} 7 \\ 11 \end{pmatrix}, \begin{pmatrix} 700 \\ 1100 \end{pmatrix}$ is still all of $\mathbb{R}^2$. (Sure, it's not a basis as there are lots of dependencies, but that doesn't change that you can make any vector in $\mathbb{R}^2$ out of them; you just have some extra vectors you don't have to use.)
- The span of the vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 10 \\ 20 \end{pmatrix}, \begin{pmatrix} 100 \\ 200 \end{pmatrix}$ is the line $y = 2x$. These are the only vectors you can make out of combinations of these three vectors.
- The span of the single vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is also the line $y = 2x$. That is, in the previous example, $\begin{pmatrix} 10 \\ 20 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 200 \end{pmatrix}$ were totally useless
- The span of the single vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is only one point: $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ itself.

Examples of spans in $\mathbb{R}^3$

Consider vectors in $\mathbb{R}^3$.

- The span of any basis, like $\begin{pmatrix} 3 \\ 421 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 53 \\ 17 \end{pmatrix}, \begin{pmatrix} 71 \\ 91 \\ 13 \end{pmatrix}$ is all of $\mathbb{R}^3$, by definition.
- The span of the vectors $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}, \begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix}$ is the line consisting of multiples of $y = 2x$ and $z = 3x$, though I think it's easier to express this as "multiples of $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$". These are the only vectors you can make out of combinations of these three vectors.
- The span of the single vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is the same line. That is, in the previous example, $\begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix}$ were totally useless.
- The span of $\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}, \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix}$ is more interesting. We saw before that these three vectors actually have a dependence, so they are not a basis and the span is not all of $\mathbb{R}^3$.

  But the span is not just a line either: it's a two-dimensional plane! In fact, all three vectors are contained inside the plane
  $$x + y = z.$$
  And the span is actually that entire plane; any point in the plane turns out to be formed as a combination, in fact just using the first two vectors is enough. For example, if I picked a random point on the plane like $\begin{pmatrix} 1337 \\ 2025 \\ 3362 \end{pmatrix}$ then it turns out the relevant combination is
  $$\begin{pmatrix} 1337 \\ 2025 \\ 3362 \end{pmatrix} = \frac{19113}{29} \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} + \frac{1386}{29} \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}$$
  (I won't explain how I got these coefficients, but you could probably figure out yourself if you wanted to).
- The span of the single vector $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is only one point: $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ itself.

What you are probably gathering from all these examples is that the span *also* has a concept of dimension: for example in $\mathbb{R}^3$, we saw an example of

- a 0-dimensional span, just the single point $\langle 0, 0, 0 \rangle$.
- a 1-dimensional span, the line consisting of multiples of $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$
- a 2-dimensional span, the plane $x + y = z$.
- a 3-dimensional span if you have a full basis.

This is worth knowing:

Tip

Remember, in $\mathbb{R}^n$, the span of a bunch of vectors is

- always d-dimensional for some $d = 0, \dots, n$ (so there are $n+1$ kinds of answers);
- always contains the origin $\mathbf{0}$.

§8.4 [RECIPE] Describing the span of several vectors

In 18.02 you might be asked to describe the span of some vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$. From the examples above, you should be able to extrapolate the recipe.

Recipe for describing the span of vectors in $\mathbb{R}^2$

- **0D case**: Are all the vectors the zero vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$? If so the span is just a single **point**.
- **1D case**: Are all the vectors pointing the same direction (i.e. multiples of each other)? If so, and there is at least one nonzero vector, the span is a **line** through the origin in the common direction of the vectors.
- **2D case**: Are there two (nonzero) vectors not pointing in the same direction (equivalently, are linearly independent)? If so, the span is **all of $\mathbb{R}^2$**.

Sample Question

What is the span of the vectors $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$, $\begin{pmatrix} 10 \\ 20 \end{pmatrix}$, $\begin{pmatrix} 100 \\ 200 \end{pmatrix}$, $\begin{pmatrix} 5000 \\ 10000 \end{pmatrix}$ in $\mathbb{R}^2$?

Solution. All the vectors are multiples of $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, so the answer is a line: the multiples of $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. □

Sample Question

What is the span of $\begin{pmatrix} 420 \\ 321 \end{pmatrix}$ and $\begin{pmatrix} 666 \\ 5 \end{pmatrix}$ in $\mathbb{R}^2$?

Solution. Because the two vectors are not multiples of each other, they are linearly independent. (The determinant lets you see this too: $\begin{vmatrix} 420 & 666 \\ 321 & 5 \end{vmatrix} = 420 \cdot 5 - 321 \cdot 666 = -211686 \neq 0$.) Hence they are a basis of $\mathbb{R}^2$ and the span is all of $\mathbb{R}^2$. □

Recipe for describing the span of vectors in $\mathbb{R}^3$

- **0D case:** Are all the vectors the zero vector $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$? If so the span is just a single **point**.
- **1D case**: Are all the vectors pointing the same direction (i.e. multiples of each other)? If so, and there is at least one nonzero vector, the span is a **line** through the origin in the common direction of the vectors.
- **2D case**: Is there more than one direction present, but you can't find three vectors which are linearly independent? If so, the span is a **plane** through the origin.
  - If you want the equation of the plane, use the cross product.
- **3D case**: Are there three vectors among them which are linearly independent from each other? If so, the span is **all of $\mathbb{R}^3$**.

(In 18.02 the reason for the 2D case is really by process of elimination: you can think of the 2D bullet as what's left over if you rule out 0D, 1D, 3D.)

Sample Question

What is the span of the vectors $\begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}$, $\begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$ and $\begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix}$ in $\mathbb{R}^3$?

Solution. All the vectors are multiples of $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, so it's a line: the multiples of $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. □

Sample Question

What is the span of $\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}$ and $\begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$?

Solution. It should be a two-dimensional plane. This follows by process of elimination: we know $d = 0$ and $d = 1$ don't apply here (none of these vectors are zero or are multiples of each other) and we can rule out $d = 3$ because we can calculate the determinant

$$\begin{vmatrix} 1 & 10 & -9 \\ 3 & 1 & 10 \\ 4 & 11 & 1 \end{vmatrix} = 0$$

to see that our three vectors are *not* linearly independent.

How do we actually find the equation of the plane? Well, really what we're asking is to find a plane through the origin passing through all of $(1, 3, 4)$, $(10, 1, 11)$, $(-9, 10, 1)$ (we're promised it exists from the determinant being zero: again, that's what it means to not be spanning). We just want a normal vector to the plane, which we can get by taking the cross product of any two of the vectors. Arbitrarily, we use the first two:

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 3 & 4 \\ 10 & 1 & 11 \end{vmatrix} = 29\mathbf{e}_1 + 29\mathbf{e}_2 - 29\mathbf{e}_3.$$

That's the normal vector we wanted, so we now know $29x + 29y - 29z = 0$ is the plane needed. This simplifies to just $\boxed{x + y - z = 0}$. □

Digression

Just for comparison, if you had used the second and third vector instead, you'd get

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 3 & 4 \\ -9 & 10 & 1 \end{vmatrix} = 37\mathbf{e}_1 + 37\mathbf{e}_2 - 37\mathbf{e}_3$$

while the second and third vectors would give

$$\begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix} \times \begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 10 & 1 & 11 \\ -9 & 10 & 1 \end{vmatrix} = -109\mathbf{e}_1 - 109\mathbf{e}_2 + 109\mathbf{e}_3$$

which are all still multiples of $\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$, so the plane is still $x + y - z = 0$.

It's not a coincidence that the magic numbers 29, 37, 109 from Equation 5 are reappearing in the normal vectors, but this time I won't explain why this is happening, and let you ruminate on it yourself if you want to figure out.

Sample Question

What is the span of $\begin{pmatrix} 3 \\ 42 \\ 18 \end{pmatrix}, \begin{pmatrix} 1 \\ 53 \\ 17 \end{pmatrix}, \begin{pmatrix} 71 \\ 91 \\ 13 \end{pmatrix}$ in $\mathbb{R}^3$?

Solution. As we mentioned above (Equation 5), you shouldn't eyeball three or more dimensions; if you get three vectors in $\mathbb{R}^3$ and want to know if they are linearly independent or not, you should always take the determinant:

$$\begin{vmatrix} 3 & 1 & 71 \\ 42 & 53 & 91 \\ 18 & 17 & 13 \end{vmatrix} = \text{ugly arithmetic} = -18522 \neq 0.$$

So the three vectors are a basis and the span is all of $\mathbb{R}^3$. □

§8.5 [TEXT] Systems of equations

As we commented earlier, a randomly chosen set of n vectors is "usually" a basis for $\mathbb{R}^n$. I want to now connect this to something else you've seen in high school: a "random" linear system of n equations and n variables "usually" has only one solution.

For example, in high school algebra, you probably were asked to solve systems of equations like

$$\begin{aligned} x_1 + 2x_2 &= 14 \\ 3x_1 + 4x_2 &= 38. \end{aligned}$$

You can do this using whatever method you're used to; you should find $(x_1, x_2) = (10, 2)$ as the only solution. And you can probably already tell that 14 and 38 could have been any numbers, and you'd still always get exactly one solution.

Why is this relevant to the stuff about basis? Well, the point is you can view the variables above as coefficients in a linear combination and consider the previous system of equations as saying

$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 14 \\ 38 \end{pmatrix}.$$

That means our observations can be rephrased in terms of our linear algebra language:

> **Idea**
>
> $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ are a basis of $\mathbb{R}^2$, so any vector like $\begin{pmatrix} 14 \\ 38 \end{pmatrix}$ can be made in exactly one way.

In other words, as long as the column vectors made bx_2 the left-hand side on a basis, there's always in fact one solution.

So what goes wrong when it's *not* a basis? Let's bring back the example we had in Equation 5.

- Consider a system like

$$\begin{aligned} x_1 + 10x_2 - 9x_3 &= 0 \\ 3x_1 + x_2 + 10x_3 &= 0 \\ 4x_1 + 11x_2 + x_3 &= 0. \end{aligned}$$

  This equation obviously has at least one solution: $x_1 = x_2 = x_3 = 0$. We saw that $\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 10 \\ 1 \\ 11 \end{pmatrix}$, $\begin{pmatrix} -9 \\ 10 \\ 1 \end{pmatrix}$ is not a basis of $\mathbb{R}^3$, and what that translates is to saying there are other solutions too: reading off Equation 5, note that $x_1 = 109k$, $x_2 = -37k$, $x_3 = -29k$ is a solution for any k.

- What if we consider something like

$$\begin{aligned} x_1 + 10x_2 - 9x_3 &= 17 \\ 3x_1 + x_2 + 10x_3 &= 42 \\ 4x_1 + 11x_2 + x_3 &= 1337 \end{aligned}$$

  instead? We still know the coefficients on the left-hand side have a dependency, so we expect something to "go wrong".

  In this case, the kind of failure is different: $\begin{pmatrix} 17 \\ 42 \\ 1337 \end{pmatrix}$ turns out to *not* be in the span of our three vectors. In fact, we earlier saw that the span of the vectors was the plane $x + y = z$, which doesn't contain $\begin{pmatrix} 17 \\ 42 \\ 1337 \end{pmatrix}$.

  So for this system, satisfying these three equations simultaneously should be impossible (that is, there are *no solutions* rather than *too many solutions*). And in fact you might be able to see this directly: if you add the first two equations and subtract the last one you can see the contradiction:

$$\begin{aligned} +[x_1 + 10x_2 - 9x_3] &= +17 \\ +[3x_1 + x_2 + 10x_3] &= +42 \\ -[4x_1 + 11x_2 + x_3] &= -1337 \\ \implies 0 &= 17 + 42 - 1337 \quad \text{which is absurd.} \end{aligned}$$

- On the other hand, if we have

$$\begin{aligned} x_1 + 10x_2 - 9x_3 &= 100 \\ 3x_1 + x_2 + 10x_3 &= 200 \\ 4x_1 + 11x_2 + x_3 &= 300 \end{aligned}$$

then the span does contain $\begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix}$, so we expect to be back in the "too many solutions" case. I won't show you how to come up with these numbers (though it's actually not that hard), but you can find one example of a solution is $(x_1, x_2, x_3) = (58, 6, 2)$. If you put that together with Equation 5 you can come up with a lot of solutions now:

$$\begin{aligned} x_1 &= 58 + 109k \\ x_2 &= 6 - 37k \\ x_3 &= 2 - 29k. \end{aligned}$$

In summary, what we've seen in this section is that for a system of n equations in n variables, we should look at the n column vectors made by the coefficients. Then

- If those vectors form a basis of $\mathbb{R}^n$ (usual case), the system of equations always has one solution.
- Otherwise, the system of equations is defective: either it has no solutions at all (is self-contradictory), or there are actually infinitely many solutions.

§8.6 [RECIPE] Number of solutions to a square system of linear equations

Now, remember that *in practice* the way we see whether or not n vectors form a basis is by considering a determinant. So we can rephrase our findings from the previous section into the following recipe using the word "determinant" in place of "basis".

Recipe for computing the number of solutions to a system of equations

Suppose you are asked to find the number of solutions to a square system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_1nx_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_2nx_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n. \end{aligned}$$

Let A be the $n \times n$ matrix formed by the a_{ij}.

1. If $\det A \neq 0$, you don't even have to look at b_i; just **output "exactly 1 solution"**.
2. If $\det A = 0$, you should **output either "zero solutions" or "infinitely many solutions"**, depending on whether there is at least one solution. To see which case you're in:
   - There's a common case $b_1 = b_2 = \dots = b_n = 0$, where the system has an obvious solution $x_1 = \dots = x_n = 0$. Thus output "infinitely many solutions".
   - When $n = 2$, you can usually tell by looking whether the two equations are redundant or not. Output "infinitely many solutions" if the two equations are multiples of each other; output "zero solutions" if the two equations contradict each other.
   - Otherwise, for $n \geq 3$, if you can't guess a solution, you should eliminate variables one by one. However, this case doesn't occur in 18.02.

Examples of what I mean when I say "tell by looking" for $n = 2$:

- The system $x + 3y = 8$ and $10x + 30y = 80$ is obviously "infinitely many", because the two equations are the same. In a case like this, **output "infinitely many"**.
- The system $x + 3y = 8$ and $10x + 30y = 42$ is obviously "no solutions", because it would imply $80 = 10 \cdot 8 = 10 \cdot (x + 3y) = 10x + 30y = 42$, which is self-contradictory. In a case like this, **output 0**.

§8.7 [EXER] Exercises

Exercise 8.1. Take your birthday and write it in eight-digit $Y_1 Y_2 Y_3 Y_4$-$M_1 M_2$-$D_1 D_2$ format. Consider the two vectors

$$\mathbf{v}_1 = \begin{pmatrix} M_1 M_2 \\ D_1 D_2 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} Y_1 Y_2 \\ Y_3 Y_4 \end{pmatrix}.$$

For example, if your birthday was May 17, 1994 you would take $\mathbf{v}_1 = \left(\begin{smallmatrix} 5 \\ 17 \end{smallmatrix}\right)$ and $\mathbf{v}_2 = \left(\begin{smallmatrix} 19 \\ 94 \end{smallmatrix}\right)$.

- Compute the span of those two vectors in $\mathbb{R}^2$.
- Find a current or former K-pop idol who gets a different answer from you when they use their birthday.

Exercise 8.2. In $\mathbb{R}^5$, consider the vector $\mathbf{v} = \langle 1, 2, 3, 4, 5 \rangle$. Compute the maximum possible number of linearly independent vectors one can find which are all perpendicular to $\mathbf{v}$.

Chapter 9. Eigenvalues and eigenvectors

In this chapter, we'll define an eigenvalue and eigenvector. The main goal of this chapter is that:

Goal

Given a 2×2 or 3×3 matrix, by the end of this chapter, you should be able to find all the eigenvalues and eigenvectors by hand.

§9.1 [TEXT] The problem of finding eigenvectors

Let's define the relevant term first:

Definition

Suppose T is a matrix or linear transformation, λ a scalar, and $\mathbf{v}$ is a vector such that

$$T(\mathbf{v}) = \lambda \mathbf{v};$$

that is, T sends $\mathbf{v}$ to a multiple of itself. Then we call λ an **eigenvalue** and $\mathbf{v}$ an **eigenvector**.

Type signature

Eigenvalues λ are always scalars.

Example

Let $T = \begin{pmatrix} 74 & 52 \\ 32 & 36 \end{pmatrix}$ and consider the vector $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Then

$$T(\mathbf{v}) = \begin{pmatrix} 74 & 52 \\ 32 & 36 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 200 \\ 100 \end{pmatrix} = 100 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 100\mathbf{v}.$$

So we would say $\mathbf{v}$ is an eigenvector with eigenvalue 100.

Of course, if $\mathbf{v}$ is an eigenvector, so are all its multiples, e.g.

$$\begin{pmatrix} 74 & 52 \\ 32 & 36 \end{pmatrix} \begin{pmatrix} 20 \\ 10 \end{pmatrix} = \begin{pmatrix} 2000 \\ 1000 \end{pmatrix} = 100 \begin{pmatrix} 20 \\ 10 \end{pmatrix}$$

so $\begin{pmatrix} 20 \\ 10 \end{pmatrix}$ is an eigenvector with the same eigenvalue 100, etc.

Remark

The stupid solution $\mathbf{v} = \mathbf{0}$ always satisfies the eigenvector equation for any λ, so we will pretty much ignore it and focus only on finding nonzero eigenvectors.

The goal of this chapter is to show, given a matrix T, how we can find its eigenvectors (besides $\mathbf{0}$).

§9.2 [TEXT] How to come up with the recipe for eigenvalues

For this story, our protagonist will be the matrix

$$A = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}.$$

Phrased another way, the problem of finding eigenvectors is, by definition, looking for λ, x, y such that

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix} \iff \begin{cases} 5x - 2y = \lambda x \\ 3x + 10y = \lambda y \end{cases}.$$

Smart-alecks will say $x = y = 0$ always works for every λ. *Are there other solutions?*

§9.2.1 Why guessing the eigenvalues is ill-fated

As an example, let's see if there are any eigenvectors $\binom{x}{y}$ with eigenvalue 100. In other words, let's solve

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 100\begin{pmatrix} x \\ y \end{pmatrix}.$$

If we solve the system of equations, we get

$$\begin{cases} 5x - 2y = 100x \\ 3x + 10y = 100y \end{cases} \implies \begin{cases} -95x - 2y = 0 \\ 3x - 90y = 0 \end{cases} \implies x = y = 0.$$

Well, that's boring. In this system of equations, the only solution is $x = y = 0$.

We can try a different guess: maybe we use 1000 instead of 100. An eigenvector with eigenvalue 1000 ought to correspond to

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1000\begin{pmatrix} x \\ y \end{pmatrix}.$$

If we solve the system of equations, we get

$$\begin{cases} 5x - 2y = 1000x \\ 3x + 10y = 1000y \end{cases} \implies \begin{cases} -995x - 2y = 0 \\ 3x - 990y = 0 \end{cases} \implies x = y = 0$$

which… isn't any better. We still don't get any solutions besides $x = y = 0$.

At this point, you should be remembering something I told you last chapter: a "random" system of equations and variables usually only has a unique solution. So if I keep picking numbers out of a hat like 100, 1000, etc., then I'm unlikely to find anything interesting. In order to get a system that doesn't just solve to $x = y = 0$, I'm going to need to cherry-pick my number λ.

§9.2.2 Cherry-picking λ

Let's try to figure out what value of λ would make the system more interesting. If we copy what we did above, we see that the general process is:

$$\begin{cases} 5x - 2y = \lambda x \\ 3x + 10y = \lambda y \end{cases} \implies \begin{cases} (5-\lambda)x - 2y = 0 \\ 3x + (10-\lambda)y = 0 \end{cases}$$

We need to cherry-pick λ to make sure that the system doesn't just solve to $x = y = 0$ like the examples we tried with 100 and 1000. But we learned how to do this in the last chapter: in order to get a degenerate system you need to make sure that

$$0 = \begin{vmatrix} 5-\lambda & -2 \\ 3 & 10-\lambda \end{vmatrix}.$$

Remark

At this point, you might notice that this is secretly an explanation of why $A - \lambda I$ keeps showing up on your formula sheet. Writing $A\mathbf{v} = \lambda\mathbf{v}$ is the same as $(A - \lambda I)\mathbf{v} = 0$, just more opaquely.

Expanding the determinant on the left-hand side gives

$$0 = \begin{vmatrix} 5-\lambda & -2 \\ 3 & 10-\lambda \end{vmatrix} = (5-\lambda)(10-\lambda) + 6 = \lambda^2 - 15\lambda + 56 = (\lambda - 7)(\lambda - 8).$$

Great! So we expect that if we choose either $\lambda = 7$ and $\lambda = 8$, then we will get a degenerate system, and we won't just get $x = y = 0$. Indeed, let's check this:

- When $\lambda = 7$, our system is

$$\begin{cases} 5x - 2y = 7x \\ 3x + 10y = 7y \end{cases} \implies \begin{cases} -2x - 2y = 0 \\ 3x + 3y = 0 \end{cases} \implies x = -y.$$

  So for example, $\left(\begin{smallmatrix} -13 \\ 13 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 37 \\ -37 \end{smallmatrix}\right)$ will be eigenvectors with eigenvalue 7:

$$A\begin{pmatrix} -13 \\ 13 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} -13 \\ 13 \end{pmatrix} = \begin{pmatrix} -91 \\ 91 \end{pmatrix} = 7\begin{pmatrix} -13 \\ 13 \end{pmatrix}.$$

  On exam, you probably answer "the eigenvectors with eigenvalue 7 are the multiples of $\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$", or "the eigenvectors with eigenvalue 7 are the multiples of $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$" if you want; these are the same thing. Or if you want to mess with the grader, "the eigenvectors with eigenvalue 7 are the multiples of $\left(\begin{smallmatrix} 100 \\ -100 \end{smallmatrix}\right)$" is fine too.

- When $\lambda = 8$, our system is

$$\begin{cases} 5x - 2y = 8x \\ 3x + 10y = 8y \end{cases} \implies \begin{cases} -3x - 2y = 0 \\ 3x + 2y = 0 \end{cases} \implies x = -\frac{2}{3}y.$$

  So for example, $\left(\begin{smallmatrix} -20 \\ 30 \end{smallmatrix}\right)$ is an eigenvector with eigenvalue 8:

$$A\begin{pmatrix} -20 \\ 30 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} -20 \\ 30 \end{pmatrix} = \begin{pmatrix} -160 \\ 240 \end{pmatrix} = 8\begin{pmatrix} -20 \\ 30 \end{pmatrix}.$$

  On exam, you should answer "the eigenvectors with eigenvalue 8 are the multiples of $\left(\begin{smallmatrix} -2 \\ 3 \end{smallmatrix}\right)$". Or you can say "the eigenvectors with eigenvalue 8 are the multiples of $\left(\begin{smallmatrix} 2 \\ -3 \end{smallmatrix}\right)$" if you want; these are the same thing. You could even say "the eigenvectors with eigenvalue 8 are the multiples of $\left(\begin{smallmatrix} 200 \\ -300 \end{smallmatrix}\right)$" and still get credit, but that's silly.

§9.3 [RECAP] Summary

To summarize the story above:

- We had the matrix $A = \left(\begin{smallmatrix} 5 & -2 \\ 3 & 10 \end{smallmatrix}\right)$ and wanted to find λ's for which the equation

$$\begin{pmatrix} 5 & -2 \\ 3 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda\begin{pmatrix} x \\ y \end{pmatrix}$$

  had solutions other than $x = y = 0$.

- We realized that guessing λ was never going to fly, so we went out of our way to cherry-pick λ to make sure the system was degenerate. The buzzwords for this are "find the roots of the characteristic polynomial", but I wanted to show that it flows naturally from the end goal.
- For the two values of λ we cherry-picked, we know the system of equations is degenerate. So we solve the two degenerate systems and see what happens.

In lectures and notes, the last two bullets are separated as two different steps, to make it into a recipe. But don't lose sight of how they're connected! I would rather call it the following interlocked thing:

- We cherry-pick λ to make sure the system doesn't just solve to $x = y = 0$.
- To do the cherry-picking, ensure $\det(A - \lambda I) = 0$ so that our system is degenerate.

§9.4 [RECIPE] Calculating all the eigenvalues

To repeat the story:

Recipe for finding the eigenvectors and eigenvalues

Given a matrix A, to find its eigenvectors and eigenvalues:

1. Find all the values of λ such that, if you subtract λ from every diagonal entry of A (that is, look at $A - \lambda I$), the resulting square matrix of coefficients has determinant 0.
2. For each λ, solve the degenerate system and output the solutions to it. (You should find there is at least a one-dimensional space of solutions.)

Remark

Eigenvectors are sometimes grouped into so-called *eigenlines* because every multiple of an eigenvector is also an eigenvector. For example, if you get $2x + y = 0$ for $\lambda = 2$, any of the following outputs is often acceptable:

- "Any multiple of $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ is an eigenvector for $\lambda = 2$"
- "Any multiple of $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ is an eigenvector for $\lambda = 2$"
- "Any multiple of $\begin{pmatrix} 100 \\ -200 \end{pmatrix}$ is an eigenvector for $\lambda = 2$"
- …

And in practice people will just say "$\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ is *the* eigenvector for $\lambda = 2$" and the "any multiple of" is understood.

Sample Question

Compute all eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$

Solution. We follow the recipe:

1. We compute $A - \lambda I$, where I is the identity matrix:

$$A - \lambda I = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 - \lambda & 1 \\ 2 & 3 - \lambda \end{pmatrix}.$$

Now, compute the determinant of $A - \lambda I$ and set it equal to zero:

$$\det(A - \lambda I) = (4 - \lambda)(3 - \lambda) - (2 \cdot 1) = (4 - \lambda)(3 - \lambda) - 2.$$

Expanding this:

$$(4 - \lambda)(3 - \lambda) = 12 - 7\lambda + \lambda^2,$$

so the equation becomes:

$$12 - 7\lambda + \lambda^2 - 2 = 0 \implies 0 = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5).$$

Solving for λ gives $\lambda = 2$ or $\lambda = 5$.

2. There are two cases:
   - For $\lambda = 5$, solve $(A - 5I)\mathbf{v} = 0$:
   $$A - 5I = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \implies \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
   This gives the equation $-x + y = 0$, so the eigenvector is $\mathbf{v}_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ as well as any multiple of it.
   - For $\lambda = 2$, solve $(A - 2I)\mathbf{v} = 0$:
   $$A - 2I = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \implies \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
   This gives the equation $2x + y = 0$, so the eigenvector is $\mathbf{v}_2 = \left(\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right)$ or any multiple of it.

In conclusion, $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ and its multiples are the eigenvectors for $\lambda = 5$ and $\left(\begin{smallmatrix} 1 \\ -2 \end{smallmatrix}\right)$ and its multiples are the eigenvectors for $\lambda = 2$. □

Sample Question

Compute the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 3 & 3 \\ 4 & -1 \end{pmatrix}.$$

Solution. We follow the recipe:

1. First, we compute $A - \lambda I$:
$$A - \lambda I = \begin{pmatrix} 3 & 3 \\ 4 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 - \lambda & 3 \\ 4 & -1 - \lambda \end{pmatrix}.$$
To find the eigenvalues, we set the determinant of $A - \lambda I$ equal to zero:
$$\det(A - \lambda I) = \det \begin{pmatrix} 3 - \lambda & 3 \\ 4 & -1 - \lambda \end{pmatrix} = (3 - \lambda)(-1 - \lambda) - (3 \cdot 4) = (3 - \lambda)(-1 - \lambda) - 12.$$
Expanding this expression:
$$(3 - \lambda)(-1 - \lambda) = -3 - 3\lambda + \lambda + \lambda^2 = \lambda^2 - 2\lambda - 3.$$
Now, substitute this into the equation:

$$\lambda^2 - 2\lambda - 3 - 12 = 0 \implies 0 = \lambda^2 - 2\lambda - 15 = (\lambda - 5)(\lambda + 3).$$

Solving for λ gives $\lambda = 5$ or $\lambda = -3$.

2. Now, we find the eigenvectors corresponding to each eigenvalue.
   - For $\lambda = 5$, we solve $(A - 5I)\mathbf{v} = 0$. First, compute $A - 5I$:

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

     We now solve the system:

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

     This gives the equations:

     $$-2x + 3y = 0, \quad 4x - 6y = 0.$$

     From the first equation, we get $x = \frac{3}{2}y$. Therefore, the eigenvectors corresponding to $\lambda = 5$ are the multiples of $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.
   - For $\lambda = -3$, we solve $(A + 3I)\mathbf{v} = 0$. First, compute $A + 3I$:

     $$A + 3I = \begin{pmatrix} 3+3 & 3 \\ 4 & -1+3 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 4 & 2 \end{pmatrix}.$$

     We now solve the system:

     $$\begin{pmatrix} 6 & 3 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

     This gives the equations:

     $$6x + 3y = 0, \quad 4x + 2y = 0.$$

     From the first equation, we get $x = -\frac{1}{2}y$. Therefore, the eigenvector corresponding to $\lambda = -3$ is $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

In conclusion, $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ and its multiples are the eigenvectors for $\lambda = 5$ and $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and its multiples are the eigenvectors for $\lambda = -3$. □

Sample Question

Compute the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 2 & 9 \\ -1 & 8 \end{pmatrix}.$$

Solution. We follow the recipe, but this time we'll find the quadratic polynomial in λ we get has a repeated root, a new phenomenon. Nothing changes much though, recipe still works fine.

1. We need to find the matrix $A - \lambda I$, where I is the identity matrix. First, we compute $A - \lambda I$:

   $$A - \lambda I = \begin{pmatrix} 2 & 9 \\ -1 & 8 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2-\lambda & 9 \\ -1 & 8-\lambda \end{pmatrix}.$$

To find the eigenvalues, we set the determinant of $A - \lambda I$ equal to zero:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 9 \\ -1 & 8-\lambda \end{vmatrix} \\ &= (2-\lambda)(8-\lambda) - 9 \cdot (-1) = (2-\lambda)(8-\lambda) + 9 \\ &= \lambda^2 - 10\lambda + 25 = (\lambda - 5)^2. \end{aligned}$$

This gives $\lambda = 5$. So we only have one case!

2. First, compute $A - 5I$:

$$A - 5I = \begin{pmatrix} 2-5 & 9 \\ -1 & 8-5 \end{pmatrix} = \begin{pmatrix} -3 & 9 \\ -1 & 3 \end{pmatrix}.$$

We now solve the system:

$$\begin{pmatrix} -3 & 9 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This gives the equations:

$$-3x + 9y = 0, \quad -x + 3y = 0.$$

From the first equation, we get $x = 3y$. Therefore, the eigenvector corresponding to $\lambda = 5$ is $\left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)$ and its multiples.

The only eigenvalue of the matrix $A = \left(\begin{smallmatrix} 2 & 9 \\ -1 & 8 \end{smallmatrix}\right)$ is $\lambda = 5$ (with multiplicity 2). The corresponding eigenvector is $\left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)$ and its multiples. □

Sample Question

Compute the eigenvalues and eigenvectors of the matrix $A = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 4 \end{smallmatrix}\right)$.

Solution. The matrix is 3×3, but that's no big deal.

1. We need to find the matrix $A - \lambda I$, where I is the identity matrix. First, we compute $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 1 & 4-\lambda \end{pmatrix}.$$

To find the eigenvalues, we set the determinant of $A - \lambda I$ equal to zero:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 1 & 4-\lambda \end{vmatrix} \\ &= (1-\lambda) \det \begin{pmatrix} 3-\lambda & 0 \\ 1 & 4-\lambda \end{pmatrix} = (1-\lambda)(3-\lambda)(4-\lambda). \end{aligned}$$

Setting this equal to 0 and solving gives $\lambda = 1$, $\lambda = 3$, $\lambda = 4$.

2. Now, we find the eigenvectors corresponding to each eigenvalue.

   - For $\lambda = 1$: We solve $(A - I)\mathbf{v} = 0$. First, compute $A - I$:

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

We now solve the system:

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

In other words $2y = 0$, $2y = 0$ and $y + 3z = 0$. From the first and second rows, we have $2y = 0$, so $y = 0$. From the third row, we have $z = 0$. There are no constraints on x at all. Thus, the eigenvector corresponding to $\lambda = 1$ is

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and all its multiples, i.e. those vectors for which the second and third component are zero.

- For $\lambda = 3$: We solve $(A - 3I)\mathbf{v} = 0$. First, compute $A - 3I$:

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

We now solve the system:

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

In other words, $-2x + 2z = 0$, $0 = 0$ and $y + z = 0$. From the third row, we have $y = -z$. From the first row, we get $-2x + 2z = 0$, so $x = z$. Thus, the eigenvector corresponding to $\lambda = 3$ is:

$$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

and its multiples.

- For $\lambda = 4$: We solve $(A - 4I)\mathbf{v} = 0$. First, compute $A - 4I$:

$$A - 4I = \begin{pmatrix} -3 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

We now solve the system:

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

In other words, $-3x + 2y = 0, -y = 0, y = 0$. Hence $x = y = 0$ and there are no constraints on z. Therefore, the eigenvector corresponding to $\lambda = 4$ is:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and its multiples, i.e. any vector for which the first two components are 0.

In conclusion, the eigenvalues of the matrix $A = \left(\begin{smallmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 4 \end{smallmatrix}\right)$ are $\lambda_1 = 1$, $\lambda_2 = 3$, and $\lambda_3 = 4$; the corresponding eigenvectors are:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and their multiples. □

Up until now I picked examples for which the solutions turn out nicely. Most of the time it's not like that though.

Sample Question

Compute the eigenvalues and eigenvectors of the matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 4 & 7 \end{smallmatrix}\right)$.

Solution. Keep going, even with terrible numbers.

1. We need to find the matrix $A - \lambda I$, where I is the identity matrix. First, we compute $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} 1 & 2 \\ 4 & 7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\lambda & 2 \\ 4 & 7-\lambda \end{pmatrix}.$$

   To find the eigenvalues, we set the determinant of $A - \lambda I$ equal to zero:

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ 4 & 7-\lambda \end{vmatrix} = (1-\lambda)(7-\lambda) - (4 \cdot 2) = (1-\lambda)(7-\lambda) - 8.$$

   Expanding this expression:

$$(1-\lambda)(7-\lambda) = 7 - 8\lambda + \lambda^2,$$

   so the equation becomes:

$$7 - 8\lambda + \lambda^2 - 8 = 0 \implies \lambda^2 - 8\lambda - 1 = 0.$$

   We now solve the quadratic equation $\lambda^2 - 8\lambda - 1 = 0$ using the quadratic formula:

$$\lambda = \frac{-(-8) \pm \sqrt{(-8)^2 - 4(1)(-1)}}{2(1)} = \frac{8 \pm \sqrt{64+4}}{2} = \frac{8 \pm \sqrt{68}}{2} = \frac{8 \pm 2\sqrt{17}}{2}.$$

   Thus, the eigenvalues are:

$$\lambda_1 = 4 + \sqrt{17}, \quad \lambda_2 = 4 - \sqrt{17}.$$

2. Now, we find the eigenvectors corresponding to each eigenvalue.
   - For $\lambda_1 = 4 + \sqrt{17}$: We solve $(A - \lambda_1 I)\mathbf{v} = 0$. First, compute $A - \lambda_1 I$:

$$A - \lambda_1 I = \begin{pmatrix} 1 - \left(4 + \sqrt{17}\right) & 2 \\ 4 & 7 - \left(4 + \sqrt{17}\right) \end{pmatrix} = \begin{pmatrix} -3 - \sqrt{17} & 2 \\ 4 & 3 - \sqrt{17} \end{pmatrix}.$$

We now solve the system:

$$\begin{pmatrix} -3 - \sqrt{17} & 2 \\ 4 & 3 - \sqrt{17} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This gives the equations:

$$\left(-3 - \sqrt{17}\right)x + 2y = 0, \quad 4x + \left(3 - \sqrt{17}\right)y = 0.$$

From the first equation, we get $y = \frac{3+\sqrt{17}}{2}x$. Therefore, the eigenvector corresponding to $\lambda_1 = 4 + \sqrt{17}$ is:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \frac{3+\sqrt{17}}{2} \end{pmatrix}.$$

- For $\lambda_2 = 4 - \sqrt{17}$, it's actually exactly the same with $\sqrt{17}$ replaced by $-\sqrt{17}$, so I won't repeat it. You get the eigenvector

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ \frac{3-\sqrt{17}}{2} \end{pmatrix}.$$

In conclusion the eigenvalues of the matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 4 & 7 \end{smallmatrix}\right)$ are:

$$\lambda_1 = 4 + \sqrt{17}, \quad \lambda_2 = 4 - \sqrt{17}.$$

The corresponding eigenvectors are:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \frac{3+\sqrt{17}}{2} \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ \frac{3-\sqrt{17}}{2} \end{pmatrix}.$$

□

§9.5 [TEXT] What to expect when solving degenerate systems

When carrying out the recipe for finding eigenvectors and eigenvalues, after cherry-picking λ, you have to solve a degenerate system of equations. Since most of the systems of equations you encounter in practice are nondegenerate, here's a few words of advice on instincts for solving the degenerate ones.

§9.5.1 Degenerate systems of two equations all look stupid

This is worth repeating: **degenerate systems of two equations all look stupid**. Earlier on, we saw the two systems

$$\begin{cases} -2x - 2y = 0 \\ 3x + 3y = 0 \end{cases} \quad \text{and} \quad \begin{cases} -3x - 2y = 0 \\ 3x + 2y = 0 \end{cases}.$$

Both look moronic to the eye, because in each equation, the two equations say the same thing. This is by design: when you're solving the eigenvector problem, *you're going out of your way to find degenerate systems* so that there will actually be solutions besides $x = y = 0$.

In particular: if you do all the steps right, **you should never wind up with $x = y = 0$ as your only solution**. That means you either didn't do the cherry-picking step correctly, or something went wrong when you were solving the system. If that happens, check your work!

§9.5.2 Degenerate systems of three equations may or may not look stupid

When you have three or more equations instead, they don't necessarily look as stupid (although they still can). To reuse the example I mentioned before, consider the system of equations

$$\begin{aligned} x + 10y - 9z &= 0 \\ 3x + y + 10z &= 0 \\ 4x + 11y + z &= 0 \end{aligned}$$

which doesn't look stupid. But again, if you check the determinant, you find out

$$\begin{vmatrix} 1 & 10 & -9 \\ 3 & 1 & 10 \\ 4 & 11 & 1 \end{vmatrix} = 0.$$

So you know *a priori* that there will be solutions besides $x = y = z = 0$.

I think 18.02 won't have too many situations where you need to solve a degenerate three-variable system of equations, because it's generally annoying to do by hand. But if it happens, you should fall back on your high school algebra and solve the system however you learned it in 9th or 10th grade. The good news is that at least one of the three equations is redundant, so you can just throw one away and solve for the other two. For example, in this case we would solve

$$\begin{aligned} x + 10y &= 9z \\ 3x + y &= -10z \end{aligned}$$

for x and y, as a function of z. I think this particular example works out to $x = -\frac{109}{29}z$, $y = \frac{37}{29}z$. And it indeed fits the third equation too.

§9.6 [SIDENOTE] Complex eigenvalues

Even in the 2×2 case, you'll find a lot of matrices M with real coefficients don't have eigenvectors. Here's one example.

Let

$$M = \begin{pmatrix} \cos(60°) & -\sin(60°) \\ \sin(60°) & \cos(60°) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

be the matrix corresponding to rotation by 60 degrees. (Feel free to replace 60 by a different number.) I claim that M has no real eigenvalues or eigenvectors.

Indeed, if $\mathbf{v} \in \mathbb{R}^2$ was an eigenvector, then $M\mathbf{v}$ needs to point in the same direction as $\mathbf{v}$, by definition. But that can never happen: M is rotation by 60°, so $M\mathbf{v}$ and $\mathbf{v}$ necessarily point in different directions — 60 degrees apart.

Nevertheless, let's boldly try this and see what goes wrong in the recipe. The answer is that you just get some complex numbers instead.

Sample Question

Compute the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

Solution. Follow the recipe, just don't be scared of complex numbers:

1. We need to find the matrix $A - \lambda I$, where I is the identity matrix. First, we compute $A - \lambda I$:

$$A - \lambda I = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} - \lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{pmatrix}.$$

To find the eigenvalues, we set the determinant of $A - \lambda I$ equal to zero:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} \frac{1}{2} - \lambda & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \lambda \end{vmatrix} \\ &= \left(\frac{1}{2} - \lambda\right)\left(\frac{1}{2} - \lambda\right) - \left(-\frac{\sqrt{3}}{2} \cdot \frac{\sqrt{3}}{2}\right) \\ &= \left(\frac{1}{2} - \lambda\right)^2 + \frac{3}{4}. \end{aligned}$$

Setting this equal to zero and solving, we get

$$\lambda = \frac{1 \pm \sqrt{3}i}{2}$$

as the two eigenvalues.

2. Now, we find the eigenvectors corresponding to each eigenvalue.
   - Choose $\lambda_1 = \frac{1+i\sqrt{3}}{2}$ first. We solve $(A - \lambda_1 I)\mathbf{v} = 0$. First, compute $A - \lambda_1 I$:

$$A - \lambda_1 I = \begin{pmatrix} \frac{1}{2} - \frac{1+i\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} - \frac{1+i\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} -\frac{i\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{i\sqrt{3}}{2} \end{pmatrix}.$$

   We now solve the system:

$$\begin{pmatrix} -\frac{i\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{i\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

   This gives the equations:

$$-\frac{i\sqrt{3}}{2} v_1 - \frac{\sqrt{3}}{2} v_2 = 0, \quad \frac{\sqrt{3}}{2} v_1 - \frac{i\sqrt{3}}{2} v_2 = 0.$$

   From the first equation, we get $v_1 = iv_2$. Therefore, the eigenvector corresponding to $\lambda_1 = \frac{1+i\sqrt{3}}{2}$ is:

$$\mathbf{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$$

and its multiples.

- When $\lambda_2 = \frac{1-i\sqrt{3}}{2}$, all the i's flip to $-i$'s and nothing else changes. The eigenvector corresponding to $\lambda_2 = \frac{1-i\sqrt{3}}{2}$ is thus

$$\mathbf{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

and its multiples.

In conclusion, the two eigenvalues are

$$\lambda_1 = \frac{1+i\sqrt{3}}{2}, \quad \lambda_2 = \frac{1-i\sqrt{3}}{2}.$$

and the corresponding eigenvectors are:

$$\mathbf{v}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

□

§9.7 [TEXT] Trace and determinant

In 18.02 the following definition is briefly mentioned, but we won't do much with it:

Definition of trace

The trace is the sum of the diagonal entries of the matrix.

Then the following two theorems are roughly true:

- The trace of a matrix equals the sum of the eigenvalues, either real or complex.
- The determinant of a matrix equals the product of the eigenvalues, either real or complex.

I say "roughly" because there is a caveat: most of the time, if you have an $n \times n$ matrix, then there will be n different eigenvalues (if you allow complex ones). You probably noticed this above. However, sometimes you'll run into a matrix for which there are fewer than n, and some of the eigenvalues are "repeated", like the example we got where $(\lambda - 5)^2 = 0 \implies \lambda = 5$. We won't define what "repeated" means here, but you need to define repetition correctly to handle these edge cases.

§9.8 [SIDENOTE] Application of eigenvectors: matrix powers

This is off-syllabus for 18.02, but I couldn't resist including it because it shows you a good use of eigenvalues in a seemingly unrelated problem, and also reinforces the idea that I keep axe-grinding:

Idea

If you have a linear operator T, and you know the outputs of T on *any* basis, that tells you all the outputs of T.

Okay, so here's the application I promised you.

Sample Question

Let M be the matrix $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. Calculate M^{100}.

At first glance, you might think this question is obviously impossible without a computer! Raising a matrix to the 100th power seems like it would require 100 matrix multiplications. But I'll show you how to do it with eigenvectors.

Solution. First, we compute the eigenvectors and eigenvalues of M. If you follow the recipe, you'll get the following results:

- The vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector with eigenvalue 2 (as is any multiple of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$), because $M\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- The vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 3 (as is any multiple of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$), because $M\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Now the trick is the following: it's really easy to apply M^{100} to the *eigenvectors*, because it's just multiplication by a constant. For example, the first few powers of M on $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ each double the vector, since they are all eigenvectors with eigenvalue 2; that is:

$$\begin{aligned} M\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ M^2\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= M\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \\ M^3\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= M\begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \end{pmatrix} \\ &\vdots \end{aligned}$$

and so on, until

$$M^{100}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2^{100}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

By the same token:

$$M^{100}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3^{100}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

So now we know the outputs of M^{100} at two linearly independent vectors. It would be sufficient, then, to use this information to extract $M^{100}(\mathbf{e}_1)$ and $M^{100}(\mathbf{e}_2)$. We can now rewrite this as

$$M^{100}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2^{100} \\ 0 \end{pmatrix}; \qquad M^{100}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = M^{100}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - M^{100}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3^{100} - 2^{100} \\ 3^{100} \end{pmatrix}.$$

Thus encoding M gives the answer:

$$M^{100} = \begin{pmatrix} 2^{100} & 3^{100} - 2^{100} \\ 0 & 3^{100} \end{pmatrix}.$$

□

§9.9 [EXER] Exercises

Exercise 9.1. Compute the eigenvalues and eigenvectors for the following matrices:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 1 \\ 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix}, \quad D = \begin{pmatrix} 6 & 1 \\ 0 & 6 \end{pmatrix}.$$

Exercise 9.2. Give an example of a 2×2 matrix T with four nonzero entries whose eigenvalues are 5 and 7. Then compute the corresponding eigenvectors.

Exercise 9.3 (*). Compute the eigenvectors and eigenvalues of the 6×6 matrix

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{pmatrix}.$$

(You can do this question without using any determinants.)

Exercise 9.4 (*). Using the procedure described in Section 9.8, show that

$$\begin{pmatrix} 4 & 3 \\ 6 & 7 \end{pmatrix}^{20} = \begin{pmatrix} 33333333333333333334 & 33333333333333333333 \\ 66666666666666666666 & 66666666666666666667 \end{pmatrix}.$$

(Each number on the right-hand side is 20 digits.)

Part Charlie: Review of complex numbers

For comparison, Part Charlie corresponds roughly to §11 of Poonen's notes.

Chapter 10. Complex numbers

§10.1 [TEXT] It's a miracle that multiplication in $\mathbb{C}$ has geometric meaning

Let $\mathbb{C}$ denote the set of complex numbers (just as $\mathbb{R}$ denotes the real numbers). It's important that realize that, **until we add in complex multiplication, $\mathbb{C}$ is just an elaborate $\mathbb{R}^2$ cosplay**.

Concept	For $\mathbb{R}^2$	For $\mathbb{C}$
Notation	$\mathbf{v}$	z
Components	$\begin{pmatrix} x \\ y \end{pmatrix}$	$x + yi$
Length	Length $\lvert \mathbf{v} \rvert$	Abs val $\lvert z \rvert$
Direction	(slope, maybe?)	argument θ
Length 1	unit vector	$e^{i\theta} = \cos\theta + i\sin\theta$
Multiply	*NONE*	✨ $z_1 z_2$ ✨

At the start of the course, I warned you about type safety, and I repeatedly stressed you that you **cannot multiply two vectors in $\mathbb{R}^n$ to get another vector**. (You had a "dot product", but it spits out a number. Honestly, you shouldn't think of dot product as a "product"; the name sucks.)

Of course, the classic newbie mistake (which you better not make on your midterm) is to define a product on vectors component-wise: why can't $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ have "product" $\begin{pmatrix} a_1 b_1 \\ \vdots \\ a_n b_n \end{pmatrix}$? Well, in 18.02, every vector definition needed a corresponding geometric picture for us to consider it worthy of attention. This definition has no geometric meaning.

However, there is a big miracle for $\mathbb{C}$. For complex numbers, you can define multiplication by

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

and there is *an amazing geometric interpretation.*

Unfortunately, AFAIK there is no English word for "complex number whose absolute value is one", the same way there is for "unit vector". For 18.02, we instead use

$$e^{i\theta} := \cos\theta + i\sin\theta$$

as the "word"; whenever you see $e^{i\theta}$, draw it as unit vector $\cos\theta + i\sin\theta$.

> **What does complex exponents mean anyway?**
>
> It's worth pointing out the notation $e^{i\theta}$ should strike you as a type-error based on what you've learned in school. What meaning does it have to raise a number to an imaginary power? Does i^i have a meaning? Does $\cos(i)$ have a meaning? If you want to know, check Chapter 55 in the Appendix.

But for 18.02, when starting out, I would actually think of the notation $e^{i\theta}$ as a *mnemonic*, i.e. a way to remember the following result:

$$\underbrace{(\cos\theta_1 + i\sin\theta_1)}_{=e^{i\theta_1}} \cdot \underbrace{(\cos\theta_2 + i\sin\theta_2)}_{=e^{i\theta_2}} = \underbrace{\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)}_{=e^{i(\theta_1+\theta_2)}}. \tag{6}$$

This is in my opinion the biggest miracle in all of precalculus. Really, I want to stress: Equation 6 is supposed to be *astonishing*. My goal by the end of this chapter is to convince you that something really powerful is happening in Equation 6 allowing you to do things that you absolutely should not expect to be able to do.

More generally, the result you need to know is:

Memorize: Complex multiplication

Suppose z_1 and z_2 are complex numbers. To describe $z_1 z_2$:

- The magnitude of $z_1 z_2$ is the *product* of the magnitudes of z_1 and z_2. That is,
$$|z_1 z_2| = |z_1|\,|z_2| \tag{7}$$
- The argument of $z_1 z_2$ is the *sum* of the arguments of z_1 and z_2. This is Equation 6.

Here's a simple example.

Example

Let's consider the complex numbers:

$$z_1 = 20 + 21i \quad \text{and} \quad z_2 = 5 + 12i.$$

Following your high school, the product $z_1 \cdot z_2$ is calculated as follows:

$$\begin{aligned} z_1 \cdot z_2 &= (20+21i)(5+12i) = 20\cdot 5 + 20 \cdot 12i + 21i \cdot 5 + 21i \cdot 12i \\ &= 100 + 240i + 105i + 252i^2 \\ &= 100 + 345i + 252(-1) \quad (\text{since} \quad i^2 = -1) \\ &= 100 + 345i - 252 = (100-252) + 345i = -152 + 345i. \end{aligned}$$

The above theorem is promising that if we had used polar form, the *angles* will add and the *magnitudes* will multiply. Let's verify this holds up.

For the magnitudes, you can do this by hand: we have $|z_1| = \sqrt{20^2+21^2} = 29$ and $|z_2| = \sqrt{5^2+12^2} = 13$, and indeed we have the miraculous $|z_1 z_2| = \sqrt{152^2 + 345^2} = 377$.

The angles here probably need a calculator to verify. For the angles, from $\arctan\left(\frac{21}{20}\right) \approx 46.04°$ and $\arctan\left(\frac{12}{5}\right) \approx 67.38°$, we have

$$\begin{aligned} z_1 &\approx 29(\cos 46.04° + i \sin 46.04°) \\ z_2 &\approx 13(\cos 67.38° + i \sin 67.38°) \end{aligned}$$

so we're expecting that

$$z_1 z_2 \approx 377(\cos 113° + i \sin 113°)$$

and indeed $\arctan\left(-\frac{345}{152}\right) \approx 113°$, as needed!

Here's a more substantial example, which shows how Equation 6 can be used to compute things that wouldn't be feasible by hand with the rectangular form.

Sample Question

Compute $(1+i)^{10}$.

Solution. The idea is that we will write $1+i$ in polar form as:

$$1+i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right).$$

Then raising powers is easy, because of Equation 6. To spell it out:

$$\begin{aligned}
(1+i)^2 &= \left(\sqrt{2}\right)^2\left(\cos\left(2\cdot\frac{\pi}{4}\right) + i\sin\left(2\cdot\frac{\pi}{4}\right)\right) \\
(1+i)^3 &= \left(\sqrt{2}\right)^3\left(\cos\left(3\cdot\frac{\pi}{4}\right) + i\sin\left(3\cdot\frac{\pi}{4}\right)\right) \\
(1+i)^4 &= \left(\sqrt{2}\right)^4\left(\cos\left(4\cdot\frac{\pi}{4}\right) + i\sin\left(4\cdot\frac{\pi}{4}\right)\right) \\
(1+i)^5 &= \left(\sqrt{2}\right)^5\left(\cos\left(5\cdot\frac{\pi}{4}\right) + i\sin\left(5\cdot\frac{\pi}{4}\right)\right) \\
&\vdots \\
(1+i)^{10} &= \left(\sqrt{2}\right)^{10}\left(\cos\left(10\cdot\frac{\pi}{4}\right) + i\sin\left(10\cdot\frac{\pi}{4}\right)\right).
\end{aligned}$$

We can simplify this now: we know

$$(1+i)^{10} = 2^5\left(\cos\left(5\cdot\frac{\pi}{2}\right) + i\sin\left(5\cdot\frac{\pi}{2}\right)\right) = 32(0+i) = \boxed{32i}. \qquad \square$$

Compare to how annoying this would be if we tried to do it by multiplying 10 times: the fastest way with repeated squaring would be something like

$$\begin{aligned}
(1+i)^2 &= (1)^2 + 2\cdot 1\cdot i + i^2 = 1 + 2i + (-1) = 2i \\
(1+i)^4 &= \left((1+i)^2\right)^2 = (2i)^2 = 4i^2 = 4\cdot(-1) = -4 \\
(1+i)^8 &= \left((1+i)^4\right)^2 = (-4)^2 = 16 \\
(1+i)^{10} &= (1+i)^8\cdot(1+i)^2 = 16\cdot 2i = 32i.
\end{aligned}$$

But you could easily imagine replacing 10 with 100 (which we'll do shortly) or even 1000000. Such a method would quickly become infeasible; whereas the polar coordinates let us avoid all this work.

§10.2 [SIDENOTE] Extracting trig identities and the Brahmagupta-Fibonacci identity

In this optional section I want to convince you that Equation 7 and Equation 6 are doing a lot of magic. To do so I'll show you two consequences of these equations that you would not expect to be true.

§10.2.1 Application 1: Equation 7 gives the Brahmagupta-Fibonacci identity

Let's start with *magnitudes*. If you don't trust your teacher (a good instinct to have sometimes 😉) you might not *believe* me the magnitudes multiply. Because let's say $z_1 = a + bi$ and $z_2 = c + di$. Then

$$z_1 z_2 = (a+bi)(c+di) = (ac-bd) + (ad+bc)i$$

doesn't look anything right. If Equation 7 is really true, it's promising that

$$\sqrt{a^2+b^2} \cdot \sqrt{c^2+d^2} = \sqrt{(ac-bd)^2+(ad+bc)^2}.$$

In other words, the equation

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2+(ad+bc)^2 \tag{8}$$

is supposed to be true for *all* real numbers a, b, c, d.

But how could that be? Equation 8 doesn't even *look* true, and if I told this to you with no context, you wouldn't believe it. It's not until you multiply out Equation 8 with brute force that you might believe me:

$$\begin{aligned}(a^2+b^2)(c^2+d^2) &= a^2c^2+b^2c^2+a^2d^2+b^2d^2 \\ (ac-bd)^2+(ad+bc)^2 &= (a^2c^2-2abcd+b^2d^2)+(a^2d^2+2abcd+b^2c^2) \\ &= a^2c^2+b^2d^2+a^2d^2+b^2c^2.\end{aligned}$$

They really are equal! The $2abcd$ has apparently cancelled out magically. (This unexpected identity is called the Brahmagupta-Fibonacci identity, if you want a name, but we won't use this name again later.)

§10.2.2 Application 2: Equation 6 gives trig addition and double-angle formulas

Let's say you still don't trust your teacher (again, good!) and even though you have grudgingly admitted Equation 7 is true, you don't believe the other equation Equation 6. Because if Equation 6 is true, then again brute-force expansion gives

$$\begin{aligned}\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2) &= (\cos\theta_1+i\sin\theta_1)(\cos\theta_2+i\sin\theta_2) \\ &= (\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+i(\sin\theta_1\cos\theta_2+\sin\theta_2\cos\theta_1).\end{aligned}$$

So for Equation 6 to be true, you would need for *any* angles θ_1 and θ_2 that

$$\begin{aligned}\cos(\theta_1+\theta_2) &= \cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2 \\ \sin(\theta_1+\theta_2) &= \sin\theta_1\cos\theta_2+\cos\theta_1\sin\theta_2.\end{aligned}$$

But this *is* true: it's the trig addition formula!

Put another way: if you have trouble remembering the trig addition formulas (like me), then Equation 6 shows you how you can derive it. Equation 6 is easy to remember, and if you do the expansion, the mysterious trig addition formula falls out.

The double angle formula is also a special case: from

$$\cos(2\theta)+i\sin(2\theta) = (\cos\theta+i\sin\theta)^2 = (\cos^2\theta-\sin^2\theta)+i\cdot 2\sin\theta\cos\theta$$

we can read off $\cos(2\theta)=\cos^2\theta-\sin^2\theta$ and $\sin(2\theta)=2\sin\theta\cos\theta$.

§10.3 [RECIPE] nth powers of complex numbers

Earlier I showed you how to calculate $(1+i)^{10}$ rapidly using polar form. You can do this in general too: the point is that

$$(r(\cos\theta+i\sin\theta))^n = r^n(\cos n\theta+i\sin n\theta). \tag{9}$$

Equation 9 is sometimes called *De Moivre's theorem*, but it's such an easy consequence of Equation 6 and Equation 7 that I don't think it really needs its own name. Nonetheless, if you see the name in other places, it's referring to Equation 9.

Recipe for raising a complex number to the nth power

Given a complex number z, to compute z^n:

1. Convert $z = r(\cos\theta + i\sin\theta)$ in polar form if it isn't already.
2. Use $z^n = r^n(\cos n\theta + i\sin n\theta)$.
3. Simplify the $\cos n\theta + i\sin n\theta$ and output the answer.

Here's the example with $(1+i)^{10}$ again, but with 10 replaced by 100 for emphasis.

Sample Question

Compute $(1+i)^{100}$.

Solution. The polar form of $1+i$ is:

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right).$$

Raising to the 100th power gives

$$\begin{aligned}(1+i)^{100} &= \left(\sqrt{2}\right)^{100}\left(\cos\left(100\cdot\frac{\pi}{4}\right) + i\sin\left(100\cdot\frac{\pi}{4}\right)\right)\\ &= 2^{50}(\cos(25\pi) + i\sin(25\pi))\\ &= 2^{50}(\cos(\pi) + i\sin(\pi)) = \boxed{-2^{50}}.\end{aligned}$$

□

Sample Question

Compute $\left(1-\sqrt{3}i\right)^{20}$.

Solution. First convert $1-\sqrt{3}i$ to polar form:

$$1 - \sqrt{3}i = 2\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right).$$

Then, when we raise to the 20th power, we get

$$\left(1-\sqrt{3}i\right)^{20} = 2^{20}\left(\cos\left(\frac{100\pi}{3}\right) + i\sin\left(\frac{100\pi}{3}\right)\right).$$

The cosine and sine cycle every 2π, so we write

$$\frac{100\pi}{3} = \frac{4\pi}{3} + 2\pi\cdot 16.$$

We only care about the "remainder" $4\frac{\pi}{3}$; we have

$$\cos\left(\frac{100\pi}{3}\right) = \cos(240°) = -\frac{1}{2}$$

$$\sin\left(\frac{100\pi}{3}\right) = \sin(240°) = -\frac{\sqrt{3}}{2}.$$

Substituting back:

$$\begin{aligned}\left(1-\sqrt{3}i\right)^{20} &= 2^{20}\left(-\frac{1}{2}-i\frac{\sqrt{3}}{2}\right) \\ &= 2^{20}\times\left(-\frac{1+\sqrt{3}i}{2}\right) = \boxed{-2^{19}\left(1+\sqrt{3}i\right)}.\end{aligned}$$

(If you care, $2^{19} = 524288$, so one could also write $-524288 - 524288\sqrt{3}i$.) □

§10.4 [TEXT] An example of nth roots of complex numbers: solving $z^5 = 243i$

This section is dedicated to z^n and is on-syllabus for exam. Specifically, you ought to be able to solve equations like $z^5 = 243i$. This section shows you how.

In this whole section, you always prefer to work in polar form. So if you get input in rectangular form, you should first convert to rectangular form. Conversely, if the answer is asked for in rectangular form, you should work with polar form anyway, and only convert to rectangular output at the end.

If you can raise to the nth power, you should be able to extract nth roots too, by running the recipe "backwards". First, I will tell you what the answer looks like:

> **! Memorize: The shape of an nth root answer**
>
> Consider solving the equation $z^n = w$, where w is a given nonzero complex number, for z. Then you should always output exactly n answers. Those n answers all have magnitude $|w|^{\frac{1}{n}}$ and arguments spaced apart by $\frac{360°}{n}$.

I think it's most illustrative if I show you the five answers to

$$z^5 = 243i$$

to start. Again, first we want to convert everything to polar coordinates:

$$z^5 = 243i = 243(\cos 90° + i\sin 90°).$$

At this point, we know that if $|z^5| = 243$, then $|z| = 3$; all the answers should have absolute 3. So the idea is to find the angles. Here are the five answers:

$$\begin{aligned}
z_1 &= 3(\cos 18° + i\sin 18°) \implies (z_1)^5 = 243(\cos 90° + i\sin 90°) \\
z_2 &= 3(\cos 90° + i\sin 90°) \implies (z_2)^5 = 243(\cos 450° + i\sin 450°) \\
z_3 &= 3(\cos 162° + i\sin 162°) \implies (z_3)^5 = 243(\cos 810° + i\sin 810°) \\
z_4 &= 3(\cos 234° + i\sin 234°) \implies (z_4)^5 = 243(\cos 1170° + i\sin 1170°) \\
z_5 &= 3(\cos 306° + i\sin 306°) \implies (z_5)^5 = 243(\cos 1530° + i\sin 1530°).
\end{aligned}$$

Here's a picture of the five numbers:

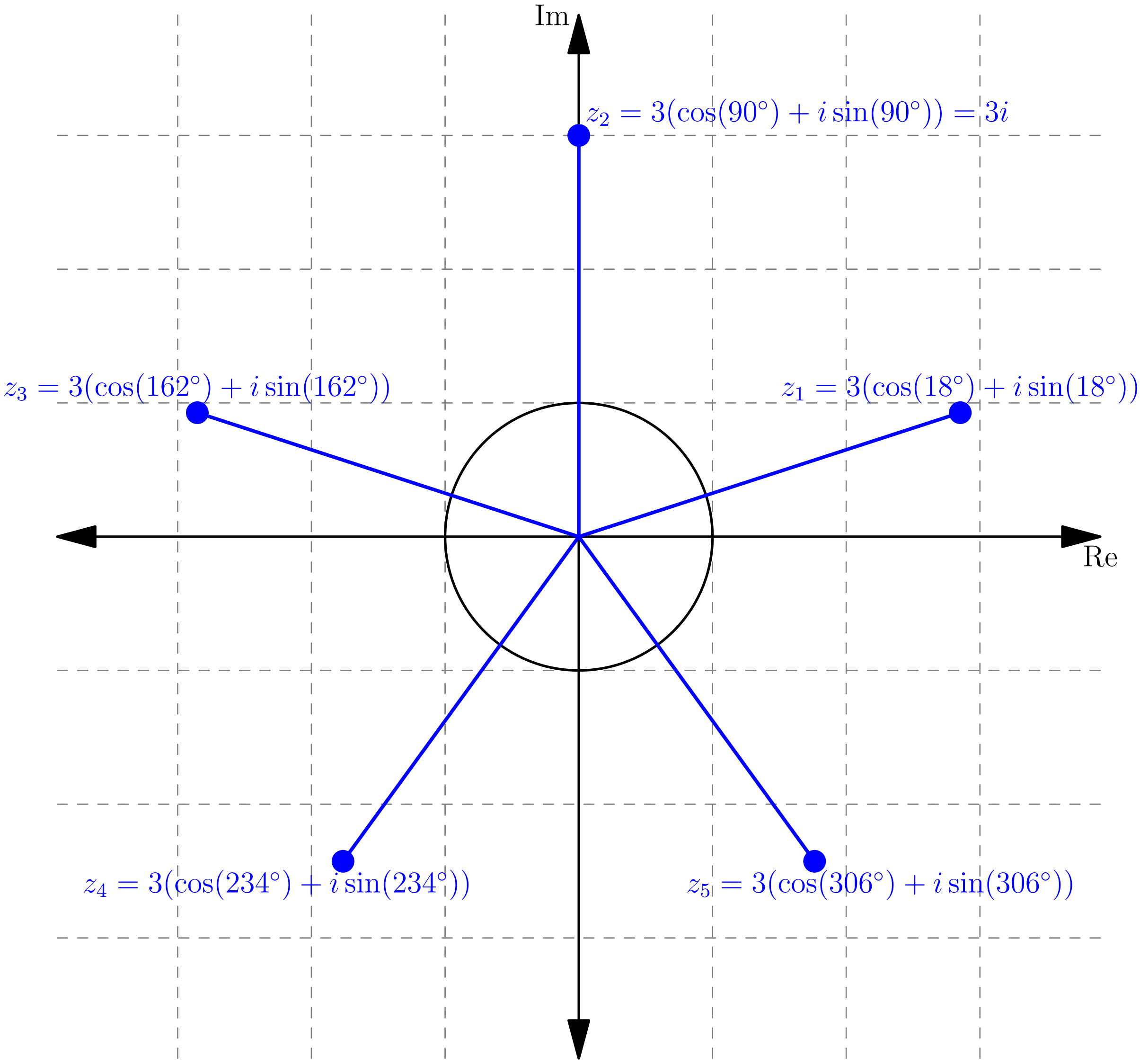

Figure 17: The five answers to $z^5 = 243i$, each of length 3.

On the right column, all the numbers are equal. Notice something interesting happening on the right-hand side. The numbers $\cos 90° + i\sin 90°$ and $\cos 450° + i\sin 450°$, etc. are all the same number; if you draw them in the plane, they'll point to the same thing. However, they give five *different* answers on the left. But if you continue the pattern one more, you start getting a cycle

$$z_6 = 3(\cos 378° + i\sin 378°) \implies (z_6)^5 = 243(\cos 1890° + i\sin 1890°).$$

This doesn't give you a new answer, because $z_6 = z_1$.

§10.5 [RECIPE] Taking the nth root of a complex number

In general, if w has argument θ, then the arguments of z satisfying $z^n = w$ start at $\frac{\theta}{n}$ and then go up in increments of $\frac{360°}{n}$. (For example, they started at $\frac{90°}{5} = 18°$ for answers to $z^5 = 243i$.) So you can describe the general recipe as:

Recipe for nth roots of complex numbers

1. Convert w to polar form; say it has angle θ.
2. One of the n answers will be $|w|^{\frac{1}{n}}\left(\cos\frac{\theta}{n}+i\sin\frac{\theta}{n}\right)$.
3. The other $n-1$ answers are obtained by increasing the angle in increments of $\frac{360°}{n}$.

Sample Question

Solve $z^5 = 243i$ for $z \in \mathbb{C}$.

Solution. We first convert to polar form as

$$243i = 243(\cos 90° + i \sin 90°)$$

and see that $243^{\frac{1}{5}} = 3$, and $\theta = 90°$. The first angle is $\frac{\theta}{5} = 18°$. So the five answers are

$$\begin{aligned} z_1 &= 3(\cos 18° + i\sin 18°) \\ z_2 &= 3(\cos 90° + i\sin 90°) \\ z_3 &= 3(\cos 162° + i\sin 162°) \\ z_4 &= 3(\cos 234° + i\sin 234°) \\ z_5 &= 3(\cos 306° + i\sin 306°). \end{aligned}$$

(As it happens, $z_2 = 3i$, which is easy to check by hand works.) □

Sample Question

Solve $z^4 = 8 + 8\sqrt{3}i$ for $z \in \mathbb{C}$.

Solution. We first convert to polar form as

$$8 + 8\sqrt{3}i = 16(\cos 60° + i\sin 60°)$$

and see that $16^{\frac{1}{4}} = 2$, and $\theta = 60°$. The first angle is $\frac{\theta}{4} = 15°$. So the four answers are

$$\begin{aligned} z_1 &= 2(\cos 15° + i\sin 15°) \\ z_2 &= 2(\cos 105° + i\sin 105°) \\ z_3 &= 2(\cos 195° + i\sin 195°) \\ z_4 &= 2(\cos 285° + i\sin 285°). \end{aligned}$$

□

Sample Question

Solve $z^3 = -1000$ for $z \in \mathbb{C}$.

Solution. We first convert to polar form as

$$-1000 = 1000(\cos 180° + i\sin 180°)$$

and see that $1000^{\frac{1}{3}} = 10$, and $\theta = 180°$. The first angle is $\frac{\theta}{3} = 60°$. So the three answers are

$$\begin{aligned} z_1 &= 10(\cos 60^\circ + i \sin 60^\circ) \\ z_2 &= 10(\cos 180^\circ + i \sin 180^\circ) \\ z_3 &= 10(\cos 300^\circ + i \sin 300^\circ). \end{aligned}$$

(As it happens, $z_2 = -10$, as expected, since $(-10)^3 = -1000$.) □

§10.6 [RECAP] Rectangular vs polar

Every complex number can be written in either *rectangular form* ($a + bi$ for $a, b \in \mathbb{R}$) or *polar form* ($re^{i\theta}$). We saw that polar form (because of Equation 6 and Equation 7) is really good if you're doing lots of multiplication. So to summarize, Table 4 tells you rules of thumb for complex numbers.

Operation	In rectangular	In polar
$z_1 \pm z_2$	✅ Component-wise like in $\mathbb{R}^2$	❌ Unless z_1 is a real multiple of z_2
$z_1 z_2$	✅ Expanding	✅ By Equation 6 + Equation 7
z_1 / z_2	✅ Write $\frac{1}{c+di} = \frac{c-di}{c^2+d^2}$ then multiply	✅ By Equation 6 + Equation 7
z^n	❌ Possible but takes forever	✅ Shown in Section 10.3
n^{th} root of z	❌ Not recommended for $n > 1$	✅ Shown in Section 10.5

§10.7 [EXER] Exercises

Exercise 10.1 (*). Without a calculator, give an example of an ordered pair (a, b) of integers satisfying

$$a^2 + b^2 = 101 \cdot 401 \cdot 901.$$

Chapter 11. Challenge review problems for Parts Alfa, Bravo, and Charlie

This set of problems is intended to be more difficult. You can try them here if you want, but don't be discouraged if you find the problems tricky. All of these are much harder than anything that showed up on the actual midterm. Solutions to these six exercises are in Chapter 44.

(Suggested usage: think about each for 20-30 minutes, then read the solution. I tried to craft problems that teach deep understanding and piece together multiple ideas, rather than just using one or two isolated recipes.)

Exercise 11.1. In $\mathbb{R}^3$, compute the projection of the vector $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ onto the plane $x + y + 2z = 0$.

Exercise 11.2 (*). Suppose A, B, C, D are points in $\mathbb{R}^3$. Give a geometric interpretation for this expression:

$$\frac{1}{6}\left|\overrightarrow{DA} \cdot \left(\overrightarrow{DB} \times \overrightarrow{DC}\right)\right|.$$

Exercise 11.3 (*). Fix a plane $\mathcal{P}$ in $\mathbb{R}^3$ which passes through the origin. Consider the linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$ where $f(\mathbf{v})$ is the projection of $\mathbf{v}$ onto $\mathcal{P}$. Let M denote the 3×3 matrix associated to f. Compute the determinant of M.

Exercise 11.4 (*). Let $\mathbf{a}$ and $\mathbf{b}$ be two perpendicular unit vectors in $\mathbb{R}^3$. A third vector $\mathbf{v}$ in $\mathbb{R}^3$ lies in the span of $\mathbf{a}$ and $\mathbf{b}$. Given that $\mathbf{v} \cdot \mathbf{a} = 2$ and $\mathbf{v} \cdot \mathbf{b} = 3$, compute the magnitudes of the cross products $\mathbf{v} \times \mathbf{a}$ and $\mathbf{v} \times \mathbf{b}$.

Exercise 11.5. Compute the trace of the 2×2 matrix M given the two equations

$$M\begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \end{pmatrix} \quad \text{and} \quad M\begin{pmatrix} 5 \\ 9 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}.$$

Exercise 11.6. There are three complex numbers z satisfying $z^3 = 5 + 6i$. Suppose we plot these three numbers in the complex plane. Compute the area of the triangle they enclose.

Part Delta: Parametric side-quest

For comparison, Part Delta corresponds roughly to §5 and §7 of Poonen's notes.

Chapter 12. Parametric equations

§12.1 [TEXT] Multivariate domains vs multivariate codomains

In 18.01, you did calculus on functions $F : \mathbb{R} \to \mathbb{R}$. So "multivariable calculus" could mean one of two things to start:

- Work with $F : \mathbb{R} \to \mathbb{R}^n$ instead (i.e. make the codomain multivariate).
- Work with $F : \mathbb{R}^n \to \mathbb{R}$ instead (i.e. make the domain multivariate).

What you should know now is **the first thing is WAY easier than the second**. This Part Delta is thus really short.

§12.2 [TEXT] Parametric pictures

From now on, we're going to usually change notation

$$\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$$
$$\mathbf{r}(t) = \begin{pmatrix} \text{function in } t \\ \vdots \\ \text{function in } t \end{pmatrix}.$$

The choice of letter t for the input variable usually means "time"; and we use $\mathbf{r}$ for the function name to remind that the output is a vector. A cartoon of this is shown in Figure 18.

</> Type signature

When you see $\mathbf{r}(t)$ or similar notation, the time variable t has type scalar. The output is in $\mathbb{R}^n$, and depending on context, you can think of it as either a point or a vector.

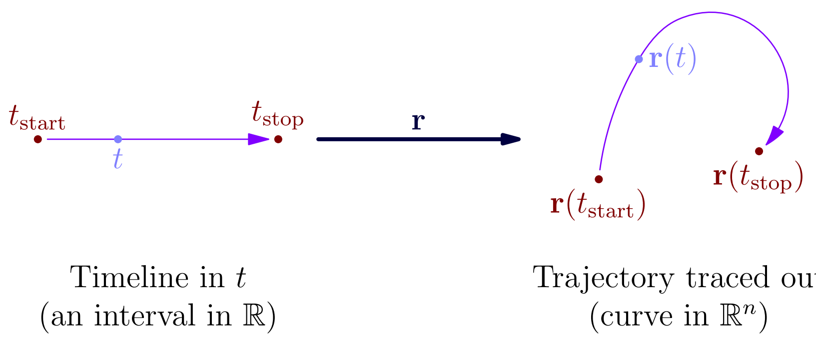

Figure 18: We think of $\mathbf{r}(t)$ as using a timeline in some variable t to trace out a path of some sort in $\mathbb{R}^n$.

Warning: $\mathbf{r}(t)$ can be drawn as either a dot or arrow, but we still use vector notation anyway in Part Delta

Unfortunately, even in cases where we think of $\mathbf{r}(t)$ as a point like $(3, 5)$, we still use boldface letter $\mathbf{r}$ and write $\binom{3}{5}$. Type enthusiasts may rightfully object to this, but this is so entrenched that it will cause confusion with other sources if I'm too pedantic.

So, don't worry too much about the difference between dot and arrow in this chapter. Throughout all of Part Delta we will not treat $(3, 5)$ and $\binom{3}{5}$ as different.

If you're drawing a picture of a parametric function, usually all the axes are components of $\mathbf{r}(t)$ and the time variable doesn't have an axis. In other words, in the picture, **all the axis variables are output components, and we treat them all with equal respect**. The input time variable doesn't show up at all. (This is in contrast to 18.01 xy-graphs, where one axis was input and one axis was output. In the next section when we talk about *level curves*, it will be the other way around, where the output variable is anonymous and every axis is an input variable we treat with equal respect.)

Example

The classic example

$$\mathbf{r}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

would be drawn as the unit circle. You can imagine a particle starting at $\mathbf{r}(0) = \binom{1}{0}$ and then moving around the unit circle counterclockwise with constant speed. It completes a full revolution in 2π time: $\mathbf{r}(2\pi) = \binom{1}{0}$.

§12.3 [TEXT] Just always use components

Why is $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ so easy that Part Delta is one chapter? Because there's pretty much only one thing you need to ever do:

Idea

TLDR Just always use components.

That is, if $\mathbf{r} : \mathbb{R} \to \mathbb{R}^3$ (say), basically 90%+ of the time what you do is write

$$\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3$$

and then just do single-variable calculus or calculations on each f_i.

- Need to differentiate $\mathbf{r}$? Differentiate each component.
- Need to integrate $\mathbf{r}$? Integrate each component.
- Need the absolute value of $\mathbf{r}$? Square root of sum of squares of components.

And so on. An example of Evan failing to do this is shown in Figure 19.

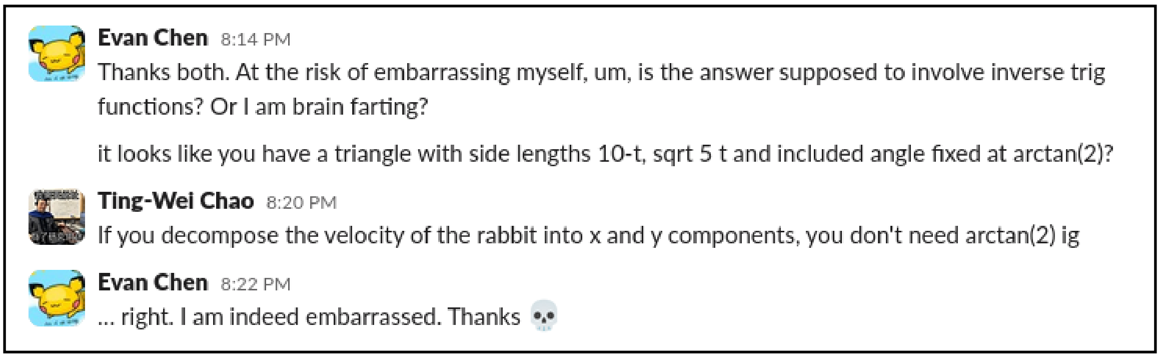

Figure 19: Seriously, just do everything componentwise.

§12.4 [RECIPE] Parametric things

I'll write this recipe with two variables, but it works equally well for three. Suppose you're given an equation $\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$. There are some things you could be asked:

Recipe/definitions for parametric stuff

- The **velocity vector** at a time t is defined as the derivative
$$\mathbf{r}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}.$$
- The **speed** at a time t is defined as the absolute value of the velocity:
$$|\mathbf{r}'(t)| = \sqrt{x'(t)^2 + y'(t)^2}.$$
- The **acceleration vector** at a time t is defined as the second derivative of each component:
$$\mathbf{r}''(t) = \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix}.$$

For three-variable $\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$, do the same thing with three components.

I don't know if there's a word for the absolute value of the acceleration vector (the way speed is the absolute value of the velocity vector).

One more thing to mention now:

Recipe/definition for arc length

The **arc length** from time t_{start} to t_{stop} is defined as the integral of the speed:
$$\text{arc length} = \int_{t=\text{start time}}^{\text{stop time}} |\mathbf{r}'(t)| \, dt.$$

(Technically, I should use "definition" boxes rather than "recipe" boxes here, since these are really the *definition* of the terms involved, and the recipes are "use the definition verbatim".)

</> **Type signature**

- Velocity $\mathbf{r}'(t)$ and acceleration $\mathbf{r}''(t)$, are vectors. In these cases, you should *always* draw them as arrows (vectors) rather than dots. That is, you should never draw velocity or acceleration as a dot.
- However, speed $|\mathbf{r}'(t)|$ and arc length are scalars (numbers).

Sample Question

Let

$$\mathbf{r}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Calculate:

- The velocity vector at time $t = \frac{\pi}{3}$.
- The speed at time $t = \frac{\pi}{3}$.
- The acceleration vector at time $t = \frac{\pi}{3}$.
- The arc length from $t = 0$ to $t = \frac{\pi}{3}$.

Solution. Let $\mathbf{r}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$. We will compute the following quantities.

Velocity vector at $t = \frac{\pi}{3}$ The velocity vector is the derivative of the position vector $\mathbf{r}(t)$ with respect to t:

$$\mathbf{v}(t) = \mathbf{r}'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}.$$

At $t = \frac{\pi}{3}$, we have:

$$\mathbf{v}\left(\frac{\pi}{3}\right) = \begin{pmatrix} -\sin\left(\frac{\pi}{3}\right) \\ \cos\left(\frac{\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Thus, the velocity vector at $t = \frac{\pi}{3}$ is:

$$\mathbf{v}\left(\frac{\pi}{3}\right) = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Speed at $t = \frac{\pi}{3}$ The speed is the magnitude of the velocity vector:

$$|\mathbf{v}(t)| = \sqrt{(-\sin(t))^2 + (\cos(t))^2} = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

Thus, the speed at $t = \frac{\pi}{3}$ (or in fact any time) is:

$$\left|\mathbf{v}\left(\frac{\pi}{3}\right)\right| = 1.$$

Acceleration vector at $t = \frac{\pi}{3}$ Differentiate the velocity vector we got earlier:

$$\mathbf{a}(t) = \mathbf{v}'(t) = \begin{pmatrix} -\cos(t) \\ -\sin(t) \end{pmatrix}.$$

At $t = \frac{\pi}{3}$, we have:

$$\mathbf{a}\left(\frac{\pi}{3}\right) = \begin{pmatrix} -\cos\left(\frac{\pi}{3}\right) \\ -\sin\left(\frac{\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}.$$

Thus, the acceleration vector at $t = \frac{\pi}{3}$ is:

$$\mathbf{a}\left(\frac{\pi}{3}\right) = \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}.$$

Arc length The arc length of a parametric curve is given by:

$$L = \int_0^{\frac{\pi}{3}} |\mathbf{r}'(t)| \, \mathrm{d}t = \int_0^{\frac{\pi}{3}} 1 \, \mathrm{d}t = \frac{\pi}{3}.$$

Thus, the arc length from $t = 0$ to $t = \frac{\pi}{3}$ is:

$$L = \frac{\pi}{3}.$$

□

§12.5 [TEXT] Constant velocity and angular velocity

In 18.02, we will see some complicated trajectories which are actually the sum of two simpler ones. So we start by describing some examples of simple trajectories in this section; then in the next section we start adding some of them together.

Constant velocity is easy: if you have a point that starts from a point A_0 and moves in a straight line with velocity $\mathbf{v}$, then the parametrization is

$$\mathbf{A}(t) = A_0 + t\mathbf{v}.$$

Sample Question

A point P starts at $(1, 2, 3)$ and moves with constant velocity 5 in the $+x$ direction. Parametrize the position $\mathbf{P}(t)$.

Solution. Just write

$$\mathbf{P}(t) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5t + 1 \\ 2 \\ 3 \end{pmatrix}.$$

□

Rotation is actually also pretty simple, but it uses the term "angular velocity" instead. If you haven't seen the term angular velocity, we describe it now.

Definition

An object is said to have *angular velocity* ω if it rotates at a rate of ω radians per unit time. For example, an angular velocity of "10π per second" means the object completes five rotations (of 2π radians each) every second.

Suppose a point P moves in a circle of radius r around $(0, 0)$ with constant angular velocity ω. Then the point can always be written as

$$(r\cos(\theta), r\sin(\theta))$$

for some angle θ that varies with t. A counterclockwise angular velocity corresponds to θ increasing by ω per unit time (hence the angle at time t is $\theta + t\omega$); clockwise is decreasing by ω per unit time instead (hence the angle at time t is $\theta - t\omega$). See Figure 20.

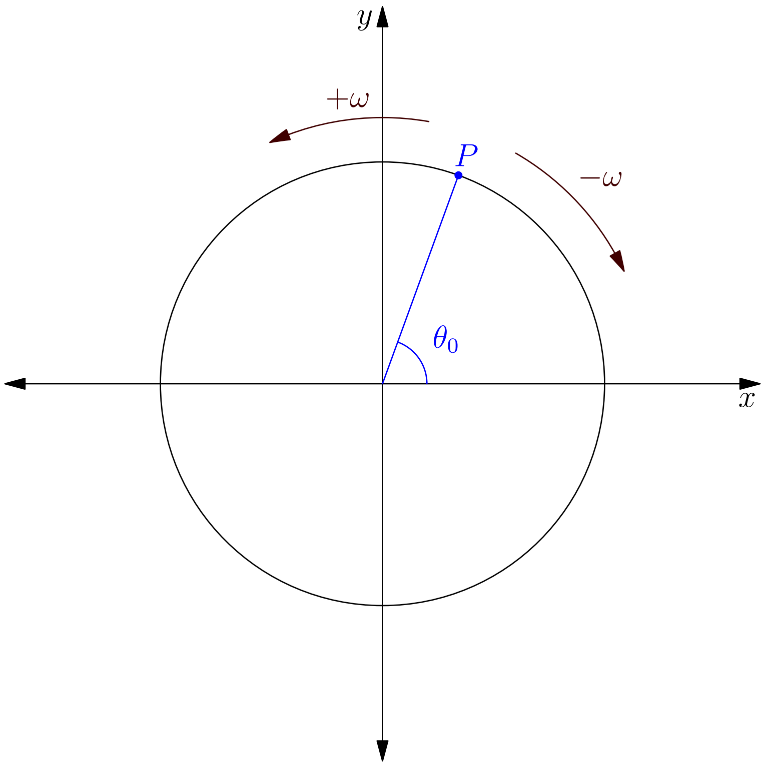

Figure 20: Rotation of a point with constant angular velocity.

Recipe for motion with constant angular velocity

1. Find the initial angle θ_0 corresponding to the position at time $t = 0$.
2. If the motion is counterclockwise, output

$$\mathbf{P}(t) = \begin{pmatrix} r\cos(\theta_0 + \omega t) \\ r\sin(\theta_0 + \omega t) \end{pmatrix}.$$

If it's clockwise instead output

$$\mathbf{P}(t) = \begin{pmatrix} r\cos(\theta_0 - \omega t) \\ r\sin(\theta_0 - \omega t) \end{pmatrix}.$$

(Note the change from $+$ to $-$.)

Sample Question

A point P moves along a circle $x^2 + y^2 = 4$ of radius 2 centered at $(0, 0)$. It starts at $\left(\sqrt{3}, 1\right)$ and moves clockwise with angular velocity ω. Parametrize the position $\mathbf{P}(t)$.

Solution. The point starts at a $\frac{\pi}{6} = 30°$ angle. So

$$\mathbf{P}(t) = \begin{pmatrix} 2\cos\left(\frac{\pi}{6} - \omega t\right) \\ 2\sin\left(\frac{\pi}{6} - \omega t\right) \end{pmatrix}.$$

Note that when $t = 0$ this indeed gives the starting point we originally had. □

§12.6 [RECIPE] Finding the parametrization of complicated-looking trajectories by adding two simpler ones

Since everything is so mechanical once you have an equation for $\mathbf{r}(t)$, there's a shape of exam question that comes up in 18.02 where you're given some weird-looking path and need to get its equation $\mathbf{r}(t)$ yourself in order to unlock things like velocity/speed/etc.

Something like 90%+ of the time if the shape is weird it's because it's the sum of two other vectors and you just add them. I'll write a recipe just for comedic value:

Recipe for decomposing paths as a sum of two things

Suppose P is a point following some weird trajectory. To parametrize $\mathbf{P}(t)$, one common approach is:

1. Find an expression for some other point of interest Q, say $\mathbf{Q}(t)$.
2. Find an expression for $\mathbf{v}(t)$, the vector pointing from Q to P.
3. Output $\mathbf{P}(t) = \mathbf{Q}(t) + \mathbf{v}(t)$.

We give a bunch of examples of this to follow. In this section of the notes only, if P is a point, I write $\mathbf{P}(t)$ for the corresponding parametric curve.

Tip

This section will feel repetitive. Pretty much all the examples look the same after a while. You have an amusement park ride, or a frisbee, or a planet rotating or something, or a wheel rolling some way or other… they're all thin flavor-text on the exact same thing over and over.

Okay, here are some examples.

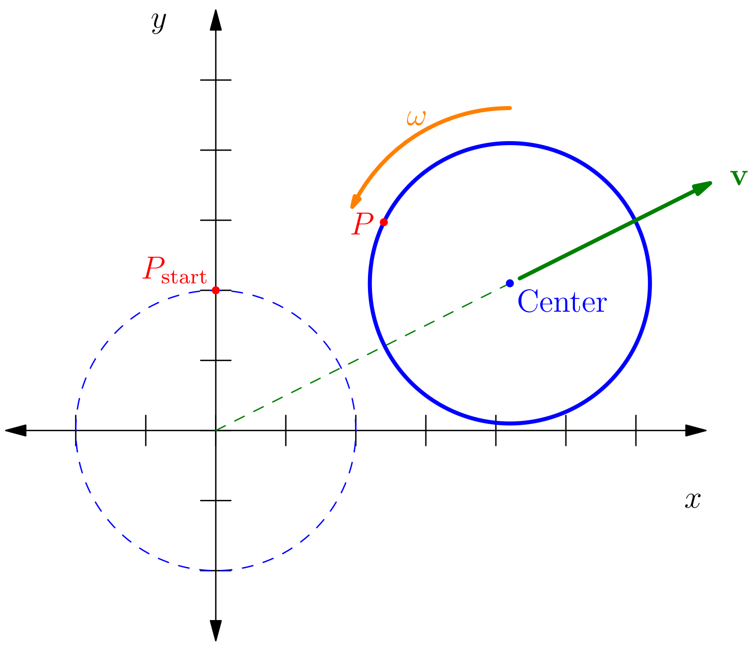

Figure 21: Spinning frisbee.

Sample Question

A frisbee has the shape of a circle of radius r, and one point P on the edge of the frisbee is marked. It's spinning in a circular motion with angular velocity ω counterclockwise and radius r about its center, while simultaneously moving in a straight line with constant velocity $\begin{pmatrix} v_x \\ v_y \end{pmatrix}$ in the plane. The frisbee initially starts at $(0, 0)$ with the marked point at $(0, r)$.

Parametrize the position of the marked point $\mathbf{P}(t)$ on the edge of the frisbee as a function of time. (See Figure 21.)

Solution. The frisbee is moving with constant velocity $\begin{pmatrix} v_x \\ v_y \end{pmatrix}$.

1. The position of the center of the frisbee as a function of time is:

$$\mathbf{O}(t) = \begin{pmatrix} v_x t \\ v_y t \end{pmatrix}.$$

   This gives the trajectory of the center of the frisbee in the plane.

2. The frisbee is also rotating about its center with angular velocity ω. The marked point on the edge of the frisbee follows a circular path around the center of the frisbee with radius r.

   Since the marked point starts at $(0, r)$ at $t = 0$, its rotational motion around the center can be described parametrically as:

$$\mathbf{v}(t) = \begin{pmatrix} r\cos\left(\frac{\pi}{2} + \omega t\right) \\ r\sin\left(\frac{\pi}{2} + \omega t\right) \end{pmatrix} = \begin{pmatrix} -r\sin(\omega t) \\ r\cos(\omega t) \end{pmatrix}.$$

   Here, ω is the angular velocity (in radians per second), and the sine and cosine terms describe the counterclockwise circular motion of the marked point around the center. (Note for $t = 0$ we get $\begin{pmatrix} 0 \\ r \end{pmatrix}$ which is what we want.)

3. To find the total position of the marked point as a function of time, we need to combine the translational motion of the frisbee's center $\mathbf{O}(t)$ with the rotational motion $\mathbf{v}(t)$. Thus, the position of the marked point at time t is the sum of the two:

$$\mathbf{P}(t) = \mathbf{O}(t) + \mathbf{v}(t).$$

   Substituting the expressions for $\mathbf{O}(t)$ and $\mathbf{v}(t)$, we get:

$$\mathbf{P}(t) = \begin{pmatrix} v_x t \\ v_y t \end{pmatrix} + \begin{pmatrix} -r\sin(\omega t) \\ r\cos(\omega t) \end{pmatrix}.$$

Simplifying, we have:

$$\mathbf{P}(t) = \begin{pmatrix} v_x t - r\sin(\omega t) \\ v_y t + r\cos(\omega t) \end{pmatrix}. \qquad \square$$

Sample Question

A planet orbits the sun in a circular path with radius R_s and *counterclockwise* angular velocity ω_s. A moon orbits the planet in a circular path with radius R_m and *clockwise* angular velocity ω_m. Parametrize the motion $\mathbf{M}(t)$ of the moon relative to the sun, assuming the sun is at the origin, the planet starts at $(R_s, 0)$, and the moon starts at $(R_s, -R_m)$. (See Figure 22.)

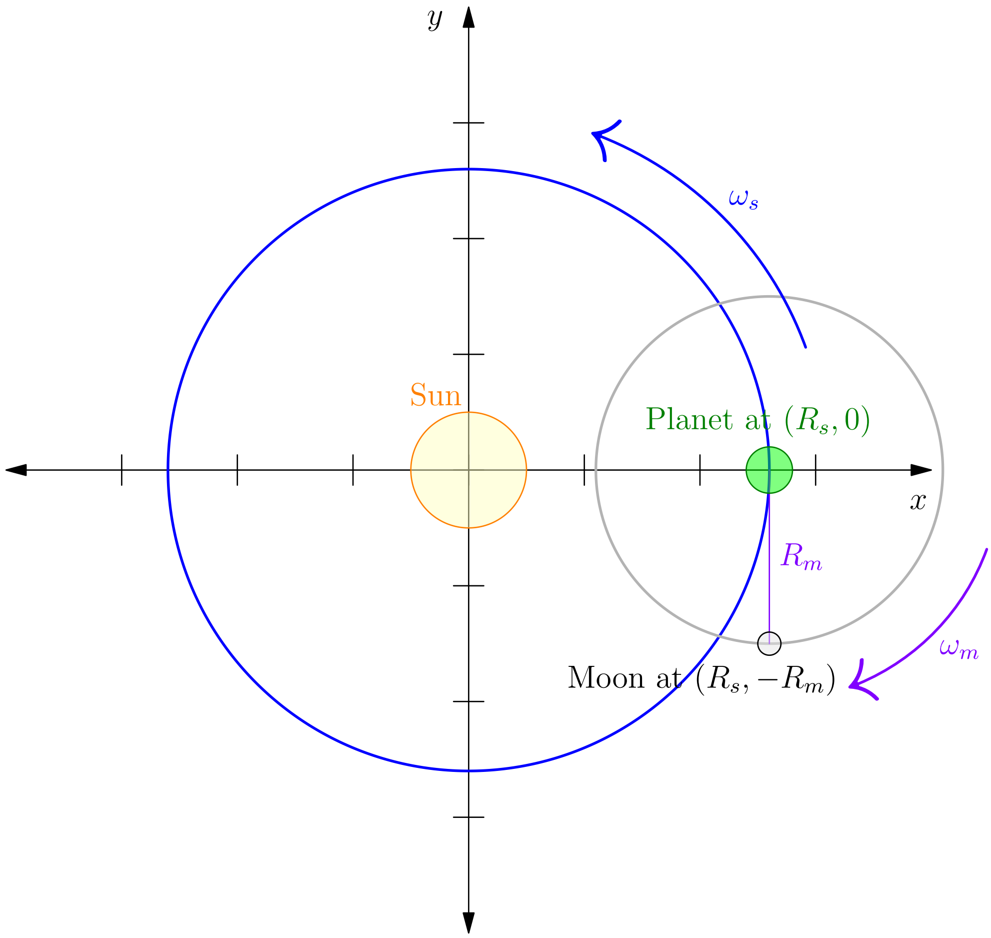

Figure 22: A planet orbits a moon in circular motion. (In real life, I think they're probably ellipses and not circles.)

Solution.

1. The planet moves in a counterclockwise orbit around the sun with radius R_s and angular velocity ω_s. The position of the planet as a function of time is:

$$\mathbf{P}(t) = \begin{pmatrix} R_s \cos(\omega_s t) \\ R_s \sin(\omega_s t) \end{pmatrix}$$

   since the planet starts due east of the sun and spins counterclockwise. This describes a counterclockwise circular motion of the planet with period $\frac{2\pi}{\omega_s}$.

2. Since the moon is orbiting the planet clockwise, the direction of its motion is reversed compared to the planet's orbit. The moon starts at $\langle 0, -R_m \rangle$ relative to the planet (due south) and moves with angular velocity ω_m.

   The position of the moon relative to the planet, moving clockwise, is given by:

$$\mathbf{v}(t) = \begin{pmatrix} R_m \cos(3\frac{\pi}{2} - \omega_m t) \\ R_m \sin(3\frac{\pi}{2} - \omega_m t) \end{pmatrix} = \begin{pmatrix} -R_m \sin(\omega_m t) \\ -R_m \cos(\omega_m t) \end{pmatrix}.$$

This describes the clockwise motion of the moon around the planet.

3. To find the total position of the moon relative to the sun, we combine the position of the planet $\mathbf{P}(t)$ and the moon's position relative to the planet $\mathbf{v}(t)$. Thus, the position of the moon relative to the sun is:

$$\mathbf{M}(t) = \mathbf{P}(t) + \mathbf{v}(t).$$

Substituting the expressions for $\mathbf{P}(t)$ and $\mathbf{v}(t)$, we get:

$$\mathbf{M}(t) = \begin{pmatrix} R_s \cos(\omega_s t) \\ R_s \sin(\omega_s t) \end{pmatrix} + \begin{pmatrix} -R_m \sin(\omega_m t) \\ -R_m \cos(\omega_m t) \end{pmatrix}.$$

Simplifying, we have:

$$\mathbf{M}(t) = \begin{pmatrix} R_s \cos(\omega_s t) - R_m \sin(\omega_m t) \\ R_s \sin(\omega_s t) - R_m \cos(\omega_m t) \end{pmatrix}.$$ □

Sample Question

A wheel of radius r starts centered at $(0, r)$ and moves in the $+x$ direction with constant speed v. Let P be a point on the rim of the wheel initially at $(0, 0)$. Parametrize the trajectory of the point $\mathbf{P}(t)$. (A picture is shown in Figure 23.)

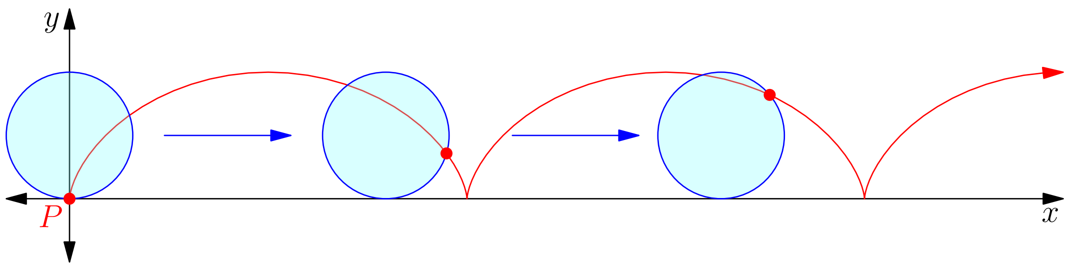

Figure 23: The cycloid formed as the wheel rolls to the right.

Solution. This problem is a little trickier because although it's easy to write the motion of the center of the wheel, it's not obvious what the angular velocity of the wheel ω should be. That will require one idea: write the length of the tire track on the ground in two ways.

1. Easy step: The wheel rolls along a straight line with constant velocity v. The position of the center of the wheel at time t is:

$$\mathbf{C}(t) = \begin{pmatrix} vt \\ r \end{pmatrix}.$$

This describes the translational motion of the center of the wheel along the horizontal axis.

2. The tricky part of the problem is determining the angular velocity of the wheel. The key idea is to look at the length of the tire track made on the ground. See Figure 24.

   - On the one hand, after time t, the length of the tire track is

$$L_{\text{tire track}} = vt$$

because the wheel covers that much distance on the ground. This is drawn in brown on Figure 24.

- On the other hand, after time t the length of the tire track should also be

$$L_{\text{tire track}} = \omega t \cdot r.$$

(It might be more natural for some of you if I write this as $\left(\frac{\omega t}{2\pi}\right) \cdot (2\pi r)$ instead, because $\frac{\omega t}{2\pi}$ is the number of full rotations made, while $2\pi r$ is the total circumference of the wheel.) This is drawn in dark blue in Figure 24.

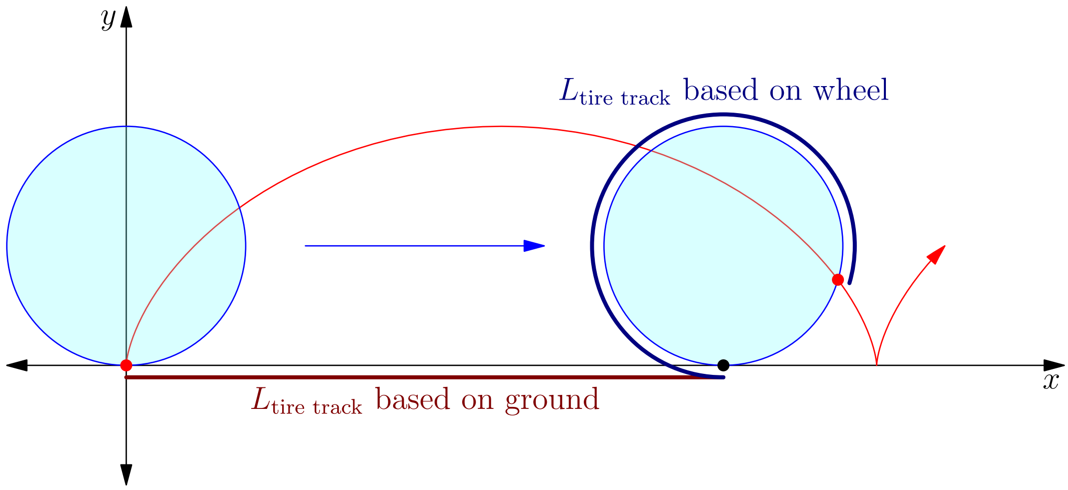

Figure 24: Calculating the length of the tire track on the ground in order to determine the angular velocity ω.

Setting $L_{\text{tire track}}$ equal in the two expressions gives

$$vt = \omega t \cdot r \implies \omega = \frac{v}{r}.$$

(We now forget about $L_{\text{tire track}}$. Its only purpose was to give us a way to get our hands on ω.)

Now that we've cleared this hurdle, the rest of the sample question is just like the earlier two. The point on the rim starts at the bottom point of the wheel at $t = 0$, due south. The rotation of the wheel is clockwise (imagine actually rolling the wheel). Therefore, the position of the point on the rim relative to the center of the wheel at time t can be parametrized as:

$$\mathbf{R}(t) = \begin{pmatrix} r\cos\left(3\frac{\pi}{4} - \omega t\right) \\ r\sin\left(3\frac{\pi}{4} - \omega t\right) \end{pmatrix} = \begin{pmatrix} -r\sin\left(\frac{v}{r}t\right) \\ -r\cos\left(\frac{v}{r}t\right) \end{pmatrix}.$$

This describes the circular motion of the point on the rim around the center of the wheel with radius r and angular velocity $\frac{v}{r}$.

3. To find the total position of the point on the rim of the wheel, we combine the translational motion of the center of the wheel $\mathbf{C}(t)$ with the rotational motion of the point on the rim $\mathbf{R}(t)$. The total position of the point on the rim at time t is:

$$\mathbf{P}(t) = \mathbf{C}(t) + \mathbf{R}(t).$$

Substituting the expressions for $\mathbf{C}(t)$ and $\mathbf{R}(t)$, we get:

$$\mathbf{P}(t) = \begin{pmatrix} vt \\ r \end{pmatrix} + \begin{pmatrix} -r\sin\left(\frac{v}{r}t\right) \\ -r\cos\left(\frac{v}{r}t\right) \end{pmatrix}.$$

Simplifying, we have:

$$\mathbf{P}(t) = \begin{pmatrix} vt - r\sin\left(\frac{v}{r}t\right) \\ r - r\cos\left(\frac{v}{r}t\right) \end{pmatrix}. \qquad \square$$

Now that you have parametric equations for each of these, you can also answer any questions solved by the methods earlier like "what is the total distance traveled" or "what is the speed at this time" or so on. Example:

Remark

The shape of $\mathbf{P}$ is called a *cycloid*, and it's shown in Figure 23. The shape looks quite scary! However, you don't actually need to know anything about the shape to compute things like the arc length (see next sample question). The geometry picture is only used to extract the algebraic expression of $\mathbf{P}(t)$. After that, you can just forget about the picture and do calculus on the expression you extracted.

Let's see this.

Sample Question

A wheel of radius 1 starts centered at $(0, 1)$ and moves in the $+x$ direction with constant speed 1. Let P be a point on the rim of the wheel initially at $(0, 0)$. Compute the total arc length of the trajectory of the point P from time $t = 0$ to $t = 2\pi$.

Solution. We just got the general equation

$$\mathbf{P}(t) = \begin{pmatrix} vt - r\sin\left(\frac{v}{r}t\right) \\ r - r\cos\left(\frac{v}{r}t\right) \end{pmatrix}$$

for a cycloid. For $v = 1$ and $r = 1$ this is

$$\mathbf{P}(t) = \begin{pmatrix} t - \sin(t) \\ r - r\cos(t) \end{pmatrix}.$$

We differentiate to get the velocity vector

$$\mathbf{P}'(t) = \begin{pmatrix} 1 - \cos(t) \\ \sin(t) \end{pmatrix}.$$

Ergo, the arc length is given by the formula

$$L = \int_0^{2\pi} \sqrt{(1 - \cos(t))^2 + \sin(t)^2}\, \mathrm{d}t.$$

This is now an 18.01 integral question. In this particular case, the square root can be simplified using trig calculation. We can expand the terms inside the square root:

$$(1 - \cos(t))^2 + \sin^2(t) = 1 - 2\cos(t) + \cos^2(t) + \sin^2(t).$$

Using the identity $\sin^2(t) + \cos^2(t) = 1$, this simplifies to:

$$1 - 2\cos(t) + 1 = 2 - 2\cos(t).$$

The trick is to use the half-angle formula to convert this to

$$1 - \cos(t) = 2\sin^2\left(\frac{t}{2}\right) \implies \sqrt{2 - 2\cos(t)} = \sqrt{4\sin^2\left(\frac{t}{2}\right)} = \left|2\sin\left(\frac{t}{2}\right)\right|.$$

Hence, the integral now becomes:

$$L = \int_0^{2\pi} \sqrt{2(1-\cos(t))}\, dt = \int_0^{2\pi} \left|2\sin\left(\frac{t}{2}\right)\right| dt.$$

Over the interval $0 \le t \le 2\pi$ we always have $\sin(\frac{t}{2}) \ge 0$, so we drop the absolute value:

$$L = \int_0^{2\pi} 2\sin\left(\frac{t}{2}\right) dt = \left[-4\cos\left(\frac{t}{2}\right)\right]_0^{2\pi} = -4\cos(\pi) + 4\cos(0) = 8.$$ □

§12.7 [TEXT] Parametrizations with flexible time

Sometimes you'll be asked to parametrize some path but not required to follow an exact time. (This happens a lot in Chapter 33, when we introduce work integrals.) In that case, you're welcome to pick any parametrization that traces out the requested path, even the start and end time. Usually the strategy is to pick one that makes subsequent calculation easier.

Sample Question

Let $\mathcal{C}$ be the line segment starting at $(0,0,0)$ and ending at $(100,200,300)$. Give any parametrization $\mathbf{r}(t)$ for $\mathcal{C}$.

Solution. The parametrization should start at $(0,0,0)$, end at $(100,200,300)$ and pass through the segment. A snapshot of some examples points on its trajectory are

$$(0,0,0) \longrightarrow (1,2,3) \longrightarrow (10,20,30) \longrightarrow (50,100,150) \longrightarrow (100,200,300)$$

among many others (like $(8\pi, 16\pi, 24\pi)$, etc.). Anyway, all the following are acceptable parametrizations:

- $\mathbf{r}(t) = (t, 2t, 3t)$ for $0 \le t \le 100$
- $\mathbf{r}(t) = (100t, 200t, 300t)$ for $0 \le t \le 1$
- $\mathbf{r}(t) = (100t^7, 200t^7, 300t^7)$ for $0 \le t \le 1$, if you enjoy making life hard for yourself.

Again, this is easiest to internalize by example: in the first one, try writing down the location of the point at $t=0$, $t=1$, $t=2$, ..., $t=100$ and verify that it's tracing out the correct thing.

In practice most people would prefer to work with the first or second one. □

Sample Question

Let $\mathcal{C}$ be the line segment starting at $(7,8,9)$ and ending at $(107,208,309)$. Give any parametrization $\mathbf{r}(t)$ for $\mathcal{C}$.

Solution. For example, the parametrization should start at $(7,8,9)$, end at $(107,208,309)$. Some examples of points along the path:

$$(7,8,9) \longrightarrow (8,10,12) \longrightarrow (17,28,39) \longrightarrow (57,108,159) \longrightarrow (107,208,309).$$

So all the following are examples of acceptable parametrizations:

- $\mathbf{r}(t) = (t+7, 2t+8, 3t+9)$ for $0 \le t \le 100$
- $\mathbf{r}(t) = (100t+7, 200t+8, 300t+9)$ for $0 \le t \le 1$
- $\mathbf{r}(t) = (100\sin(t)+7, 200\sin(t)+8, 300\sin(t)+9)$ for $0 \le t \le \frac{\pi}{2}$, if you really like trig functions.

I recommend the first one. $\square$

Sample Question

Let $\mathcal{C}$ denote the arc of the parabola $y = x^2$ starting from $(-1, 1)$ and moving right to $(1, 1)$.

Solution. Just to make things concrete, examples of points we expect to pass through in our path are

$$(-1,1) \longrightarrow \left(-\frac{1}{2}, \frac{1}{4}\right) \longrightarrow \left(-\frac{1}{3}, \frac{1}{9}\right) \longrightarrow (0,0) \longrightarrow \left(\frac{1}{3}, \frac{1}{9}\right) \longrightarrow \left(\frac{1}{2}, \frac{1}{4}\right) \longrightarrow (1,1).$$

All of the following are thus examples:

- $\mathbf{r}(t) = (t, t^2)$ for $-1 \le t \le 1$. (Yes, negative time is okay!)
- $\mathbf{r}(t) = (t-1, (t-1)^2)$ for $0 \le t \le 2$ if you're allergic to negative times.
- $\mathbf{r}(t) = (2t-1, (2t-1)^2)$ for $0 \le t \le 1$.
- $\mathbf{r}(t) = (\log(t), \log(t)^2)$ for $\frac{1}{e} \le t \le e$ if you have nothing better to do with your day.

I recommend the first one. $\square$

Sample Question

Let $\mathcal{C}$ be the path traced out by following the parabola $y = \frac{x^2}{10} + 1$ starting from $(-2, 1.4)$ and ending at $(3, 1.9)$. (See Figure 25.) Give any parametrization $\mathbf{r}(t)$ for $\mathcal{C}$.

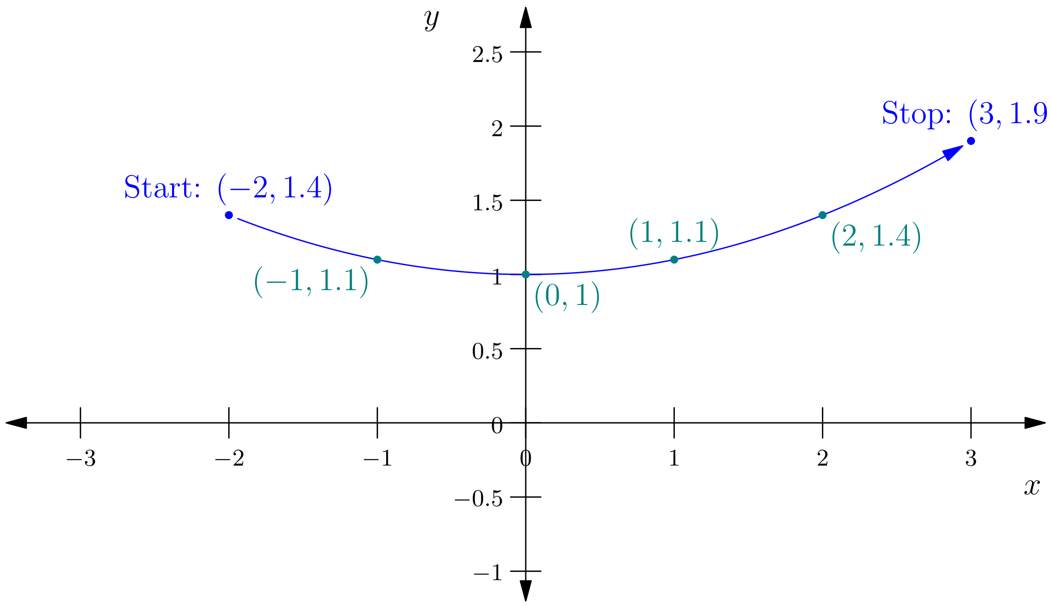

Figure 25: Walking along the parabola $y = \frac{x^2}{10} + 1$. I recommend the parametrization $\mathbf{r}(t) = \left(t, \frac{t^2}{10} + 1\right)$ for $-2 \le t \le 3$.

Solution. Examples of points passed through in this trajectory are:

$$(-3, 1.9) \longrightarrow (-2, 1.4) \longrightarrow (-1, 1.1) \longrightarrow (0, 1) \longrightarrow (1, 1.1) \longrightarrow (2, 1.4) \longrightarrow (3, 1.9).$$

In situations like this where the one coordinate just moves from one end to the other along the path, one common strategy is to just use that coordinate as t and then figure out the other coordinates from there.

All of the following are examples of acceptable parametrizations:

- $\mathbf{r}(t) = \left(t, \frac{t^2}{10} + 1\right)$ for $-2 \le t \le 3$.
- $\mathbf{r}(t) = \left(t - 2, \frac{(t-2)^2}{10} + 1\right)$ for $0 \le t \le 5$ if you're allergic to negative times.
- $\mathbf{r}(t) = \left(5t - 2, \frac{(5t-2)^2}{10} + 1\right)$ for $0 \le t \le 1$ if you really like the end time to be 1.
- $\mathbf{r}(t) = \left(5 \cdot 2^t - 7, \frac{(5 \cdot 2^t - 7)^2}{10} + 1\right)$ for $0 \le t \le 1$ if you want to torment graders.

I think most people in practice would prefer the first one. □

Sample Question

Let $\mathcal{C}$ be the $120°$ arc of the unit circle starting from $(0, -1)$ and ending at $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, going counterclockwise. (See Figure 26.) Give any parametrization $\mathbf{r}(t)$ for $\mathcal{C}$.

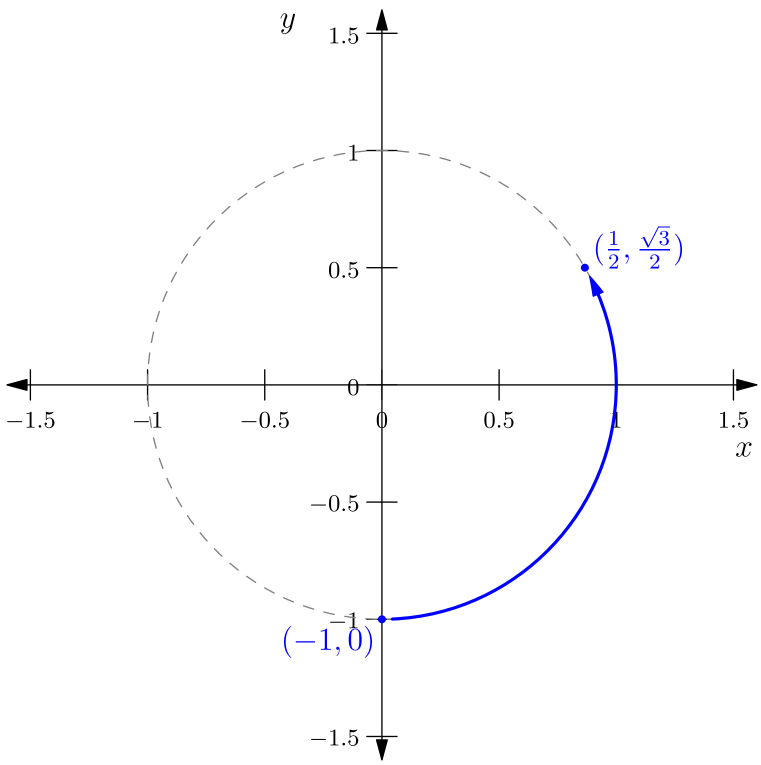

Figure 26: An $120°$ arc of the unit circle parametrized by $\mathbf{r}(t) = (\cos t, \sin t)$ for $-\frac{\pi}{2} \le t \le \frac{\pi}{6}$.

Solution. All the following are examples of acceptable parametrizations:

- $\mathbf{r}(t) = (\cos t, \sin t)$ for $-\frac{\pi}{2} \le t \le \frac{\pi}{6}$.
- $\mathbf{r}(t) = (\cos t, \sin t)$ for $\frac{3}{2}\pi \le t \le \frac{13}{6}\pi$, if you insist on using nonnegative t.
- $\mathbf{r}(t) = \left(\cos\left(t + \frac{3}{2}\pi\right), \sin\left(t + \frac{3}{2}\pi\right)\right)$ for $0 \le t \le \frac{2}{3}\pi$.
- $\mathbf{r}(t) = \left(\sqrt{1 - t^2}, t\right)$ for $-1 \le t \le \frac{\sqrt{3}}{2}$ (not recommended).

Again, I recommend the simplest (first) one. □

§12.8 [EXER] Exercises

Exercise 12.1. Compute the arc length of the part of the parabola $y = x^2 - x - 12$ between $(-3, 0)$ and $(4, 0)$.

You will probably need the following antiderivative fact not commonly seen in 18.01:

$$\int \sqrt{u^2+1}\,\mathrm{d}u = \frac{u}{2}\sqrt{u^2+1} + \frac{\log\left(u+\sqrt{u^2+1}\right)}{2} + C.$$

Exercise 12.2. At an amusement park, a teacup ride consists of teacups rotating clockwise around a fixed center while each individual teacup rotates counterclockwise. (See Figure 27 if you've never seen one of these before.) The teacup ride is specified in $\mathbb{R}^2$ as follows:

- The teacup ride revolves around $(0, 0)$ with radius R and angular velocity ω_{ride} *clockwise*.
- Each individual teacup rotates *counterclockwise* with angular velocity ω_{cup} and radius r.
- Initially, at $t = 0$, the center of the teacup is at $(R, 0)$, and a toddler is positioned at the rightmost point on the edge of the teacup relative to its center.

Compute the *velocity* vector of the toddler at time t.

Figure 27: You know, one of these teacup ride things. Image from Dreamland Amusements.

Exercise 12.3. A helicopter in $\mathbb{R}^3$ is moving upward with constant speed 5 in the $+z$ direction while its rotor blades are spinning with *clockwise* angular velocity $\frac{\pi}{3}$ and radius 2 in the horizontal plane. Let P be a point on the tip of the blade, initially at $(r, 0, 0)$.

- Parametrize the motion of a point on the tip of one of the blades as a function of time, assuming the helicopter starts at height $z = 0$ and the blade points along the positive x-axis at $t = 0$.
- Calculate the distance traveled by P from time $t = 0$ to time $t = 18$.

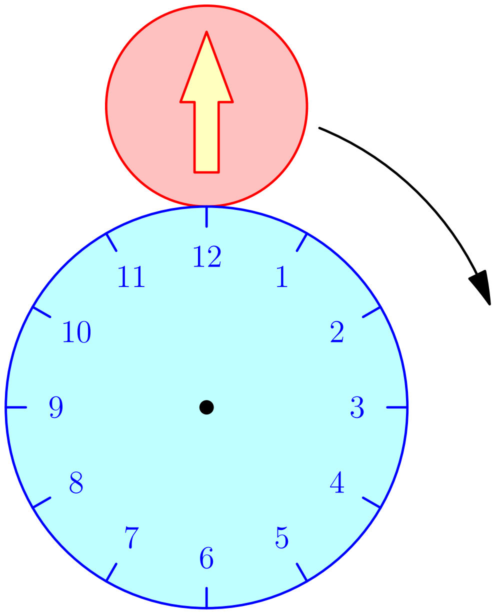

Figure 28: The clock problem from the AMC 10A in 2015.

Exercise 12.4 (*) (AMC 10A 2015)**.** In Figure 28, there's a circular clock with radius 20 cm and a circular disk of radius 10 cm externally tangent at the 12 o'clock position. The disk has an arrow painted that points directly up and rolls clockwise. At what point on the clock face will the disk be tangent when the arrow is next pointing in the upward vertical direction?

Part Echo: Multivariable differentiation

For comparison, Part Echo corresponds roughly to §8 and §12.1-§12.3 of Poonen's notes.

Chapter 13. Level curves (aka contour plots)

§13.1 [TEXT] Level curves replace xy-graphs

In high school and 18.01, you were usually taught to plot single-variable functions in two dimensions, so $f(x) = x^2$ would be drawn as a parabola $y = x^2$, and so on. You may have drilled into your head that x is an input and y is an output.

However, for 18.02 we'll typically want to draw pictures of functions like $f(x, y) = x^2 + y^2$ in a different way[14], using what's known as a *level curve*.

Definition

For any number c and function $f(x, y)$ the level curve for the value c is the plot of points for which $f(x, y) = c$.

The contrast to what you're used to is that:

- In high school and 18.01, the variables x and y play different roles, with x representing the input and $y = f(X)$ representing output.
- In 18.02, when we draw a function $f(x, y)$ both x and y are inputs; we treat them all with equal respect. Meanwhile, the *output* of the function does *not* have a variable name. If we really want to refer to it, we might sometimes write $f = 2$ as a shorthand for "the level curve for output 2".

To repeat that in table format:

18.01 xy-graphs	**18.02 level curves**
x is input	Both variables are inputs
y is output	No variable name for output

Table 5: Comparison between 18.01 xy-graphs and 18.02 level curve pictures.

We give some examples.

[14]This is a lot like how we drew planes in a symmetric form earlier. In high school algebra, you drew 2D graphs of one-variable functions like $y = 2x + 5$ or $y = x^2 + 7$. So it might have seemed a bit weird to you that we wrote planes instead like $2x + 5y + 3z = 7$ rather than, say, $z = \frac{7-2x-5y}{3}$. But this form turned out to be better, because it let us easily access the normal vector (which here is $\langle 2, 5, 3 \rangle$). The analogy carries over here.

Example: the level curves of $f(x, y) = y - x^2$

To draw the level curves of the function $f(x, y) = y - x^2$, we begin by recalling that a level curve corresponds to the points (x, y) such that the function takes on a constant value, say c. For our function, this becomes:

$$y - x^2 = c$$

which rearranges to

$$y = x^2 + c.$$

Let's talk through some values of c.

- As an example, if $c = 0$, then some points on the level curve would be $(-3, 9)$, $(-2, 4)$, $(-1, 1)$, $(0, 0)$, $(1, 1)$, $(2, 4)$, $(3, 9)$, or even $\left(\sqrt{5}, 5\right)$ and $\left(\sqrt{11}, 11\right)$. You probably already recognize what's happening: $y = x^2$ happens to be an equation you met before in 18.01 form, so you get a parabola. (More generally, if you get an equation in 18.01 form where y is a function of x, you can sketch it like you did before).
- If we change the value of $c = 2$, this equation represents a family of parabolas. For example, the level curve for 2 will be the parabola with points like $(-2, 6)$, $(-1, 3)$, $(0, 2)$, $(1, 3)$, $(2, 6)$.

In general, as c varies, the level curves are parabolas that shift upward or downward along the y -axis. The shape of these curves is determined by the quadratic term x^2, which indicates that all level curves will have the same basic "U" shape.

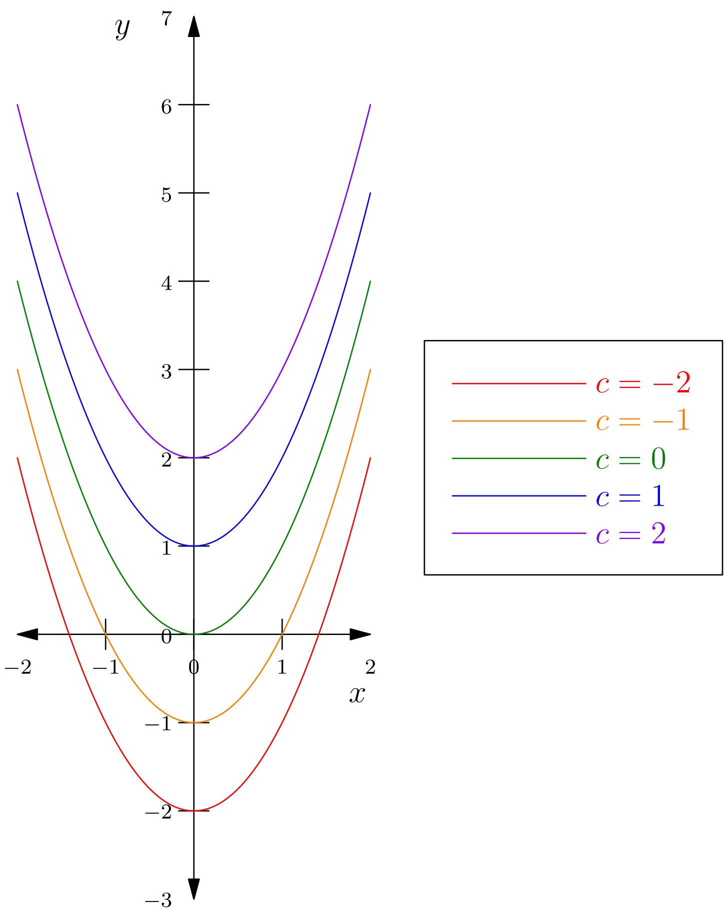

Figure 29: The level curves of $f(x, y) = y - x^2$.

Example: the level curves of $f(x, y) = x - y^2$

Let's draw level curves for $f(x, y) = x - y^2$. This example is exactly like the previous one, except the roles of x and y are flipped.

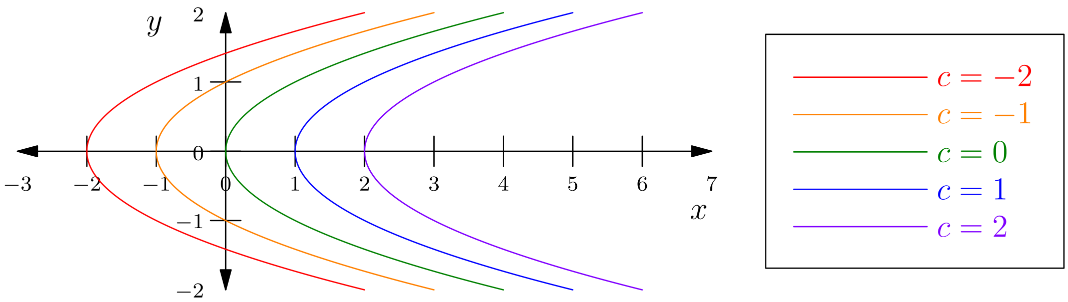

Figure 30: The level curves of $f(x, y) = x - y^2$.

Example: the level curves of $f(x, y) = x^2 + y^2$

Let's draw level curves of $f(x, y) = x^2 + y^2$ For each c we want to sketch the curve

$$x^2 + y^2 = c.$$

When $c < 0$, no points at all appear on this curve, and when $c = 0$ the only point is the origin $(0, 0)$. For $c > 0$ this equation represents a family of circles centered at the origin $(0, 0)$, with radius $\sqrt{c}$. For example:

- For $c = 1$, the level curve is a circle with radius 1.
- For $c = 4$, the level curve is a circle with radius 2.
- For $c = 9$, the level curve is a circle with radius 3.

As c increases, the circles expand outward from the origin. These concentric circles represent the level curves of the function $f(x, y) = x^2 + y^2$.

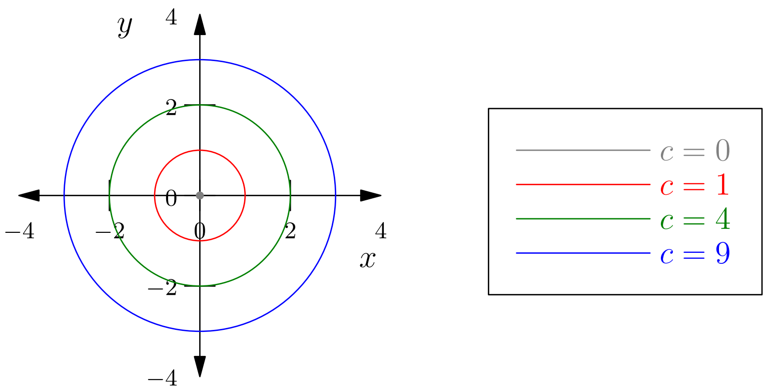

Figure 31: Four of the level curves for $f(x, y) = x^2 + y^2$.

Example: the level curves of $f(x, y) = |x| + |y|$

Let's draw level curves of $f(x, y) = |x| + |y|$. To draw the level curve for c, we are looking at

$$|x| + |y| = c.$$

Like before, if $c < 0$ there are no pairs (x, y) at all and for $c = 0$ there is only a single point.

This equation represents a family of polygons. Specifically, for a given value of c, the points that satisfy this equation form a diamond shape centered at the origin. Indeed, in the first quadrant (where the absolute values don't do anything) it represents the line segment joining $(0, c)$ to $(c, 0)$.

So for example,

- When $c < 0$, there are no points.
- For $c = 0$, the level curve is just the point $(0, 0)$.
- For $c = 1$, the level curve is a diamond with vertices at $(1, 0)$, $(-1, 0)$, $(0, 1)$, and $(0, -1)$.
- For $c = 2$, the level curve is a larger diamond with vertices at $(2, 0)$, $(-2, 0)$, $(0, 2)$, and $(0, -2)$.
- For $c = 3$, the diamond expands further, with vertices at $(3, 0)$, $(-3, 0)$, $(0, 3)$, and $(0, -3)$.

As c increases, the diamonds expand outward, maintaining their shape but increasing in size. Each level curve is a square rotated by 45 degrees.

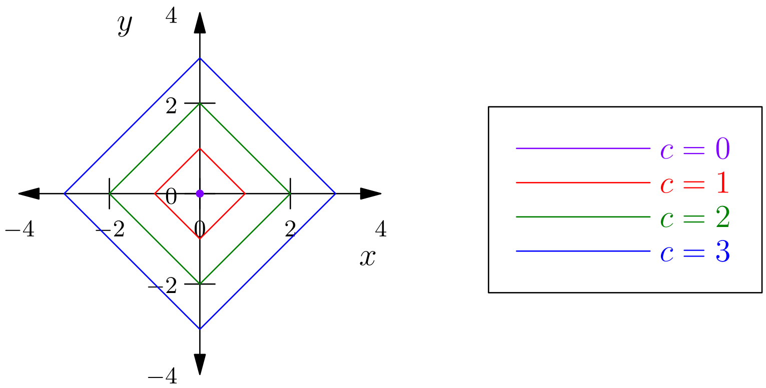

Figure 32: Four of the level curves for $f(x, y) = |x| + |y|$.

§13.2 [RECIPE] Drawing level curves

Despite the fact this chapter is labeled "recipe", there isn't an easy method that works for every function. **You have to do it in an ad-hoc way depending on the exact function you're given**. For many functions you'll see on an exam, it'll be pretty easy.

To summarize the procedure, given an explicit function like $f(x, y)$ and the value of c, one tries to plot all the points (x, y) in space with $f(x, y) = c$. We gave three examples right above, where:

- The level curves of $f(x, y) = y - x^2$ were easy to plot because for any given c, the equation just became an xy-plot like in 18.01.
- The level curves of $f(x, y) = x - y^2$ were similar to the previous example, but the roles of x and y were flipped.
- To draw the level curves of $f(x, y) = x^2 + y^2$, you needed to know that $x^2 + y^2 = r^2$ represents a circle of radius r centered at $(0, 0)$.

- To draw the level curves of $f(x,y) = |x| + |y|$, we had to think about it an ad-hoc manner where we worked in each quadrant; in Quadrant I we figured out that we got a line, and then we applied the same image to the other quadrants to get diamond shapes.

So you can see it really depends on the exact f you are given. If you wrote a really nasty function like $f(x,y) = e^{\sin xy} + \cos(x+y)$, there's probably no easy way to draw the level curve by hand.

§13.3 [TEXT] Level surfaces are exactly the same thing, with three variables instead of two

Nothing above really depends on having exactly two variables. If we had a three-variable function $f(x,y,z)$, we could draw *level surfaces* for a value of c by plotting all the points in $\mathbb{R}^3$ for which $f(x,y,z) = c$.

Example: Level surface of $f(x,y,z) = x^2 + y^2 + z^2$

If $f(x,y,z) = x^2 + y^2 + z^2$, then the level surface for the value c will be a sphere with radius $\sqrt{c}$ if $c \geq 0$. (When $c < 0$, the level surface is empty.)

Example: Level surface of $f(x,y,z) = x + 2y + 3z$

If $f(x,y,z) = x + 2y + 3z$, all the level surfaces of f are planes in $\mathbb{R}^3$, which are parallel to each other with normal vector $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

§13.4 [EXER] Exercises

Exercise 13.1. Draw 2D level curves for some values for the following functions:

- $f(x,y) = \frac{3}{2}x + y$
- $f(x,y) = xy$
- $f(x,y) = \sin(x^2 + y^2)$
- $f(x,y) = e^{y-x^2}$
- $f(x,y) = \max(x,y)$ (i.e. f outputs the larger of its two inputs, so $f(3,5) = 5$ and $f(2,-9) = 2$, for example).

Exercise 13.2 (*). Give an example of a polynomial function $f(x,y)$ for which the level curve for the value 100 consists of exactly seven points.

Chapter 14. Partial derivatives

§14.1 [TEXT] The point of differentiation is linear approximation

In 18.01, when $f : \mathbb{R} \to \mathbb{R}$, you defined a **derivative** $f'(p)$ at each input $p \in \mathbb{R}$, which you thought of as the **slope** of the **tangent line** at p. Think $f(5.01) \approx f(5) + f'(5) \cdot 0.01$. This slope roughly tells you, if you move a slight distance away from the input p, this is how fast you expect f to change. To drill the point home again, in 18.01, we had

$$f(p + \varepsilon) = f(p) + f'(p) \cdot \varepsilon.$$

See figure below.

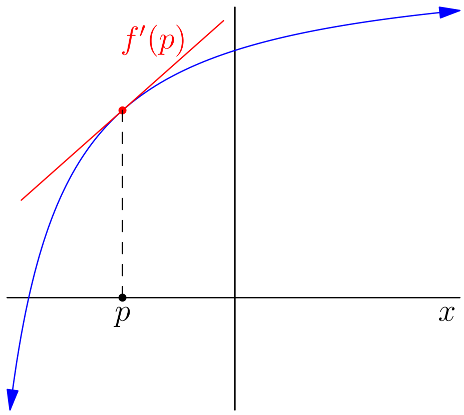

Figure 33: In 18.01, the slope $f'(p)$ tells you how quickly f changes near p.

The 18.01 derivative had type "scalar". But for a *two-variable* function, that's not enough. For concreteness, let's take

$$f(x, y) = x^2 + y^2$$

as our example function (for which we have drawn level curves before), and consider some point $P = (3, 4)$, so that $f(3, 4) = 25$.

Then, what would a point "close" to $(3, 4)$ mean? The point $(3.01, 4)$ is close, but so is $(3, 4.01)$. For that matter, so is $(3.006, 4.008)$ — that's also a point at distance 0.01 away! So having a single number isn't enough to describe the rate of change anymore.

For a two-variable function, we would really want *two* numbers, in the sense that we want to fill in the blanks in the equation

$$f(3 + \varepsilon_x, 4 + \varepsilon_y) \approx 25 + (\text{slope in } x\text{-direction}) \cdot \varepsilon_x + (\text{slope in } y\text{-direction}) \cdot \varepsilon_y.$$

> **Idea**
>
> For an n-variable functions, we have a rate of change in *each* of the n directions. Therefore, **we need n numbers and not just one**.

The first blank corresponds to what happens if you imagine y is held in place at 4, and we're just changing the x-value to 3.01. The second blank is similar. So we need a way to calculate these; the answer to our wish is what's called a *partial derivative*.

§14.2 [TEXT] Computing partial derivatives is actually just 18.01

The good news about partial derivatives is that **they're actually really easy to calculate**. You pretty much just need to do what you were taught in 18.01 with one variable changing while pretending the others are constants.

Here's the definition:

Definition

Suppose $f(x, y)$ is a two-variable function. Then the *partial derivative with respect to* x, which we denote either f_x or $\frac{\partial f}{\partial x}$, is the result if we differentiate f while treating x as a variable and y as a constant. The partial derivative $f_y = \frac{\partial f}{\partial y}$ is defined the same way.

Similarly, if $f(x, y, z)$ is a three-variable function, we write $f_x = \frac{\partial f}{\partial x}$ for the derivative when y and z are fixed.

Type signature

Each partial derivative has the same type signature as f. That is:

- Given $f : \mathbb{R}^n \to \mathbb{R}$ which accepts **points** in $\mathbb{R}^n$ and outputs **scalars**.
- Then the partial derivative $\frac{\partial f}{\partial x} = f_x$ also accepts **points** in $\mathbb{R}^n$ and outputs **scalars**.

But that's a lot of words. I think this is actually better explained by example. In fact you could probably just read the examples and ignore the definition above.

Example: partial derivatives of $f(x, y) = x^3y^2 + \cos(y)$

Let $f(x, y) = x^3y^2 + \cos(y)$.

Let's compute f_x. Again, pretend y is a constant, so look at the function

$$x \mapsto y^2 \cdot x^3 + \cos(y).$$

If we differentiate with respect to x, then x^3 becomes $3x^2$, and $\cos(y)$ goes to 0 (it doesn't have any x stuff in it). So

$$f_x = y^2 \cdot 3x^2.$$

Similarly, let's compute f_y. This time we pretend x is a constant, and look at

$$y \mapsto x^3 \cdot y^2 + \cos(y).$$

This time y^2 becomes $2y$, and $\cos(y)$ has derivative $-\sin(y)$. So

$$f_y = x^3 \cdot 2y - \sin(y).$$

Example: partial derivatives of $f(x, y, z) = e^{xyz}$

Let $f(x, y, z) = e^{xyz}$ for a three-variable example. To compute f_x, think of the function

$$x \mapsto e^{yz \cdot x}$$

where we pretend y and z are constants. Then the derivative is with respect to x is just $yze^{yz \cdot x}$ (just like how the derivative of e^{3x} is $3e^{3x}$). In other words,

$$f_x(x, y, z) = yz \cdot e^{xyz}.$$

For analogous reasons:

$$f_y(x, y, z) = xz \cdot e^{xyz}$$
$$f_z(x, y, z) = xy \cdot e^{xyz}.$$

Example: partial derivatives of $f(x, y) = x^2 + y^2$ and linear approximation

Let's go back to

$$f(x, y) = x^2 + y^2$$

which we used in our earlier example as motivation, at the point $P = (3, 4)$.

Let's fill in the numbers for the example $f(x, y) = x^2 + y^2$ we chose. By now, you should be able to compute that

$$f_x(x, y) = 2x$$
$$f_y(x, y) = 2y$$

Now, let's zoom in on just the point $P = (3, 4)$. We know that

$$f(P) = 3^2 + 4^2 = 25$$
$$f_x(P) = 2 \cdot 3 = 6$$
$$f_y(P) = 2 \cdot 4 = 8.$$

So our approximation equation can be written as

$$(3 + \varepsilon_x)^2 + (4 + \varepsilon_y)^2 \approx 25 + 6\varepsilon_x + 8\varepsilon_y. \tag{10}$$

If you manually expand both sides, you can see this looks true. The two sides differ only by ε_x^2 and ε_y^2, and the intuition is that if ε_x and ε_y were small numbers, then their squares will be negligibly small.

We'll return to Equation 10 later when we introduce the gradient.

§14.3 [RECIPE] Computing partial derivatives

You probably can already figure out the recipe from the sections above, but let's write it here just for completeness.

Recipe for calculating partial derivatives

To compute the partial derivative of a function $f(x, y)$ or $f(x, y, z)$ or $f(x_1, \ldots, x_n)$ with respect to one of its input variables,

1. Pretend all the other variables are constants, and focus on just the variable you're taking the partial derivative to.
2. Calculate the derivative of f with respect to just that variable like in 18.01.
3. Output the derivative you got.

This is easy, and only requires 18.01 material.

We just saw three examples where we computed the partials for $f(x, y) = x^3y^2 + \cos(y)$, $f(x, y, z) = e^{xyz}$, and $f(x, y) = x^2 + y^2$. Here are a bunch more examples that you can try to follow along:

Sample Question

Calculate the partial derivatives of $f(x, y, z) = x + y + z$.

Solution. The partial derivative with respect to x is obtained by differentiating

$$x \mapsto x + y + z.$$

Since we pretend y and z are constants, we just differentiate x to get 1. The same thing happens with y and z. Hence

$$\begin{aligned} f_x(x, y, z) &= 1 \\ f_y(x, y, z) &= 1 \\ f_z(x, y, z) &= 1. \end{aligned}$$

□

Sample Question

Calculate the partial derivatives of $f(x, y, z) = xy + yz + zx$.

Solution. We differentiate with respect to x first, where we view as the function

$$x \mapsto (y + z)x + yz$$

pretending that y and z are constants. This gives derivative $f_x(x, y, z) = y + z$. Similarly, $f_y(x, y, z) = x + z$ and $f_z(x, y, z) = x + y$. So

$$\begin{aligned} f_x(x, y, z) &= y + z \\ f_y(x, y, z) &= z + x \\ f_z(x, y, z) &= x + y. \end{aligned}$$

□

Sample Question

Calculate the partial derivatives of $f(x, y) = x^y$, where we assume $x, y > 0$.

Solution. If we view y as a constant and x as a variable, then

$$x \mapsto x^y$$

is differentiated by the "power rule" to get yx^{y-1}. However, if we view x as constant and y as a variable, then

$$y \mapsto x^y = e^{\log x \cdot y}$$

ends up with derivative $\log x \cdot e^{\log x \cdot y} = \log x \cdot x^y$. (Remember, in this book log denotes the *natural log*.) Hence

$$\begin{aligned} f_x(x,y) &= yx^{y-1} \\ f_y(x,y) &= \log x \cdot x^y. \end{aligned}$$

□

§14.4 [EXER] Exercises

Exercise 14.1. Compute all the partial derivatives of the following functions, defined for $x, y, z > 0$:

- $f(x,y,z) = \frac{x}{y} + \frac{y}{z} + \frac{z}{x}$
- $f(x,y,z) = \sin(xyz)$
- $f(x,y,z) = x^y + y^z + z^x$.

Chapter 15. The gradient

The gradient of $f : \mathbb{R}^n \to \mathbb{R}$, denoted ∇f, is the single most important concept in the entire "Multivariable differentiation" part. Although its definition is actually quite easy to compute, I want to give a proper explanation for where it comes from.

Throughout this chapter, remember two important ideas:

- The goal of the derivative is to approximate a function by a linear one.
- Everything you used slopes for before, you should use normal vectors instead.

If you want spoilers for what's to come, see the following table.

Thing	18.01	18.02
Input	$f : \mathbb{R} \to \mathbb{R}$	$f : \mathbb{R}^n \to \mathbb{R}$
Output	$f' : \mathbb{R} \to \mathbb{R}$	$\nabla f : \mathbb{R}^n \to \mathbb{R}^n$
Think of as	Slope (rise/run)	Measures change in *each* of n directions
Approximation	Multiply by small run	Dot product with small displacement
Picture	Slope of tangent in xy-graph	Normal vector to tangent of level curve

Table 6: How to think of ∇f for multivariable functions, compared to the derivative in 18.01.

§15.1 [TEXT] The gradient rewrites linear approximation into a dot product

In 18.01, when $f : \mathbb{R} \to \mathbb{R}$ was a function and $p \in \mathbb{R}$ was an input, we thought of the single number $f'(p)$ as the slope to interpret it geometrically. Now that we're in 18.02, we have n different rates of change, but we haven't talked about how to think of it geometrically yet.

It turns out the correct definition is to take the n numbers and make them into a vector. Bear with me for just one second:

Definition

If $f(x, y)$ is a two-variable function (so $f : \mathbb{R}^2 \to \mathbb{R}$), the **gradient** of f, denoted ∇f, is the function $\mathbb{R}^2 \to \mathbb{R}^2$ obtained by taking the two partial derivatives as the coordinates:

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix}.$$

The case of n variables is analogous; for example if $f(x, y, z)$ is a three-variable function, then

$$\nabla f(x, y, z) = \begin{pmatrix} f_x(x, y, z) \\ f_y(x, y, z) \\ f_z(x, y, z) \end{pmatrix}.$$

</> **Type signature**

The types are confusing here. To continue harping on type safety:

- Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ accepts **points** in $\mathbb{R}^2$ and outputs **scalars** in $\mathbb{R}$.
- Then $\nabla f : \mathbb{R}^2 \to \mathbb{R}^2$ accepts **points** in $\mathbb{R}^2$ and outputs **vectors** in $\mathbb{R}^2$.

Keep the distinction between points and vectors in mind when drawing pictures. We'll always draw points as dots, and vectors as arrows.

The reason for defining this gradient is that it lets us do linear approximation with a **dot product**, and consequently dot products are going to be super important throughout this chapter. Let me show you how. Let's go back to our protagonist

$$f(x, y) = x^2 + y^2$$

at the point $P = (3, 4)$. Way back in Equation 10 (on page 122), we computed $f_x(P) = 2 \cdot 3 = 6$ and $f_y(P) = 2 \cdot 4 = 8$ and used it to get the approximation

$$\begin{aligned} f(P + \langle \varepsilon_x, \varepsilon_y \rangle) &= f(\langle 3, 4 \rangle + \langle \varepsilon_x, \varepsilon_y \rangle) \\ &= (3 + \varepsilon_x)^2 + (4 + \varepsilon_y)^2 \approx 25 + 6\varepsilon_x + 8\varepsilon_y. \end{aligned}$$

Now the idea that will let us do geometry is to replace the pair of numbers ε_x and ε_y with a single "small displacement" vector $\mathbf{v} = \begin{pmatrix} \varepsilon_x \\ \varepsilon_y \end{pmatrix}$, and the pair of numbers 6 and 8 with the vector $\begin{pmatrix} 6 \\ 8 \end{pmatrix}$ instead, so that **the approximation part just becomes a dot product**:

$$f\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \mathbf{v}\right) \approx f\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right) + \begin{pmatrix} 6 \\ 8 \end{pmatrix} \cdot \mathbf{v}.$$

⚠ **Warning: the directional derivative sucks**

In some places you see the abbreviation $D_{\mathbf{v}} f(P) := \nabla f(P) \cdot \mathbf{v}$ and the name "directional derivative" for it. I hate this term, because some people have different notations and definitions (according to Wikipedia, some authors require $\mathbf{v}$ to be a unit vector, etc.).

So I will always just write the dot product $\nabla f(P) \cdot \mathbf{v}$ instead, which is unambiguous and means you have one less symbol to remember. The gradient does everything directional derivative can do, and does it better.

In full abstraction, we can rewrite linear approximation as:

❗ **Memorize: Linear approximation**

Suppose f is differentiable at a point P. Then for small displacement vectors $\mathbf{v}$, **linear approximation** promises that

$$f(P + \mathbf{v}) \approx f(P) + \nabla f(P) \cdot \mathbf{v}.$$

In other words the net change from $f(P)$ to $f(P + \mathbf{v})$ is approximated by the dot product $\nabla f(P) \cdot \mathbf{v}$.

Up until now, all we've done is rewrite the earlier equation with a different notation; so far, nothing new has been introduced. Why did we do all this work to use different symbols to say the same thing?

The important idea is what I told you a long time ago: **anything you used to think of in terms of slopes, you should rethink in terms of normal vectors**. It turns out that to complete the analogy to differentiation, the normal vector is going to be that gradient $\nabla f(P)$, and we'll see why in just a moment (spoiler: it's because of the dot product). For now, you should just know that $\nabla f(P)$ is *going to be* the right way to draw pictures of all n rates of change at once, although I haven't explained why yet.

Before going on, let's write down the recipes and some examples just to make sure the *definition* of the gradient makes sense, then I'll explain why the gradient is the normal vector we need to complete our analogy.

§15.2 [RECIPE] Calculating the gradient

Recipe for calculating the gradient

1. Compute every partial derivative of the given function.
2. Output the vector whose components are those partial derivatives.

Sample Question

Consider the six functions

$$\begin{aligned} f_1(x,y) &= x^3y^2+\cos(y), & f_2(x,y,z) &= e^{xyz} \\ f_3(x,y) &= x^2+y^2, & f_4(x,y,z) &= x+y+z \\ f_5(x,y,z) &= xy+yz+zx & f_6(x,y) &= x^y \end{aligned}$$

from back in Section 14.2 and Section 14.3. Compute their gradients.

Solution. Take the partial derivatives we already computed and make them the components:

$$\begin{aligned} \nabla f_1(x,y) &= \begin{pmatrix} 3x^2y^2 \\ 2x^3y-\sin(y) \end{pmatrix}, & \nabla f_2(x,y) &= \begin{pmatrix} yze^{xyz} \\ xze^{xyz} \\ xye^{xyz} \end{pmatrix}, \\ \nabla f_3(x,y) &= \begin{pmatrix} 2x \\ 2y \end{pmatrix}, & \nabla f_4(x,y,z) &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\ \nabla f_5(x,y,z) &= \begin{pmatrix} y+z \\ x+z \\ x+y \end{pmatrix}, & \nabla f_6(x,y) &= \begin{pmatrix} yx^{y-1} \\ \log(y)\cdot x^y \end{pmatrix}. \end{aligned}$$ □

(Remember log is the natural log, not base 10.)

§15.3 [RECIPE] Linear approximation

We actually could have stated an equivalent recipe right after we defined partial derivatives, but conceptually I think it's better to think of everything in terms of the gradient, so I waited until after I had defined the gradient to write the recipe.

Recipe for linear approximation

To do linear approximation of $f(P + \mathbf{v})$ for a small displacement vector $\mathbf{v}$:

1. Compute $\nabla f(P)$, the gradient of f at the point P.
2. Take the dot product $\nabla f(P) \cdot \mathbf{v}$ to get a number, the approximate change.
3. Output $f(P)$ plus the change from the previous step.

Sample Question

Let $f(x, y) = x^2 + y^2$. Approximate the value of $f(3.01, 4.01)$ by using linear approximation from $(3, 4)$.

Solution. Compute the gradient by taking both partial derivatives:

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

So the gradient vector at the starting point is given by

$$\nabla f(3, 4) = \begin{pmatrix} 2 \cdot 3 \\ 2 \cdot 4 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}.$$

The target point $(3.01, 4.01)$ differs from the starting point $(3, 4)$ by the displacement $\mathbf{v} = (0.01, 0.01)$. So the approximate change in f is given by

$$\underbrace{\begin{pmatrix} 6 \\ 8 \end{pmatrix}}_{=\nabla f(3,4)} \cdot \underbrace{\begin{pmatrix} 0.01 \\ 0.01 \end{pmatrix}}_{=\mathbf{v}} = (6 \cdot 0.01 + 8 \cdot 0.01) = 0.14.$$

Therefore,

$$f(3.01, 4.01) \approx \underbrace{f(3, 4)}_{=25} + 0.14 = 25.14.$$

□

Sample Question

Let $f(x, y) = x^3 - y^3$. Approximate the value of $f(2.01, -1.01)$ by using linear approximation from $(2, -1)$.

Solution. Compute the gradient by taking both partial derivatives:

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 \\ -3y^2 \end{pmatrix}.$$

So the gradient vector at the starting point $(2, -1)$ is given by

$$\nabla f(2, -1) = \begin{pmatrix} 3(2)^2 \\ -3(-1)^2 \end{pmatrix} = \begin{pmatrix} 12 \\ -3 \end{pmatrix}.$$

The target point $(2.01, -1.01)$ differs from the starting point $(2, -1)$ by the displacement $\mathbf{v} = (0.01, -0.01)$. So the approximate change in f is given by

$$\underbrace{\begin{pmatrix} 12 \\ -3 \end{pmatrix}}_{=\nabla f(2,-1)} \cdot \underbrace{\begin{pmatrix} 0.01 \\ -0.01 \end{pmatrix}}_{=\mathbf{v}} = (12 \cdot 0.01 + (-3) \cdot (-0.01)) = 0.15.$$

Therefore,

$$f(2.01, -1.01) \approx \underbrace{f(2,-1)}_{=9} + 0.15 = \boxed{9.15}.$$

□

Sample Question

Let $f(x,y) = e^x \sin(y) + 777$. Approximate the value of $f(0.04, 0.03)$ by using linear approximation from the point $(0,0)$.

Solution. Compute the gradient by taking both partial derivatives:

$$\nabla f(x,y) = \begin{pmatrix} e^x \sin y \\ e^x \cos y \end{pmatrix}.$$

So the gradient vector at the starting point $(0,0)$ is given by

$$\nabla f(0,0) = \begin{pmatrix} e^0 \sin 0 \\ e^0 \cos 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The target point $(0.04, 0.03)$ differs from the starting point $(0,0)$ by $(0.04, 0.03)$. So the approximate change in f is given by

$$\underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{=\nabla f(0,0)} \cdot \underbrace{\begin{pmatrix} 0.04 \\ 0.03 \end{pmatrix}}_{=\mathbf{v}} = 0 \cdot 0.04 + 1 \cdot 0.03 = 0.03.$$

Therefore,

$$f(0.04, 0.03) \approx \underbrace{f(0,0)}_{=777} + 0.03 = \boxed{777.03}.$$

□

§15.4 [TEXT] Gradient descent

At the end of Section 15.1, we promised the geometric definition of the dot product would pay dividends. We now make good on that promise.

The motivating question here is:

Question

Let $f(x,y) = x^2 + y^2$. Imagine we're standing at the point $P = (3,4)$. We'd like to take a step 0.01 away in some direction of our choice. For example, we could go to $(2.99, 4)$, or $(3, 4.01)$ or $(2.992, 4.006)$, or any other point on the circle we've marked in the figure below. (For the third point, note that $\sqrt{(3-2.992)^2 - (4-4.006)^2} = 0.01$, so that point is indeed 0.01 away.)

- Which way should we step if we want to maximize the f-value at the new point?
- Which way should we step if we want to the f-value to stay about the same?
- Which way should we step if we want to minimize the f-value at the new point?

You can see a cartoon of the situation in Figure 34. Note that this figure is not to scale, because 0.01 is too small to be legibly drawn, so the black circle is drawn much larger than it actually is.

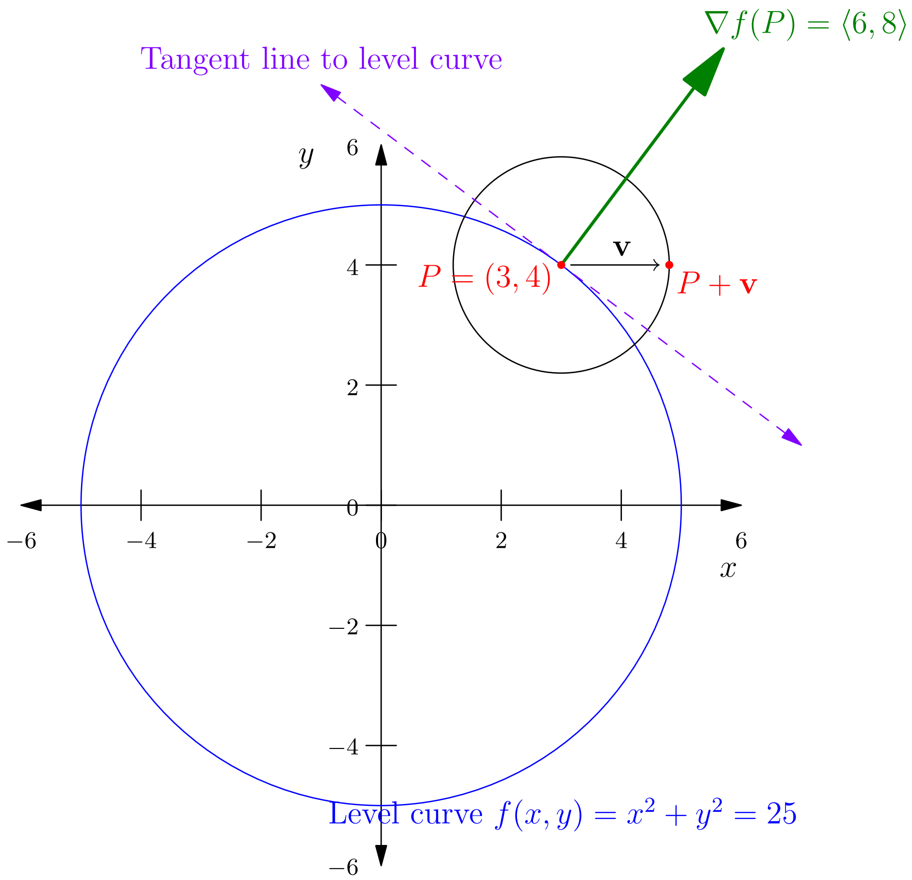

Figure 34: Starting from $P = (3,4)$, we make a step $\mathbf{v}$ away, where $|\mathbf{v}| = 0.01$. Not to scale.

To answer the question, we use the geometric interpretation of the dot product now. Remember that the change in f is approximated by

$$f(P + \mathbf{v}) - f(P) \approx \nabla f(P) \cdot \mathbf{v}.$$

The geometric definition of the dot product is that it equals

$$\nabla f(P) \cdot \mathbf{v} = |\nabla f(P)|\ |\mathbf{v}| \cos\theta$$

where θ is the included angle. But $|\nabla f(P)|$ is fixed (in this example, it's $\sqrt{6^2 + 8^2} = 10$) and $|\mathbf{v}|$ is fixed as well (in this example we chose it to be the small number 0.01).

So actually all we care about is the angle θ! Think about that for a moment. Then remember how the cosine function works:

- $\cos(0°) = 1$ is the most positive value of the cosine, and that occurs when $\mathbf{v}$ and $\nabla f(P)$ point the same direction.
- $\cos(180°) = -1$ is the most negative value of the cosine, and that occurs when $\mathbf{v}$ and $\nabla f(P)$ point opposite directions.
- If $\nabla f(P)$ and $\mathbf{v}$ are perpendicular (so $\theta = 90°$ or $\theta = 270°$), then the dot product is zero.

Translation:

Memorize

- Move **along** the gradient to increase f as quickly as possible.
- Move **against** the gradient to decrease f as quickly as possible.
- Move **perpendicular to** the gradient to avoid changing f by much either direction.

§15.5 [TEXT] Normal vectors to the tangent line/plane

We only need to add one more idea: *keeping f about the same should correspond to moving along the tangent line or plane.*

Indeed, in the 2D case, the tangent line is the line that "hugs" the level curve the closest, so we think of it as the direction causing f to avoid much change. The same is true for a tangent plane to a level surface in the 3D case; the plane hugs the curve near the point P. So that means the last bullet could be rewritten as

Memorize

The gradient $\nabla f(P)$ is normal to the tangent line/plane at P. It points towards the direction that increases f.

Example

In the previous example with a level curve, the gradient pointed away from the interior. This is not true in general. For example, imagine instead the function

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

The point $(3, 4)$ lies on the level curve of $f(3, 4) = \frac{1}{25}$. The level curve of $f(x, y)$ with value $\frac{1}{25}$ is *also* a circle of radius 5, because it corresponds to the equation $\frac{1}{x^2+y^2} = \frac{1}{25}$.

However, the gradient looks quite different: with enough calculation one gets

$$\nabla f(x, y) = \begin{pmatrix} \frac{-2x}{(x^2+y^2)^2} \\ \frac{-2y}{(x^2+y^2)^2} \end{pmatrix}.$$

Evaluating at $(3, 4)$, we get

$$\nabla f(3, 4) = \begin{pmatrix} -\frac{6}{625} \\ -\frac{8}{625} \end{pmatrix}.$$

Hence, for the function $f(x, y) = \frac{1}{x^2+y^2}$, drawing the figure analogous to Figure 34 gives something that looks quite similar, except the green arrow points the *other* way and is way smaller. This makes sense: as you move *towards* the origin, you expect $\frac{1}{x^2+y^2}$ to get larger. See Figure 35.

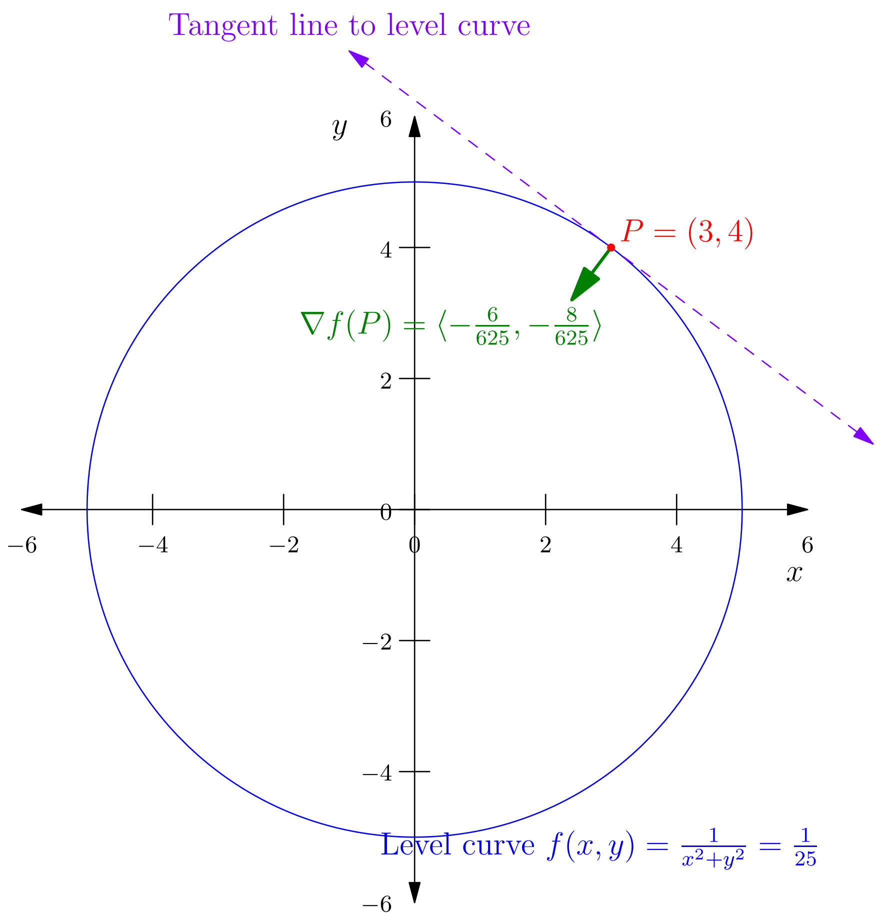

Figure 35: Similar picture but for $f(x,y) = \frac{1}{x^2+y^2}$. It looks very similar to Figure 34, but now the gradient points the other way and has much smaller absolute value, indicating that the value of f increases as we go *towards* the center (but only slightly). Not to scale.

> **Remark**
>
> Back in the 3D geometry in the linear algebra part of the course, we usually neither knew nor cared what the sign and magnitude of the normal vector was. That is, when asked "what is a normal vector to the plane $x - y + 2z = 8$?", you could answer $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ or $\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}$ or even $\begin{pmatrix} -100 \\ 100 \\ -200 \end{pmatrix}$.
>
> But this doesn't apply to the gradient anymore: while it is a normal vector to the tangent line/plane, the magnitude carries additional information we shouldn't just throw away.

§15.6 [RECIPE] Computing tangent lines/planes to level curves/surfaces

At this point, we can compute tangent lines and planes easily. We apply the old recipe in Section 5.4 (finding a plane given a point with a known normal vector) with $\nabla f(P)$ as the normal vector. To spell it out:

Recipe: Tangent line/plane to level curve/surface

To find the tangent line/plane to a level curve/surface of a function f at point P:

1. Compute the gradient ∇f. This is a normal vector, so it tells you the left-hand side for the equation of the line/plane.
2. Adjust the right-hand side so it passes through P, like in Section 5.4.

Sample Question

Compute the tangent line to $x^2 + y^2 = 25$ at the point $(3, 4)$.

Solution. Let $f(x, y) = x^2 + y^2$, so we are looking at the level curve for 25 of f. We have seen already that

$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \implies \nabla f(3,4) = \begin{pmatrix} 6 \\ 8 \end{pmatrix}.$$

Hence, the tangent line should take the form

$$6x + 8y = d$$

for some d. To pass through $P = (3, 4)$, we need $d = 6 \cdot 3 + 8 \cdot 4$, so the answer is

$$\boxed{6x + 8y = 50}.$$

□

Sample Question

Compute the tangent line to $y = x^2 + 5$ at the point $(3, 14)$.

Solution. Isn't this an 18.01 question? Yes, but the level curves work fine here to. We think of this parabola as the level curve of $f(x, y) = y - x^2$ for the value 5. The gradient is then

$$\nabla f = \begin{pmatrix} -2x \\ 1 \end{pmatrix} \implies \nabla f(3, 14) = (-6, 1).$$

Hence the tangent line should take the line $-6x + y = d$ for some d. We need to pass through $(3, 14)$, so we take $d = (-6) \cdot 3 + 14 = -4$ to get the answer

$$\boxed{-6x + y = -4}.$$

(Written in 18.01 form this would be $y = 6x - 4$, which shouldn't be a surprise, because we know the derivative of $x^2 + 5$ at $x = 3$ is 6.) □

§15.7 [RECAP] A recap of Part Echo on Multivariable Differentation

Let's summarize the last few sections.

- We replaced the old graphs we used in 18.01 with level curve and level surface pictures in Chapter 13. These new pictures differed from 18.01 pictures because all the variables on the axes are inputs now, and we treat them all with equal respect.

- We explained in Chapter 14 how to take a partial derivative of $f(x, y)$ or $f(x, y, z)$, which measures the change in just one of the variables.
- We used these partial derivatives to define the gradient ∇f in Chapter 15. This made linear approximation into a dot product, where $f(P + \mathbf{v}) \approx f(P) + \nabla f(P) \cdot \mathbf{v}$ for a small displacement $\mathbf{v}$.
- Using the geometric interpretation of a dot product, $\nabla f(P)$ was a normal vector to the level curve of f passing through P, and:
  - Going along the gradient increases f most rapidly
  - Going against the gradient decreases f most rapidly
  - Going perpendicular to the gradient puts you along the tangent line or plane at P.

§15.8 [EXER] Exercises

Exercise 15.1. Compute the equation of the tangent plane to the sphere $x^2 + y^2 + z^2 = 14$ at the point $(1, 2, 3)$.

Exercise 15.2. The level curve of a certain differentiable function $f(x, y)$ for the value -7 turns out to be a circle of radius 2 centered at $(0, 0)$.

- Give an example of one such function f.
- What are all possible vectors that $\nabla f(1.2, -1.6)$ could be?
- Do linear approximation to estimate $f(1.208, -1.594)$ starting from the point $(1.2, -1.6)$.

Exercise 15.3. For each part, either give an example of $f : \mathbb{R}^2 \to \mathbb{R}$ or show that none exist.

- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle x, y \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle 100x, y \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle y, x \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle 100y, x \rangle$?

Exercise 15.4 (*). Let a, b, c, d be nonzero real numbers and let

$$f(x, y) = ae^{x+y} + be^{x-y}.$$

Suppose the level curve of f for the value c is tangent to the line $y = 5x$ at the origin, and also passes through $(0, d)$. Compute d.

Chapter 16. Anti-gradients

This chapter is actually usually not taught until much later in 18.02, in Part India. However, I'm going to stick it in here, while your brain hasn't been infested with integrals yet, since it's a standalone question. However, if you prefer to follow 18.02 more strictly, you could skip this chapter for now and come back a month or two later once you actually need to know how to do it.

The goal of this chapter is to do ∇ backwards:

Goal

If you know ∇f, can you go back and find f?

§16.1 [TEXT] There's still $+C$ everywhere

I'll note right away that you still have the $+C$ from 18.01. To elaborate, you might remember in 18.01 that for $\int x^2 \, \mathrm{d}x = \frac{x^3}{3} + C$ for any constant C, and we usually just ignore the $+C$ because it does nothing.

For 18.02 we'll do the same thing. For example, if $f_1(x, y) = x^2 + y^2$ and $f_2(x, y) = x^2 + y^2 + 17$, then they have the same gradient:

$$\nabla f_1 = \nabla f_2 = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

However, we just agree to not care about the constant; if asked to find a potential function, and f is any acceptable answer, then so is $f + 100$. But there are will be no other answers besides $f + C$ for various C.

§16.2 [TEXT] Guessing works pretty well

Sometimes you might be able to just guess f, and if so, good for you. See if you can guess the answers to the following ones:

$$\nabla f = \begin{pmatrix} x \\ y \end{pmatrix}$$
$$\nabla f = \begin{pmatrix} ye^{xy} \\ xe^{xy} \end{pmatrix}$$
$$\nabla f = \begin{pmatrix} yz \\ zx \\ xy \end{pmatrix}.$$

§16.3 [TEXT] Antiderivative method, if you're promised there is one

How to find an anti-gradient with two variables

1. Let f denote the gradient function.
2. Integrate the given $\frac{\partial f}{\partial x}$ with respect to x to get some equation of the form $f(x, y) =$ expression $+ C_1(y)$ for some function $C_1(y)$.
3. Integrate the given $\frac{\partial f}{\partial y}$ with respect to y to get some equation of the form $f(x, y) =$ expression $+ C_2(x)$ for some function $C_2(x)$.
4. Stitch them together and output a function f.

With three variables, it's similar, but more work.

How to find an anti-gradient with three variables

1. Let f denote the gradient function.
2. Integrate the given $\frac{\partial f}{\partial x}$ with respect to x to get some equation of the form $f(x, y, z) =$ expression $+ C_1(y, z)$ for some function $C_1(y, z)$.
3. Integrate the given $\frac{\partial f}{\partial y}$ with respect to y to get some equation of the form $f(x, y, z) =$ expression $+ C_2(x, z)$ for some function $C_2(x, z)$.
4. Integrate the given $\frac{\partial f}{\partial z}$ with respect to z to get some equation of the form $f(x, y, z) =$ expression $+ C_3(x, y)$ for some function $C_3(x, y)$.
5. Stitch everything together to output f.

Let's do a two-variable example first.

Sample Question

We are given the gradient of a function $f(x, y)$:

$$\nabla f(x, y) = \begin{pmatrix} x + \cos y \\ -x \sin y \end{pmatrix}$$

Recover f.

Solution. This means:

$$\frac{\partial f}{\partial x} = x + \cos y \quad \text{and} \quad \frac{\partial f}{\partial y} = -x \sin y.$$

Integrate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ with respect to x and y:

$$f(x, y) = \int \frac{\partial f}{\partial x}\, \mathrm{d}x = \int \cos y \, \mathrm{d}x = \frac{x^2}{2} + x \cos y + C_1(y).$$
$$f(x, y) = \int \frac{\partial f}{\partial y}\, \mathrm{d}y = \int -x \sin y \, \mathrm{d}y = x \cos y + C_2(x).$$

Stitching these together to get the final expression for $f(x, y)$ as:

$$f(x, y) = \boxed{\frac{x^2}{2} + x \cos y + C}$$

for any constant C. □

Sample Question

We are given the gradient of a function $f(x, y)$:

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 + 4xy + y^2 \\ 2x^2 + 2xy - 3y^2 \end{pmatrix}$$

Recover f.

Solution. This means:

$$\frac{\partial f}{\partial x} = 3x^2 + 4xy + y^2 \quad \text{and} \quad \frac{\partial f}{\partial y} = 2x^2 + 2xy - 3y^2.$$

Integrate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ with respect to x and y:

$$f(x,y) = \int \frac{\partial f}{\partial x} \, dx = \int (3x^2 + 4xy + y^2) \, dx = x^3 + 2x^2y + xy^2 + C_1(y)$$
$$f(x,y) = \int \frac{\partial f}{\partial y} \, dy = \int (2x^2 + 2xy - 3y^2) \, dy = 2x^2 + xy^2 - y^3 + C_2(x).$$

Stitching this together gives

$$f(x,y) = \boxed{x^3 + 2x^2y + xy^2 - y^3 + C}.$$

□

Here's a three-variable version.

Sample Question

We are given the gradient of a function $f(x,y,z)$:

$$\nabla f(x,y,z) = \begin{pmatrix} y^2 - \sin(x) \\ 2xy + 4yz \\ e^z + 2y^2 \end{pmatrix}.$$

Recover f.

Solution. Again, integrate with respect to all three components:

$$f(x,y,z) = \int (y^2 - \sin x) \, dx = y^2x + \cos x + C_1(y,z)$$
$$f(x,y,z) = \int (2xy + 4yz) \, dy = xy^2 + 2zy^2 + C_2(x,z)$$
$$f(x,y,z) = \int (e^z + 2y^2) \, dy = e^z + 2y^2z + C_3(x,y).$$

Here again $C_1(y,z)$ is some function depending only on y and z; similarly for C_2 and C_3. Now stitch everything together:

$$f(x,y,z) = \boxed{y^2x + \cos x + 2y^2z + e^z + C}.$$

□

§16.4 [TEXT] Actually most of the time no potential function f exists

So far this might feel like 18.01 integration beefed up to many variables. But something is actually different. Up until now I've picked gradients for which there was an answer.

But most of the time that's not true: **the thing that's different in 18.02 is that for a randomly written question this task is *really* impossible**. That's actually a major difference.

Digression: There's a huge difference between "not easy to write the answer" and "no such function exists"

Weren't there impossible integrals in 18.01? Well, it depends on what you mean by "impossible". For example, a question you won't see in 18.01 is

$$\int \cos(x^2)\, \mathrm{d}x$$

which can be translated to "find a function f such that $f'(x) = \cos(x^2)$". The reason you're not asked this in 18.01 is because, while such a function f *does exist*, it can't be expressed in a way that makes sense to 18.01 students.

But I mean, you could always cheat and write

$$f(t) = \int_0^t \cos(x^2)\, \mathrm{d}x.$$

The right-hand side really evaluates to some number for every t, e.g. if you do numerical analysis $f(1) \approx 0.904524$, $f(2) \approx 0.461461$, etc. So there really is some function f whose derivative is $\cos(x^2)$. It just doesn't have any good way to write it down. That means, if you tried to solve the question in 18.01 methods, you just eventually run out of methods and ideas to try.

In the 18.02 version, we're about to see that even simple analogous questions might have answer "no such function exists". So when you try to solve an impossible anti-gradient question, something really different will happen: rather than running out of methods and ideas, you can follow the usual method and then *reach a contradiction.*

Question

Determine whether or not there exists a differentiable function $f(x, y)$ such that

$$\nabla f = \begin{pmatrix} 2y \\ x \end{pmatrix}.$$

Solution. We can imagine we follow through the same method as before. Integration gives

$$f(x, y) = \int 2y\, \mathrm{d}x = 2xy + C_1(y)$$
$$f(x, y) = \int x\, \mathrm{d}y = xy + C_2(x).$$

For these to be equal we need $2xy + C_1(y) = xy + C_2(x)$, so $xy + C_1(y) = C_2(x)$, which is impossible! What's going on?

(It's tempting to write $C_1(y) = C$ and $C_2(x) = xy + C$, but that's a type error. These new functions C_1 and C_2 can only depend on their arguments. Indeed, look carefully at everything we wrote for C_i previously: for example, whenever we wrote $C_1(y)$ or $C_1(y, z)$, we never allow it to depend on x.)

In mathematics there's a concept of *proof by contradiction*: if you start from an assumption, and then do some logic and reasoning to reach an impossible conclusion, then the starting assumption was wrong. Here, the starting assumption that there was *some* function f such that $\nabla f = \begin{pmatrix} 2y \\ x \end{pmatrix}$. Starting from this assumption we found that there were functions $C_1(y)$ and $C_2(x)$ such that $C_2(x) - C_1(y) = xy$ holds for all real numbers x and y. So our assumption was wrong: there can't be such function f. Not like 18.01 where "f exists but is hard to write down"; the function f literally cannot exist. □

§16.5 [TEXT] Shortcut for weeding out impossible questions

Okay, so I bet you're all wondering now, "how can I tell if the question is impossible?".

Well, one strategy would just be to **run the recipe I showed you and see if it works out**.

- If you find a function f that works, great.
- If you run into a contradiction, well, now you know it's impossible.

But that's a lot of work. We'd like a shortcut, and there is one.

The idea is that for functions (for which the partial derivatives are continuous), the *partial derivatives* commute. What that means, if say $f(x, y)$ is a two-variable function, the following is true:

! Memorize: partial derivatives commute

If f_x and f_y are both continuous then

$$f_{xy} = f_{yx}.$$

If you like ∂ notation better, this could also be written as

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}.$$

In other words, if you try to differentiate with respect to x, then with respect to y, you get the same thing as y first then x. Sometimes people write this as

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

This result is at least a bit surprising, and I actually don't expect you to believe me without seeing some examples. So let's see some examples:

Example showing that the order of differentiation doesn't matter

Let $f(x, y) = x^7 y^3$.

- If we differentiate with respect to x then to y we get $f_x = 7x^6y^3 \implies \left(\frac{\partial}{\partial y}\right) f_x = 21x^6y^2$.
- Do it the other order: $f_y = 3x^6y^2 \implies \left(\frac{\partial}{\partial x}\right) f_y = 21x^6y^2$.
- Either way we get the same result

$$f_{xy} = f_{yx} = 21x^6y^2.$$

Another example showing that the order of differentiation doesn't matter

Let $f(x,y) = \cos(x+y)y^8$.

- First, differentiate f with respect to x:

$$f_x = -\sin(x+y)y^8 \implies \frac{\partial}{\partial y} f_x = \frac{\partial}{\partial y}(-\sin(x+y)y^8)$$
$$= -\cos(x+y)y^8 - 8y^7 \sin(x+y)$$

- Do it the other order: First, differentiate f with respect to y:

$$f_y = (-\sin(x+y)y^8 + 8y^7 \cos(x+y)) \implies \frac{\partial}{\partial x} f_y = \frac{\partial}{\partial x}(-\sin(x+y)y^8 + 8y^7 \cos(x+y))$$
$$= -\cos(x+y)y^8 - 8y^7 \sin(x+y)$$

- Either way, we get the same result:

$$f_{xy} = f_{yx} = -\cos(x+y)y^8 - 8y^7 \sin(x+y)$$

Okay, so how about the example we gave earlier?

Question

Determine whether or not there exists a differentiable function $f(x,y)$ such that

$$\nabla f = \begin{pmatrix} 2y \\ x \end{pmatrix}.$$

Well, if there was such an f, and we got a mismatch, then

$$f_{xy} = \frac{\partial}{\partial y} f_x = \frac{\partial}{\partial y}(2y) = 2$$
$$f_{yx} = \frac{\partial}{\partial x} f_y = \frac{\partial}{\partial x} x = 1.$$

So via proof by contradiction, no such f could exist.

§16.6 [RECIPE] Ruling out the existence of an anti-gradient

As it turns out, this test I described is good enough for 18.02 — it will catch all impossible questions. Specifically, the following theorem is true.

Memorize: Criteria for 2D anti-gradient to exist

Consider two functions $\binom{p(x,y)}{q(x,y)}$ defined on all of $\mathbb{R}^2$, where p and q are continuously differentiable. Then there exists f such that

$$\nabla f = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$$

if and only if

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}.$$

You should think of this as $f_{xy} = f_{yx}$. We'll see this again much later in Part India, but in different language: "the 2D scalar curl of f is zero".

The 3D version tests all the possible pairs:

Memorize: Criteria for 3D anti-gradient to exist

Consider three functions $\begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}$ defined on all of $\mathbb{R}^3$, where p, q, r are continuously differentiable. Then there exists f such that

$$\nabla f = \begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}$$

if and only if all three of the following equations hold:

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}, \quad \frac{\partial p}{\partial z} = \frac{\partial r}{\partial x}, \quad \frac{\partial q}{\partial z} = \frac{\partial r}{\partial y}.$$

The above three equations should be remembered as $f_{xy} = f_{yx}$, $f_{yz} = f_{zy}$, $f_{zy} = f_{xy}$. We'll see this also in Part India again hidden in a different name: "the 3D scalar curl of f is zero". In that part, the given right-hand side will be called called a **vector field** and the function f will be called a **potential function** for it. But ignore those names for now.

§16.7 [EXER] Exercises

Exercise 16.1. Suppose $f(x,y)$ is a differentiable function and that

$$\nabla f(x,y) = \begin{pmatrix} x^2 + axy + 2y^2 + y + 1 \\ x^2 + x + bxy + y^2 + 2 \end{pmatrix}$$

for some constants a and b. Compute the constants a and b, and determine f.

Part Foxtrot: Optimization

For comparison, Part Foxtrot corresponds roughly to §9 and §12.4-§12.6 of Poonen's notes.

Chapter 17. Critical points

§17.1 [TEXT] Critical points in 18.01

First, a comparison to 18.01. Way back when you had a differentiable single-variable function $f : \mathbb{R} \to \mathbb{R}$, and you were trying to minimize it, you used the following terms:

18.01 term	Meaning
Global minimum	Minimum of the function f across the entire region you're considering
Local minimum	A point at which f is smaller than any nearby points in a small neighborhood
Critical point	A point where $f'(x) = 0$

Table 7: 18.01 terminology for critical points

Each row includes all the ones above it, but not vice-versa. Here's a picture of an example showing these for a random function $f(x) = -\frac{1}{5}x^6 - \frac{2}{7}x^5 + \frac{2}{3}x^4 + x^3$. From left to right in Figure 36, there are four critical points:

- A local maximum (that isn't a global maximum), drawn in blue.
- A local minimum (that isn't a global minimum), draw in green.
- An critical inflection point — neither a local minimum *nor* a local maximum. Drawn in orange.
- A global maximum, drawn in purple.

Note there's no global minimum at all, since the function f goes to $-\infty$ in both directions as $x \to -\infty$ or $x \to +\infty$.

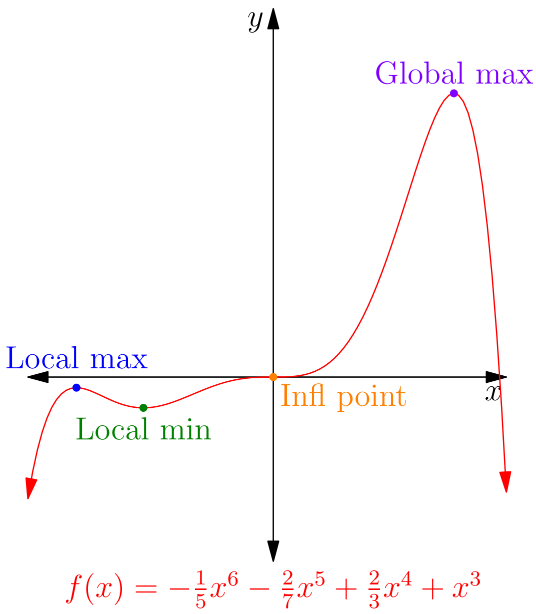

Figure 36: Some examples of critical points in an 18.01 graph of a single variable function.

§17.2 [TEXT] Critical points in 18.02

In 18.02, when we consider $f : \mathbb{R}^n \to \mathbb{R}$ the only change we make is:

Definition

For 18.02, we generalize the definition of **critical point** to be a point P for which $\nabla f(P) = \mathbf{0}$ is the zero vector. (The other two definitions don't change.)

As soon as I say this I need to carry over the analogous warnings from 18.01:

Warning

- Keep in mind that each of the implications

$$\text{Global minimum} \implies \text{Local minimum} \implies \text{Critical point, i.e. } \nabla f = \mathbf{0}$$

  is true only one way, not conversely. So a local minimum may not be a global minimum; and a point with gradient zero might not be a minimum, even locally. You should still find all the critical points, just be aware a lot of them may not actually be min's or max's.
- There may not be *any* global minimum or maximum at all, like we just saw.

Definition

In 18.02, a critical point that isn't a local minimum or maximum is called a **saddle point**.

Example

The best example of a saddle point to keep in your head is the origin for the function

$$f(x, y) = x^2 - y^2.$$

Why is this a saddle point? We have $f(0, 0) = 0$, and the gradient is zero too, since

$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \implies \nabla f(0, 0) = \begin{pmatrix} 2 \cdot 0 \\ 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The problem is that the small changes in x and y clash in sign. Specifically, if we go a little bit to either the left or right in the x-direction, then f will *increase* a little bit, e.g.

$$f(0.1, 0) = f(-0.1, 0) = 0.01 > 0.$$

But the $-y^2$ term does the opposite: if we go a little bit up or down in the y-direction, then f will *decrease* a little bit.

$$f(0, 0.1) = f(0, -0.1) = -0.01 < 0.$$

So the issue is the clashing signs of small changes in x and y directions. This causes f to neither be a local minimum nor local maximum.

There's actually nothing special about $\pm x$ and $\pm y$ in particular; I only used those to make arithmetic easier. You can see Figure 37 for values of f at other nearby points.

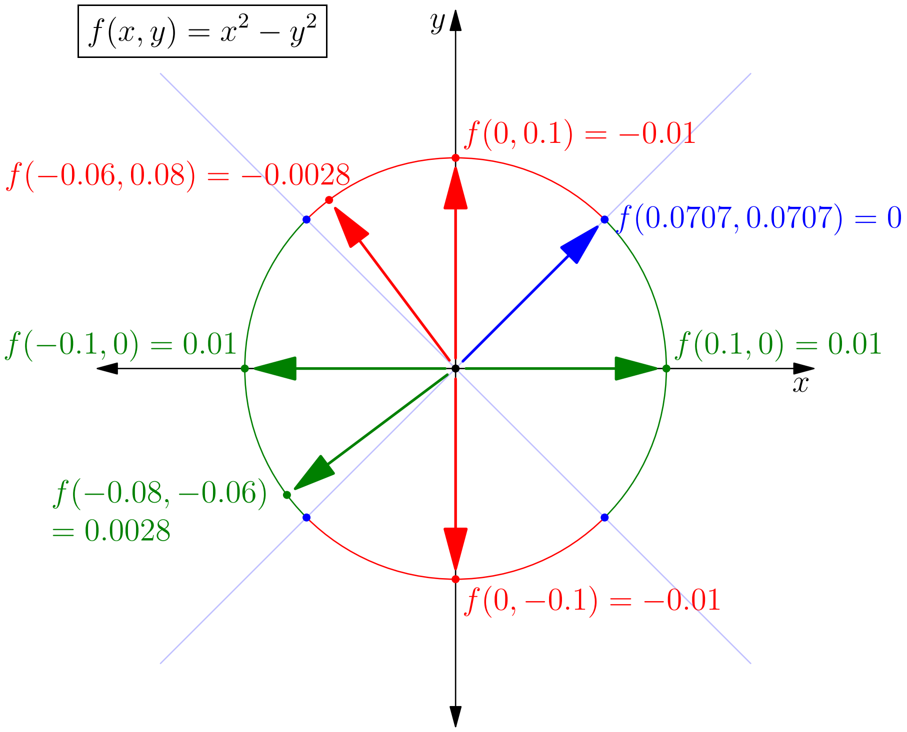

Figure 37: Values of $f(x,y) = x^2 - y^2$ at a distance of 0.1 from the saddle point $(0,0)$. Green values are positive and red ones are negative. It's a saddle point because there are both.

> **Remark**
>
> The name "saddle point" comes from the following picture: if one looks at the surface
>
> $$z = x^2 - y^2$$
>
> then near $(0,0)$ you have something that looks like a horse saddle. It curves upwards along the x-direction, but downwards along the y-direction.

We'll get to the recipe for distinguishing between saddle points and local minimums and maximums in a moment; like in 18.01, there is something called the second derivative test. First, one digression and a few examples of finding critical points.

§17.3 [SIDENOTE] Saddle points are way more common than critical inflection points

At first glance, you might be tempted to think that a saddle point is the 18.02 version of the critical inflection point. However, that analogy is actually not so good for your instincts, and **saddle points feel quite different from 18.01 critical inflection points**. Let me explain why.

In 18.01, it was *possible* for a critical point to be neither a local minimum or maximum, and we called these critical inflection points. However, in 18.01 this was actually really rare. To put this in perspective, suppose we considered a random 18.01 function of the form

$$f(x) = \square x^3 + \square x^2 + \square x + \square$$

where each square was a random integer between -1000000 and 1000000 inclusive. Of the approximately 10^{25} functions of this shape, you will find that while there are plenty of critical points, the chance of finding a critical inflection point is something like 10^{-15} — far worse than the lottery. (Of course, if you *know* where to look, you can find them: $f(x) = x^3$ has a critical inflection point at the origin, for example.)

In 18.02 this is no longer true. If we picked a random function of a similar form

$$f(x) = \square x^3 + \square x^2 + \square x + \square y^3 + \square y^2 + \square y + \square$$

where we fill each square with a number from -1000000 to 1000000 then you'll suddenly see saddle points everywhere. For example, when I ran this simulation 10000 times, among the critical points that showed up, I ended up with about

- 24.6% local minimums
- 25.3% local maximums
- 50.1% saddle points.

And the true limits (if one replaces 10^6 with N and takes the limit as $N \to \infty$) are what you would guess from the above: 25%, 25%, 50%. (If you want to see the code, it's in the Appendix, Chapter 56.)

Why is the 18.02 situation so different? It comes down to this: in 18.02, you can have two clashing directions. For the two experiments I've run here, consider the picture in Figure 38. Here P is a critical point, and we consider walking away from it in one of two directions. I'll draw a blue $+$ arrow if f increases, and a red $-$ arrow if f decreases.

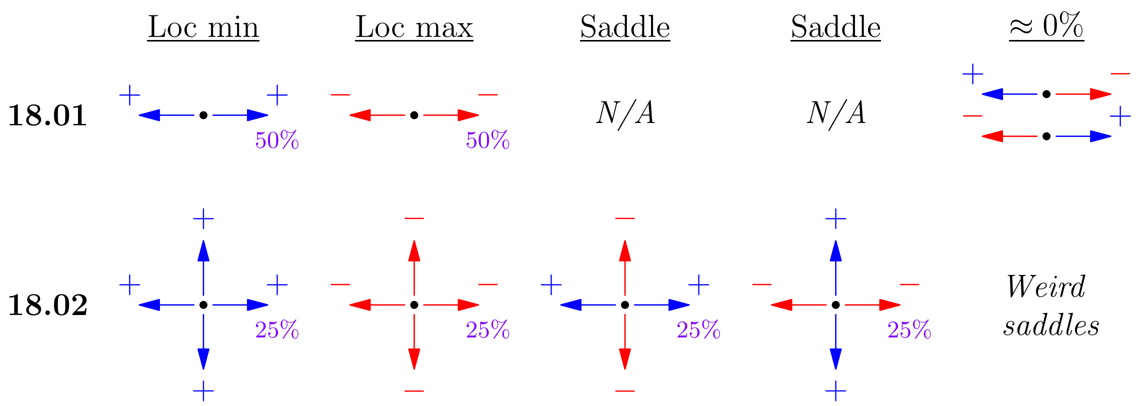

Figure 38: Why the 18.01 and 18.02 polynomial experiments have totally different outcomes.

In the 18.01 experiment, we saw that two arrows pointing *opposite* directions almost always have the same color. So in 18.01, when we could only walk in one direction, that meant almost every point was either a local minimum or a local maximum. But the picture for 18.02 is totally different because there's nothing that forces the north/south pair to have the same sign as the east/west pair. For a "random" function, if you believe the colors are equally likely, then half the time the arrows don't match colors and you end up with a saddle point.

This whole section was for two-variable functions $P(x) + Q(y)$, so it's already a simplification. If you ran an analogous three-variable experiment defined similarly for polynomials $f(x, y, z) = P(x) + Q(y) + R(z)$:

- 12.5% local minimums
- 12.5% local maximums
- 75.0% saddle points.

If we return to the world of *any* two-variable function, the truth is even more complicated than this. In this sidenote I only talked about functions $f(x, y)$ that looked like $P(x) + Q(y)$ where P and Q were polynomials. The x and y parts of the function were completely unconnected, so we only looked in the four directions north/south/east/west. But most two-variable functions have some more dependence between x and y, like $f(x,y) = x^2y^3$ or $f(x) = e^x \sin(y)$ or similar. Then you actually need to think about more directions than just north/south/east/west.

Digression

For example, Poonen's lecture notes (see question 9.22) show a weird *monkey saddle*: the point $(0,0)$ is a critical point of

$$f(x,y) = xy(x-y)$$

where the values of f nearby split into six regions, alternating negative and positive, in contrast to Figure 37 where there were only four zones on the circle. (See also Wikipedia for monkey saddle.) Poonen also invites the reader to come up with an *octopus saddle* (which sounds like it needs sixteen regions, eight down ones for each leg of the octopus).

§17.4 [RECIPE] Finding critical points

For finding critical points, on paper you can just follow the definition:

Recipe for finding critical points

To find the critical points of $f : \mathbb{R}^n \to \mathbb{R}$

1. Compute the gradient ∇f.
2. Set it equal to the zero vector and solve the resulting system of n equations in n variables.

The thing that might be tricky is that you have to solve a system of equations. Depending on how difficult your function is to work with, that might require some creativity in order to get the algebra right. We'll show some examples where the algebra is really simple, and examples where the algebra is much more involved.

Sample Question

Compute the critical points of $f(x,y,z) = x^2 + 2y^2 + 3z^2$.

Solution. The gradient is

$$\nabla f(x,y,z) = \begin{pmatrix} 2x \\ 4y \\ 6z \end{pmatrix}.$$

In order for this to equal $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, we need to solve the three-variable system of equations

$$\begin{aligned} 2x &= 0 \\ 4y &= 0 \\ 6z &= 0 \end{aligned}$$

which is so easy that it's almost insulting: $x = y = z = 0$. The only critical point is $(0,0,0)$. $\square$

Sample Question

Compute the critical points of $f(x,y) = xy(6-x-y)$.

Solution. This example is a lot more annoying than the previous one, despite having fewer variables, because casework is forced upon you. You need to solve four systems of linear equations, not just one, as you'll see.

We expand

$$f(x,y) = 6xy - x^2y - xy^2.$$

So

$$\nabla f = \begin{pmatrix} 6y - 2xy - y^2 \\ 6x - x^2 - 2xy \end{pmatrix}.$$

Hence, the resulting system of equations to solve is

$$\begin{aligned} y(6-2x-y) &= 0 \\ x(6-2y-x) &= 0. \end{aligned}$$

The bad news is that these are quadratic equations. Fortunately, they come in factored form, so we can rewrite them as OR statements:

$$\begin{aligned} y(6-2x-y) = 0 &\implies (y = 0 \quad \text{OR} \quad 2x+y=6) \\ x(6-2y-x) = 0 &\implies (x = 0 \quad \text{OR} \quad x+2y=6). \end{aligned}$$

So actually there are $2^2 = 4$ cases to consider, and we have to manually tackle all four. These cases fit into the following 2×2 table; we solve all four systems of equations.

	Top eqn. gives $y = 0$	Top eqn. gives $2x+y=6$
Bottom eqn. gives $x = 0$	$\begin{cases} y=0 \\ x=0 \end{cases} \implies (x,y) = (0,0)$	$\begin{cases} 2x+y=6 \\ x=0 \end{cases} \implies (x,y) = (0,6)$
Bottom eqn. gives $x+2y=6$	$\begin{cases} y=0 \\ x+2y=6 \end{cases} \implies (x,y) = (6,0)$	$\begin{cases} 2x+y=6 \\ x+2y=6 \end{cases} \implies (x,y) = (2,2)$

,

So we get there are four critical points, one for each case: $(0,0)$, $(0,6)$, $(6,0)$ and $(2,2)$. $\square$

§17.5 [TEXT] General advice for solving systems of equations

In the last example with $f(x,y) = xy(6-x-y)$, we saw the solving a system of equations is not necessarily an easy task. In general, solving a system of generic equations, even when the number of variables equals the number of unknowns, can be disproportionately difficult in the number of variables.

Digression: Even simple-looking systems can be challenging

It's easy to generate examples of systems that can't be solved by hand. But it's also possible to generate examples of systems that look innocent, *and* can be solved by hand in a nice way, *but* for which finding that nice way is extremely challenging.

One example of such a system of equations is

$$\begin{aligned} x^3 &= 3x - 12y + 50 \\ y^3 &= 12y + 3z - 2 \\ z^3 &= 27z + 27x. \end{aligned}$$

There is a way to solve it by hand, but it's quite hard to come up with, even for the best high school students in the world. (The source of the problem is the USA Team Selection Test 2009.)

This means that you need to put away your chef hat for a moment and put on your problem-solving cap: The good news is that it's all high-school algebra: no calculus involved, no derivatives, etc. The bad news is that it's tricky. You really have to think.

Tips on systems of equations

- When solving a system of equations, **treat it like a self-contained algebra puzzle**. That means you cannot just blindly follow a recipe, but need to actually think.
- Possible strategy in some situations: try to isolate one variable in terms of others. For example, if you see $x^2 + x + 2y = 7$, one strategy is to rewrite it as $y = \frac{1}{2}(7 - (x^2 + x))$ and then use that substitution to kill all the x's for your system. This reduces the number of variables by 1, at the cost of some work.
- If there's symmetry in the system of equations, see if you can exploit it to save work.
- Try to factor things when you spot factors. For example, if you see $xy - x = 0$, write it as $x(y - 1) = 0$, then either $x = 0$ or $y = 1$.
- If you are taking square roots of both sides, That is, if $a^2 = b^2$, you conclude $a = \pm b$, not $a = b$.
- Be careful in making sure you don't miss cases if you start getting OR statements. In the last example, there were $2^2 = 4$ cases. You can easily imagine careless students accidentally forgetting a case.
- See if you can "guess" some obvious solutions to start (e.g. all-zero). If so, note them down so you know that they should show up later.

I also need one warning: be really careful about **division by zero**. For example, in the example from last section, careless students might try to divide by y and x to get

$$\begin{aligned} 6y - 2xy - y^2 = 0 &\implies 2x + y = 6 \\ 6x - 2xy - x^2 = 0 &\implies x + 2y = 6. \end{aligned}$$

But this is wrong, because x and y could be zero too! If you make this mistake you're only getting to one of the four critical points. This is important enough I'll box it:

Warning: Watch for division by zero

Any time you divide both sides of an equation, ask yourself if you the expression you're dividing by could be 0 as well. If so, that case needs to be handled separately.

I'm going to give two examples, each with three variables, to show these ideas in the tip I just mentioned. Fair warning: these are deliberately a bit trickier, to give some space to show ideas. Don't worry if you can't do these two yourself. The exam ones will probably tone down this algebra step a bit.

Sample Question

Compute all the critical points of the function

$$f(x, y, z) = x^3 + y^3 + z^3 - 3xyz.$$

Solution. We first compute the gradient:

$$\nabla f = \begin{pmatrix} 3x^2 - 3yz \\ 3y^2 - 3zx \\ 3z^2 - 3xy \end{pmatrix}.$$

The critical points occur when $\nabla f = \mathbf{0}$, which gives us the system of equations (after dividing by 3):

$$\begin{aligned} x^2 &= yz \\ y^2 &= zx \\ z^2 &= xy. \end{aligned}$$

We'd like to divide out by the variables, but this would be division by zero. Indeed, note $(0, 0, 0)$ is a solution!

- If $x = 0$, then it follows $z = 0$ from the last equation, then $y = 0$ from the second.
- By symmetry, if *any* of the three variables is zero, then all three are.

Now let's suppose all the variables are nonzero. Then we can write the first equation safely as $z = \frac{x^2}{y}$ and use that to get rid of z in the second equation:

$$y^2 = \left(\frac{x^2}{y}\right) x \Rightarrow x^3 = y^3.$$

Similarly, we get $y^3 = z^3$ and $z^3 = x^3$.

So in fact $x = y = z$, because we can safely take cube roots of real numbers. And any triple with $x = y = z$ works fine.

In conclusion, every point of the form (t, t, t) is a critical point — an infinite family of critical points!
□

Sample Question

Compute all the critical points of the function

$$f(x, y, z) = z(x - y)(y - z) - 2xz.$$

Solution. The gradient is given by

$$\nabla f = \begin{pmatrix} z(y-z) - 2z \\ z(-2y + x + z) \\ y(x-y) - 2z(x-y) - 2x \end{pmatrix}.$$

That looks scary, but it turns out the first two equations factor. Cleaning things up, we get:

$$\begin{aligned} z(y - z - 2) &= 0 \\ z(-2y + x + z) &= 0 \\ y(x-y) - 2(x-y)z - 2x &= 0. \end{aligned}$$

In the first equation, we have cases on $z = 0$ and $y = z + 2$.

- First case: If $z = 0$, then both the first and second equation are true and give no further information. So we turn to the last equation, which for $z = 0$ says

  $$y(x - y) - 2x = 0.$$

  This is a linear equation in x that we can isolate:

  $$(y-2)x - y^2 = 0 \implies (y-2)x = y^2.$$

  Again, before dividing by $y - 2$, we check the cases:

  - If $y = 2$, we get an obvious contradiction $0 = 4$.
  - So we can assume $y \neq 2$ and $x = \frac{y^2}{y-2}$.

  Hence, for *any* real number $y \neq 2$, we get a critical point

  $$\left(\frac{y^2}{y-2}, y, 0\right).$$

- Now assume $z \neq 0$. Then we can safely divide by z in the first two equations to get

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

  Our strategy now is to write everything in terms of z. The first equation tells us $y = z + 2$, so the second equation says

  $$x = 2(z+2) + z = z + 4.$$

  We have one more equation, so we make the two substitutions everywhere and expand:

  $$\begin{aligned} 0 &= (z+2)((z+4) - (z+2)) - 2((z+4) - (z+2))z - 2(z+4) \\ &= 2(z+2) - 4z - 2(z+4) = -4z - 4 \\ &\implies z = -1. \end{aligned}$$

  Hence, we get one more critical point $(3, 1, -1)$.

In conclusion, the answer is

$$\left(\frac{y^2}{y-2}, y, 0\right) \quad \text{for every} \quad y \neq 2 \text{ , plus one extra point } (3, 1, -1). \qquad \square$$

§17.6 [RECIPE] The second derivative test for two-variable functions

Earlier we classified critical points by looking at nearby points. Technically speaking, we did not give a precise definition of "nearby", just using small numbers like 0.01 or 0.1 to make a point. So in 18.02, the exam will want a more systematic theorem for classifying critical points as local minimum, local maximum, or saddle point.

I thought for a bit about trying to explain why the second derivative test works, but ultimately I decided to not include it in these notes. Here's some excuses why:

> **Digression**
>
> The issue is that getting the "right" understanding of this would require me to talk about *quadratic forms*. However, in the prerequisite parts Alfa and Bravo of these notes, we only did linear algebra, and didn't cover quadratic forms in this context at all. I hesitate to introduce an entire chapter on quadratic forms (which are much less intuitive than linear functions) and *then* tie that to eigenvalues of a 2×2 matrix just to justify a single result not reused later. (Poonen has some hints on quadratic forms in section 9 of his notes if you want to look there though.)
>
> The other downside is that even if quadratic forms are done correctly, the second derivative test doesn't work in all cases anyway, if the changes of the function near the critical point are sub-quadratic (e.g. degree three). And multivariable Taylor series are not on-syllabus for 18.02.

So to get this section over with quickly, I'll just give the result. I'm sorry this will seem to come out of nowhere.

> **Recipe: The second derivative test**
>
> Suppose $f(x, y)$ has a critical point at P. We want to tell whether it's a local minimum, local maximum, or saddle point. Assume f has a continuous second derivative near P.
>
> 1. Let $A = f_{xx}(P)$, $B = f_{xy}(P) = f_{yx}(P)$, $C = f_{yy}(P)$. These are the partial derivatives of the partial derivatives of f (yes, I'm sorry), evaluated at P. If you prefer gradients, you could write this instead as
>
> $$\nabla f_x(P) = \begin{pmatrix} A \\ B \end{pmatrix}, \quad \nabla f_y(P) = \begin{pmatrix} B \\ C \end{pmatrix}.$$
>
> 2. If $AC - B^2 \neq 0$, output the answer based on the following chart:
>   - If $AC - B^2 < 0$, output "saddle point".
>   - If $AC - B^2 > 0$ and $A, C > 0$, output "local minimum".
>   - If $AC - B^2 > 0$ and $A, C < 0$, output "local maximum".
> 3. If $AC - B^2 = 0$, the second derivative test is inconclusive. Any of the above answers are possible, including weird/rare saddle points like the monkey saddle. You have to use a different method instead.

The quantity $AC - B^2$ is sometimes called the *Hessian determinant*; it's the determinant of the matrix $\begin{pmatrix} A & B \\ B & C \end{pmatrix}$.

Tip

It is indeed a theorem that if f is differentiable twice continuously, then $f_{xy} = f_{yx}$. That is, if you take a well-behaved function f and differentiate with respect to x then differentiate with respect to y, you get the same answer as if you differentiate with respect to y and respect to x. You'll see this in the literature written sometimes as

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} f = \frac{\partial}{\partial y}\frac{\partial}{\partial x} f.$$

Warning: This is only for two variables

The second derivative test only works for $f(x, y)$. There is no analog for $f(x, y, z)$ in this class.

Sample Question

Use the second derivative test to classify the critical point $(0, 0)$ of the function

$$f(x, y) = x^3 + x^2 + y^3 - y^2.$$

Solution. Start by computing the partial derivatives:

$$\nabla f = \begin{pmatrix} 3x^2 + 2x \\ 3y^2 - 2y \end{pmatrix} \implies \begin{cases} f_x = 3x^2 + 2x \\ f_y = 3y^2 - 2y \end{cases}.$$

We now do partial differentiation a second time on each of these. Depending on your notation, you can write this as either

$$\nabla f_x = \begin{pmatrix} 6x + 2 \\ 0 \end{pmatrix} \quad \nabla f_y = \begin{pmatrix} 0 \\ 6y - 2 \end{pmatrix}$$

or

$$f_{xx} = 6x + 2, \quad f_{xy} = f_{yx} = 0, \quad f_{yy} = 6y - 2.$$

Again, the repeated $f_{xy} = f_{yx}$ is either $\frac{\partial}{\partial y}(6x + 2) = 0$ or $\frac{\partial}{\partial x}(6y - 2) = 0$; for well-behaved functions, you always get the same answer for f_{xy} and f_{yx}.

At the origin, we get

$$\begin{aligned} A &= 6 \cdot 0 + 2 = 2 \\ B &= 0 \\ C &= 6 \cdot 0 - 2 = -2. \end{aligned}$$

Since $AC - B^2 = -4 < 0$, we output the answer "saddle point". $\square$

Sample Question

Compute the critical points of $f(x, y) = xy + y^2 + 2y$ and classify them using the second derivative test.

Solution. Start by computing the gradient:

$$\nabla f = \begin{pmatrix} y \\ x + 2y + 2 \end{pmatrix}.$$

Solve the system of equations $y = 0$ and $x + 2y + 2 = 0$ to get just $(x, y) = (-2, 0)$. Hence this is the only critical point.

We now compute the second derivatives:

$$\begin{aligned} f_{xx} &= \frac{\partial}{\partial x}(y) = 0 \\ f_{xy} = f_{yx} &= \frac{\partial}{\partial y}(y) = \frac{\partial}{\partial x}(x + 2y + 2) = 1 \\ f_{yy} &= \frac{\partial}{\partial y}(x + 2y + 2) = 2. \end{aligned}$$

These are all constant functions in this example; anyway, we have $A = 0$, $B = 1$, $C = 2$, and $AC - B^2 = -1 < 0$, so output "saddle point". $\square$

§17.7 [EXER] Exercises

Exercise 17.1. Compute the critical point(s) of $f(x, y) = x^3 + 2y^3 - 6xy$ and classify them as local minimums, local maximums, or saddle points.

Exercise 17.2. Compute the critical point(s) of $f(x, y, z) = x^2 + y^3 + z^4$ and classify them as local minimums, local maximums, or saddle points.

Exercise 17.3 (*). Does there exist a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ such that *every* point in $\mathbb{R}^2$ is a saddle point?

Exercise 17.4 (*). Give an example of a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ with the following property: every lattice point (x, y) (i.e. a point where both x and y are integers) is a saddle point, and there are no other saddle points. For example, $(2, -7)$, $(100, 100)$, and $(-42, -13)$ should be saddle points, but $\left(\frac{1}{2}, 0\right)$, $\left(\pi, -\sqrt{2}\right)$, and $\left(\sqrt{7}, \sqrt{11}\right)$ should not be.

Chapter 18. Regions

In 18.02, you'll be asked to find global minimums or maximums over a **constraint region** $\mathcal{R}$, which is only a subregion of $\mathbb{R}^n$. For example, if you have a three-variable function $f(x, y, z)$ given to you, you may be asked questions like

- What is the global maximum of f (if any) across all of $\mathbb{R}^3$?
- What is the global maximum of f (if any) across the octant[15] $x, y, z > 0$?
- What is the global maximum of f (if any) across the cube given by $-1 \leq x, y, z \leq 1$?
- What is the global maximum of f (if any) across the sphere $x^2 + y^2 + z^2 = 1$?
- ... and so on.

It turns out that thinking about constraint regions is actually half the problem. In 18.01 you usually didn't have to think much about it, because the regions you got were always intervals, and that made things easy. But in 18.02, you will need to pay much more attention.

Warning: if you are proof-capable, read the grown-up version

This entire chapter is going to be a lot of wishy-washy terms that I don't actually give definitions for. If you are a high-school student preparing for a math olympiad, or you are someone who can read proofs, read the version at `https://web.evanchen.cc/handouts/LM/LM.pdf` instead. We use open/closed sets and compactness there to do things correctly.

§18.1 [TEXT] Constraint regions

Digression: An English lesson on circle vs disk, sphere vs ball

To be careful about some words that are confused in English, I will use the following conventions:

- The word **circle** refers to a one-dimensional object with no inside, like $x^2 + y^2 = 1$. It has no area.
- The word **open disk** refers to points strictly inside a circle, like $x^2 + y^2 < 1$
- The word **closed disk** refers to a circle and all the points inside it, like $x^2 + y^2 = 1$ or $x^2 + y^2 < 1$.
- The word **disk** refers to either an open disk or a closed disk.

Similarly, a **sphere** refers only to the surface, not the volume, like $x^2 + y^2 + z^2 = 1$. Then we have **open ball**, **closed ball**, and **ball** defined in the analogous way.

In 18.02, all the constraint regions we encounter will be made out of some number (possibly zero) of equalities and inequalities. We provide several examples.

Examples of regions in $\mathbb{R}$

In $\mathbb{R}$:

- All of $\mathbb{R}$, with no further conditions.
- An open interval like $-1 < x < 1$ in $\mathbb{R}$.
- A closed interval like $-1 \leq x \leq 1$ in $\mathbb{R}$.

[15] Like "quadrant" with xy-graphs. If you've never seen this word before, ignore it.

Examples of two-dimensional regions in $\mathbb{R}^2$

In $\mathbb{R}^2$, some two-dimensional regions:
- All of $\mathbb{R}^2$, with no further conditions.
- The first quadrant $x, y > 0$, not including the axes
- The first quadrant $x, y \geq 0$, including the positive x and y axes.
- The square $-1 < x < 1$ and $-1 < y < 1$, not including the four sides of the square.
- The square $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$, including the four sides.
- The open disk $x^2 + y^2 < 1$, filled-in unit disk without its circumference.
- The closed disk $x^2 + y^2 \leq 1$, filled-in unit disk including its circumference.

Examples of one-dimensional regions in $\mathbb{R}^2$

In $\mathbb{R}^2$, some one-dimensional regions:
- The unit circle $x^2 + y^2 = 1$, which is a circle of radius 1, not filled.
- Both $x^2 + y^2 = 1$ and $x, y > 0$, a quarter-arc, not including $(1, 0)$ and $(0, 1)$.
- Both $x^2 + y^2 = 1$ and $x, y \geq 0$, a quarter-arc, including $(1, 0)$ and $(0, 1)$.
- The equation $x + y = 1$ is a line.
- Both $x + y = 1$ and $x, y > 0$: a line segment not containing the endpoints $(1, 0)$ and $(0, 1)$.
- Both $x + y = 1$ and $x, y \geq 0$: a line segment containing the endpoints $(1, 0)$ and $(0, 1)$.

I could have generated plenty more examples for $\mathbb{R}^2$, and I haven't even gotten to $\mathbb{R}^3$ yet. That's why the situation of constraint regions requires more thought in 18.02 than 18.01, (whereas in 18.01 there were pretty much only a few examples that happened).

In order to talk about the regions further, I have to introduce some new words. The three that you should care about for this class are the following: "boundary", "limit cases", and "dimension".

Warning: This is all going to be waving hands furiously

As far as I know, in 18.02 it's not possible to give precise definitions for these words. So you have to play it by ear. All the items below are rules of thumb that work okay for 18.02, but won't hold up in 18.100/18.900.

- The **boundary** is usually the points you get when you choose any one of the $\leq$ and $\geq$ constraints and turn it into and $=$ constraint. For example, the boundary of the region cut out by $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$ (which is a square of side length 2) are the four sides of the square, where either $x = \pm 1$ or $y = \pm 1$.
- The **limit cases** come in two forms:
  - If any of the variables can go to $\pm\infty$, all those cases are usually limit cases.
  - If you have any $<$ and $>$ inequalities, the cases where the variables approach those strict bounds are usually limit cases.
- The **dimension** of $\mathcal{R}$ is the hardest to define in words but easiest to guess. I'll give you two ways to guess it:
  - Geometric guess: pick a point P in $\mathcal{R}$ that's not on the boundary. Look at all the points of $\mathcal{R}$ that are close to P, i.e. a small neighborhood.

  - Say $\mathcal{R}$ is one-dimensional if the small neighborhood could be given a *length*.
  - Say $\mathcal{R}$ is two-dimensional if the small neighborhood could be given an *area*.
  - Say $\mathcal{R}$ is three-dimensional if the small neighborhood could be given a *volume*.
- Algebraic guess: the dimension of a region in $\mathbb{R}^n$ is usually equal to n minus the number of $=$ in constraints.

Overall, trust your instinct on dimension; you'll usually be right.

The table below summarizes how each constraint affects each of the three words above.

Constraint	Boundary	Limit case	Dimension
$\leq$ or $\geq$	Change to $=$ to get boundary	No effect	No effect
$<$ or $>$	No effect	Approach for limit case	No effect
$=$	No effect	No effect	Reduces dim by one

Table 8: Effects of the rules of thumb.

Let's use some examples.

Example: the circle, open disk, and closed disk

See Figure 39.

- The circle $x^2 + y^2 = 1$ is a **one-dimensional** shape. Again, we consider this region to be *one-dimensional* even though the points live in $\mathbb{R}^2$. The rule of thumb is that with 2 variables and 1 equality, the dimension should be $2 - 1 = 1$.

  Because there are no inequality constraints at all, and because x and y can't be larger than 1 in absolute value, there is no **boundary** and there are no **limit cases**.
- The open disk $x^2 + y^2 < 1$ is **two-dimensional** now, since it's something that makes sense to assign an area. (Or the rule of thumb that with 2 variables and 0 equalities, the dimension should be $2 - 0 = 2$.)

  There is one family of **limit cases: when $x^2 + y^2$ approaches 1^-**. But there is no boundary.
- The closed disk $x^2 + y^2 \leq 1$ is also **two-dimensional**. Because x and y can't be larger than 1 in absolute value, and there were no $<$ or $>$ constraints, there are no limit cases to consider. But there is a **boundary of $x^2 + y^2 = 1$**.

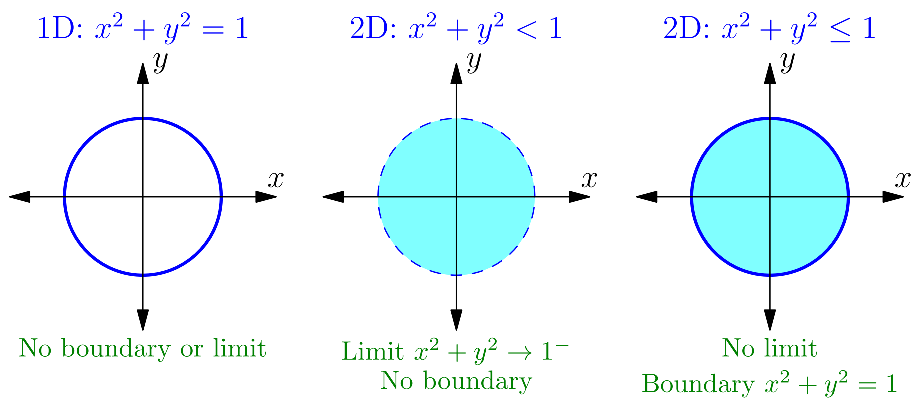

Figure 39: Pictures of $x^2 + y^2 = 1$, $x^2 + y^2 < 1$, and $x^2 + y^2 \leq 1$.

In compensation for the fact that I'm not giving you true definitions, I will instead give you a pile of examples, their dimensions, boundaries, and limit cases. See Table 9, Table 10, Table 11.

Region	Dim.	Boundary	Limit cases
All of $\mathbb{R}$	1D	No boundary	$x \to \pm\infty$
$-1 < x < 1$	1D	No boundary	$x \to \pm 1$
$-1 \le x \le 1$	1D	$x = \pm 1$	No limit cases

Table 9: Examples of regions inside $\mathbb{R}$ and their properties.

Region	Dim.	Boundary	Limit cases
All of $\mathbb{R}^2$	2D	No boundary	$x \to \pm\infty$ or $y \to \pm\infty$
$x, y > 0$	2D	No boundary	$x \to 0^+$ or $y \to 0^+$ or $x \to +\infty$ or $y \to +\infty$
$x, y \ge 0$	2D	$x = 0$ or $y = 0$	$x \to +\infty$ or $y \to +\infty$
$-1 < x < 1$ $-1 < y < 1$	2D	No boundary	$x, y \to \pm 1$
$-1 \le x \le 1$ $-1 \le y \le 1$	2D	$x = \pm 1$ or $y = \pm 1$	No limit cases
$x^2 + y^2 < 1$	2D	No boundary	$x^2 + y^2 \to 1^-$
$x^2 + y^2 \le 1$	2D	$x^2 + y^2 = 1$	No limit cases
$x^2 + y^2 = 1$	1D	No boundary	No limit cases
$x^2 + y^2 = 1$ $x, y > 0$	1D	No boundary	$x \to 0^+$ or $y \to 0^+$
$x^2 + y^2 = 1$ $x, y \ge 0$	1D	$(1, 0)$ and $(0, 1)$	No limit cases
$x + y = 1$	1D	No boundary	$x \to \pm\infty$ or $y \to \pm\infty$
$x + y = 1$ $x, y > 0$	1D	No boundary	$x \to 0^+$ or $y \to 0^+$
$x + y = 1$ $x, y \ge 0$	1D	$(1, 0)$ and $(0, 1)$	No limit cases

Table 10: Examples of regions inside $\mathbb{R}^2$ and their properties

Region	Dim.	Boundary	Limit cases
All of $\mathbb{R}^3$	3D	No boundary	Any var to $\pm\infty$
$x, y, z > 0$	3D	No boundary	Any var to 0 or ∞
$x, y, z \geq 0$	3D	$x = 0$ or $y = 0$ or $z = 0$	Any var to ∞
$x^2 + y^2 + z^2 < 1$	3D	No boundary	$x^2 + y^2 + z^2 \to 1^-$
$x^2 + y^2 + z^2 \leq 1$	3D	$x^2 + y^2 + z^2 = 1$	No limit cases
$x^2 + y^2 + z^2 = 1$	2D	No boundary	No limit cases
$x^2 + y^2 + z^2 = 1$ $x, y, z > 0$	2D	No boundary	$(1, 0)$ and $(0, 1)$
$x^2 + y^2 + z^2 = 1$ $x, y, z \geq 0$	2D	Three quarter-circle arcs[16]	No limit cases
$x + y + z = 1$	2D	No boundary	Any var to $\pm\infty$
$x + y + z = 1$ $x, y, z > 0$	2D	No boundary	Any var to 0^+
$x + y + z = 1$ $x, y, z \geq 0$	2D	$x = 0$ or $y = 0$ or $z = 0$	No limit cases

Table 11: Examples of regions inside $\mathbb{R}^3$ and their properties

Digression on intentionally misleading constraints that break the rule of thumb

I hesitate to show these, but here are some examples where the rules of thumb fail:

- An unusually cruel exam-writer might rewrite the unit circle as
$$x^2 + y^2 \leq 1 \quad \text{and} \quad x^2 + y^2 \geq 1$$
instead of the more natural $x^2 + y^2 = 1$. Then if you were blindly following the rules of thumb, you'd get the wrong answer.
- In $\mathbb{R}^3$ the region cut out by the single equation
$$x^2 + y^2 + z^2 = 0$$
is actually 0-dimensional, because there's only one point in it: $(0, 0, 0)$.

That said, intentionally misleading constraints like this are likely off-syllabus for 18.02.

§18.2 [RECIPE] Working with regions

This is going to be an unsatisfying recipe, because it's just the rules of thumb. But again, for 18.02, the rules of thumb should work on all the exam questions.

[16] To be explicit, the first quarter circle is $x^2 + y^2 = 1$, $x, y \geq 0$ and $z = 0$. The other two quarter-circle arcs are similar.

Recipe: The rule of thumb for regions defined by equations and inequalities

Given a region $\mathcal{R}$ contained in $\mathbb{R}^n$, to guess its dimension, limit cases, and boundary:

- The dimension is probably n minus the number of $=$ constraints.
- The limit cases are obtained by turning $<$ and $>$ into limits, and considering when any of the variables can go to $\pm\infty$.
- The boundary is obtained when any $\leq$ and $\geq$ becomes $=$.

See Table 8 for a summary of these rules of thumbs, and again consult Table 9, Table 10, Table 11 for a lot of examples.

Chapter 19. Optimization problems

Now that we understand both critical points of f and regions $\mathcal{R}$, we turn our attention to the problem of finding global minimums and maximums.

§19.1 [TEXT] The easy and hard cases

Suppose you have a function $f : \mathbb{R}^n \to \mathbb{R}$ that you can compute ∇f for, and a region $\mathcal{R}$. We're going to distinguish between two cases:

- The **easy case** is if $\mathcal{R}$ has dimension n as well. The rule of thumb says there should be zero "=" constraints.
- The **hard case** is if $\mathcal{R}$ has dimension $n - 1$. Rule of thumb says there should be one "=" constraint. In the hard case, we will use **Lagrange multipliers**.

We won't cover the case where $\mathcal{R}$ has dimension $n - 2$ or less in 18.02 (i.e. two or more constraints), although it can be done.

§19.2 [RECIPE] The easy case

Recipe for optimization without Lagrange Multipliers

Suppose you want to find the optimal values of $f : \mathbb{R}^n \to \mathbb{R}$ over a region $\mathcal{R}$, and $\mathcal{R}$ has dimension n.

0. Figure out the boundary and limit cases for the region $\mathcal{R}$. (You don't need to look at f for this step.)
1. Evaluate f on all the **critical points** of f in the region $\mathcal{R}$.
2. Evaluate f on all the **boundary points** of f in the region $\mathcal{R}$.
3. Evaluate f on all the **limit cases** of f in the region $\mathcal{R}$.
4. Output the points in the previous steps that give the best values, or assert the optimal value doesn't exist (if points in step 3 do better than steps 1-2).

If there are any points at which ∇f is undefined, you should check those as well. However, these seem to be pretty rare for the examples that show up in 18.02.

Warning: Don't underestimate the boundary!

In 18.01, you probably only optimized functions over intervals $I = [a, b]$, in which case the boundary was just two inputs a and b. In 18.02, the situation is completely different: the boundary (if it is nonempty) will often have *infinitely many* points. So it can take a lot of work to do the boundary case! Don't underestimate the possible complexity of Step 2.

In particular, Step 2 might even require you to use Lagrange multipliers, i.e. that one step of the easy case is an entire instance of the hard case. For that reason, the naming "easy case" and "hard case" is a bit of a misnomer.

We'll start with an example with no boundary for which the limit cases are easy to examine, so Step 2 and Step 3 are mostly harmless. Later on we'll do more examples where Step 2 and Step 3 are more intricate.

Sample Question

Compute the minimum and maximum possible value, if they exist of

$$f(x, y) = x + y + \frac{8}{xy}$$

over $x, y > 0$.

Solution. The region $\mathcal{R}$ is the first quadrant which is indeed two-dimensional (no $=$ constraints), so we're in the easy case and the recipe applies here. We check all the points in turn:

0. $\mathcal{R}$ has no boundary and limit cases when any variable approaches 0 or $+\infty$.
1. To find the critical points, calculate the gradient

$$\nabla f(x, y) = \begin{pmatrix} 1 - \frac{8}{x^2 y} \\ 1 - \frac{8}{xy^2} \end{pmatrix}$$

and then set it equal to $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. This gives us the simultaneous equations

$$1 = \frac{8}{x^2 y} = \frac{8}{xy^2}.$$

This implies $x^2 y = xy^2$ or $x = y$ (we have $x, y > 0$ in $\mathcal{R}$, so we're not worried about division by zero) and so the only critical point is $(x, y) = (2, 2)$.

2. The region $\mathcal{R}$ has no boundary, so there are no boundary points to check.
3. The region $\mathcal{R}$ has four different kinds of limit cases:
   - $x \to 0^+$
   - $y \to 0^+$
   - $x \to +\infty$
   - $y \to +\infty$.

   In fact all four of these cases cause $f \to +\infty$. In each of the first two cases, the term $\frac{8}{xy}$ in f causes $f \to +\infty$. In the case $x \to \infty$, the term x causes $f \to +\infty$. In the case $y \to \infty$, the term y causes $f \to +\infty$.

Putting these together:

- The global minimum is $(2, 2)$, at which $f(2, 2) = 6$.
- There is no global maximum, since we saw limit cases where $f \to +\infty$. □

§19.3 [TEXT] Lagrange multipliers

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function we're optimizing over some region $\mathcal{R}$. We now turn to the case where $\mathcal{R}$, is dimension $n - 1$, because of a single constraint of the form $g(x, y) = c$ or $g(x, y, z) = c$.

We need a new definition of critical point. To motivate it, let's consider a particular example in Figure 40. Here $n = 2$, and

- $f(x, y) = x^2 + y^2$, and
- $g(x, y) = c$ is the red level curve shown in the picture below;
- $\mathcal{R}$ is just the level curve $g(x, y) = c$ (no further $<$ or $\leq$ constraints).

Trying to optimize f subject to $g(x, y) = c$ in this picture is the same as finding the points on the level curve which are furthest or closest to the origin. I've marked those two points as P and Q in the figure. The trick to understanding how to get them is to *also* draw the level curves for f that pass through P and Q: then we observe that the level curves for f that get those minimums and maximums ought to be tangent to $g(x, y) = c$ at P and Q.

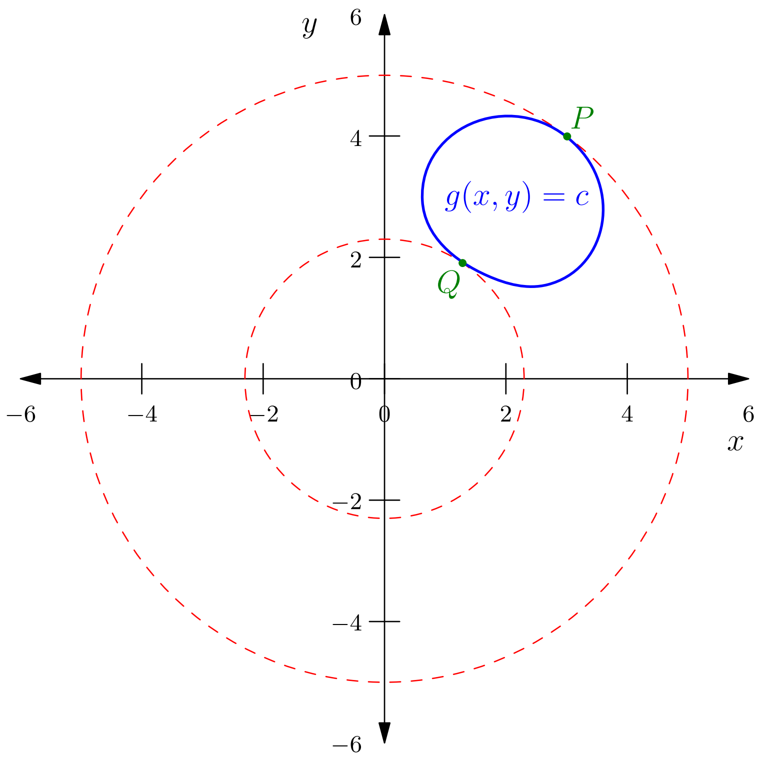

Figure 40: An example of a LM-type optimization problem, where one finds points on $g(x, y) = c$ which optimize f

Now how can we check whether there's a tangency? Answer: look at the gradient! We expect that ∇f and ∇g, at the points P and Q, should point in the same direction. So that gives us the strategy: look for the points where ∇f and ∇g point the same way.

I don't think the following term is an official name, but I like it, and I'll use it:

Definition

An **LM-critical point** is a point P on the curve/surface $g(P) = c$ such that *either*

- $\nabla f(P) = \lambda \nabla g(P)$ for some scalar λ; OR
- $\nabla g(P) = \mathbf{0}$.

Note that there are *two* possible situations. If you want, you can think about this as requiring that $\nabla f(P)$ and $\nabla g(P)$ are *linearly dependent*, so it's only one item. However, in practice, people end up usually breaking into cases like this.

Digression

The parameter λ is the reason for the name "Lagrange multipliers"; it's a scalar multiplier on ∇g. Personally, I don't think this name makes much sense.

Also, some sources will define an LM-critical point in the following (in my opinion, more confusing) way. Let's say $f(x, y, z)$ is a three-variable function. Define a *four-variable* "Lagrangian function" $L(x, y, z, \lambda) = f(x, y, z) - \lambda g(x, y, z)$. Then an LM-critical point is a point for which either $\nabla g = \mathbf{0}$ or $\nabla L = \mathbf{0}$, i.e. a normal critical point of L. It can be checked this is equivalent to the original definition, but I personally find this unnatural. However, if you like this definition, feel free to use it instead.

Now that we have this, we can describe the recipe for the "hard" case. The only change is to replace the old critical point definition (where $\nabla f(P) = \mathbf{0}$) with the LM-critical point definition.

§19.4 [RECIPE] Lagrange multipliers

Recipe for Lagrange multipliers

Suppose you want to find the optimal values of $f : \mathbb{R}^n \to \mathbb{R}$ over a region $\mathcal{R}$, and $\mathcal{R}$ has dimension $n - 1$ due to a single constraint $g = c$ for some $g : \mathbb{R}^n \to \mathbb{R}$.

0. Figure out the boundary and limit cases for the region $\mathcal{R}$. (You don't need to look at f for this step.)
1. Evaluate f on all the **LM-critical points** of f that lie on the region $\mathcal{R}$.
2. Evaluate f on all the **boundary points** of f of the region $\mathcal{R}$.
3. Evaluate f on all the **limit cases** of f of the region $\mathcal{R}$.
4. Output the points in the previous steps that give the best values, or assert the optimal value doesn't exist (if points in step 3 do better than steps 1-2).

If there are any points at which ∇f or ∇g are undefined, you should check those as well. However, these seem to be pretty rare for the examples that show up in 18.02.

Again, this is the same recipe as Section 19.2, except we changed "critical point" to "LM-critical point".

Tip

Remember how finding critical points could lead to systems of equations that required quite a bit of algebraic skill to solve? The same is true for Lagrange multipliers, but even more so, because of the new parameter λ that you have to care about. So the reason this is called the "hard case" isn't because the 18.02 ideas needed are different, but because the algebra can become quite involved in finding LM-critical points.

In fact, in high school math competitions, the algebra can sometimes become so ugly that the method of Lagrange multipliers is sometimes jokingly called "Lagrange *murderpliers*" to reflect the extreme amount of calculation needed for some problems.

Sample Question

Compute the minimum and maximum possible value, if they exist, of

$$f(x,y,z) = x + y + z$$

over $x, y, z > 0$ satisfying the condition $xyz = 8$.

Solution. We carry out the recipe.

0. The region $\mathcal{R}$ is two-dimensional, consisting of strict inequalities $x, y, z > 0$ and the condition $g(x,y,z) = xyz = 8$. So there is no boundary, but there are limit cases if any variables approaches 0 or $+\infty$.
1. To find the LM-critical points, we need to compute both ∇f and ∇g. We do so:

$$\nabla f(x,y,z) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
$$\nabla g = (yz, zx, xy).$$

Now, there are no points with $\nabla g = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ in the region $\mathcal{R}$, because in $\mathcal{R}$ all the variables are constrained to be positive. So we now solve the system

$$\begin{aligned} 1 &= \lambda yz \\ 1 &= \lambda zx \\ 1 &= \lambda xy \end{aligned}$$

and see what values it takes.

The trick to solving the system of equations is to divide the first two to get rid of the parameter λ, which we don't really care about, to get

$$\frac{1}{1} = \frac{\lambda yz}{\lambda zx} = \frac{y}{x}.$$

So we must have $x = y$. Similarly, we find $y = z$ and $z = x$.

Hence the LM-critical point must have $x = y = z$. Since $xyz = 8$, it follows the only LM-critical point is $(2,2,2)$. Evaluating f here gives $f(2,2,2) = 6$.

2. The region $\mathcal{R}$ has no boundary, because no $\le$ or $\ge$ constraints are present.
3. The region $\mathcal{R}$ has limit cases when any of the variables x, y, z either approach 0 or $+\infty$. However, remember that $xyz = 8$. So if any variable approaches 0, some other variable must become large. Consequently, in every limit case, we find that $f \to +\infty$.

Collating all these results:

- The unique global minimum is $(2,2,2)$ at which $f(2,2,2) = 6$.
- There is no global maximum. $\square$

Remark

If you're paying close enough attention, you might realize this sample question we just did is a thin rewriting of the example in Section 19.2. This illustrates something: sometimes you can rewrite a hard-case optimization problem in 3 variables to an easy-case one with 2 variables.

§19.5 [TEXT] Advice for solving systems of equations (reprise)

Back in Section 17.5, when we were finding all the critical points, I reminded you to be careful about division-by-zero, handle cases carefully, and gave you some advice for solving systems. Solving the system for *LM-critical* points is a similar situation: systems of equations are hard, and you have to treat it like a self-contained high school algebra puzzle.

For systems of equations generated by Lagrange multipliers, there's a new feature: a variable λ whose value is never used, but which appears in every equation besides the constraint $g = c$. So a couple tips specific to Lagrange Multiplier systems:

Tip: Advice for Lagrange Multiplier systems

In addition to the tips in Section 17.5, here are some strategies that sometimes help:

- It might make sense to try to get rid of λ ASAP, if that's easy to do. After all, we don't actually care what λ is.
- Alternatively, you can try to kill every variable *except* λ! This is commonly used if each equations involves only one non-λ variable. That is, solve for x in terms of λ; do the same for y and z. Then plug these in the original constraint equation to solve for λ, and hence extract (x, y, z).

Note the second advice bullet is the opposite of the first advice bullet! Again, systems of equations can't be solved by blindly following recipes. You should use whatever method you think makes sense for the given problem. You don't need anyone's permission to use so-and-so approach.

Warning: Make sure you only divide by nonzero things!

Remember: watch out for division by zero! For example, if you get to the equation $2\lambda y = 10y$, for example, this implies EITHER $\lambda = 5$ OR $y = 0$.

During the 2024 midterm, an equation like this appeared as part of a step in a standard Lagrange multipliers question. Something like 50%-80% of students who got this equation would forget one of the two cases (which one they forgot about varied). Don't let this be you! Whenever you try to cancel, check for division by zero!

Here's an example where a good idea is to kill λ ASAP:

Sample Question

Use Lagrange multipliers to find the smallest possible value of $x^2 + xy + y^2 + y$ subject to the constraint $x + 2y = 3$.

Solution. We want to minimize the function $f(x,y) = x^2 + xy + y^2 + y$ subject to the constraint $g(x,y) = x + 2y = 3$.

0. The constraint region has no boundary, but limit cases along the line $x + 2y = 3$ if either $x \to +\infty$ and $y \to -\infty$ and either $x \to -\infty$ and $y \to +\infty$.
1. Computing the gradients gives

$$\begin{aligned} \nabla f &= \begin{pmatrix} 2x+y \\ x+2y+1 \end{pmatrix} \\ \nabla g &= \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \end{aligned}$$

   Note ∇g is never $\mathbf{0}$. Thus, the Lagrange multiplier equation becomes:

$$\begin{pmatrix} 2x+y \\ x+2y+1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

   This gives the following system of equations:

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

   Combining these two equations gives us an easy way to get rid of λ:

$$x + 2y + 1 = 2 \cdot \underbrace{(2x+y)}_{=\lambda} = 4x + 2y.$$

   Cancel $2y$ on both sides gives us x:

$$x + 1 = 4x \implies x = \frac{1}{3}$$

   Now substitute $x = \frac{1}{3}$ into the constraint $x + 2y = 3$ to get:

$$2y = 3 - \frac{1}{3} = \frac{8}{3} \implies y = \frac{4}{3}.$$

   In other words, the only LM-critical point is $\left(\frac{1}{3}, \frac{4}{3}\right)$, at which point we have

$$f\left(\frac{1}{3}, \frac{4}{3}\right) = \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{4}{3}\right) + \left(\frac{4}{3}\right)^2 + \frac{4}{3} = \frac{1}{9} + \frac{4}{9} + \frac{16}{9} + \frac{4}{3} = \frac{11}{3}.$$

2. There is no boundary to check.
3. If any of the variables goes to $+\infty$ (and hence the other goes to $-\infty$), the value of f will become large too.[17]

In conclusion, the global minimum is $f\left(\frac{1}{3}, \frac{4}{3}\right) = \frac{11}{3}$. □

And here's an example where we kill every variable *except* λ:

[17] This is actually a bit tricky to see, because in this limit case you have two positive terms x^2 and y^2 and one negative term xy. One idea is to write

$$f(x,y) \approx \frac{3}{4}x^2 + \left(\frac{x}{2} + y\right)^2 = \frac{3}{4}y^2 + \left(x + \frac{y}{2}\right)^2$$

for large x and y. The first expression shows that if x is big, then so is f; The first expression shows that if y is big, then so is f.

Sample Question

Use Lagrange multipliers to find the smallest possible value of $x^2 + y^2 + z^2 + y - z$ subject to the constraint $x + 2y + 3z = 4$.

Solution. We want to minimize the function $f(x,y,z) = x^2 + y^2 + z^2 + y - z$ subject to the constraint $(x,y,z) = x + 2y + 3z = 4$.

0. The $g = 4$ region is a plane with no boundary but limit cases if any variable becomes $\pm\infty$.
1. Let's find all the LM-critical points. We start by calculating all the gradients:

$$\nabla f = \begin{pmatrix} 2x \\ 2y+1 \\ 2z-1 \end{pmatrix}$$
$$\nabla g = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Note the gradient ∇g is never $\mathbf{0}$. Thus, the Lagrange multiplier equation becomes:

$$\begin{pmatrix} 2x \\ 2y+1 \\ 2z-1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

or

$$\begin{aligned} 2x &= \lambda, \\ 2y+1 &= 2\lambda, \\ 2z-1 &= 3\lambda. \end{aligned}$$

Let's get rid of every variable besides λ, by solving in λ:

$$\begin{aligned} 2x = \lambda &\implies x = \frac{\lambda}{2} \\ 2y+1 = 2\lambda &\implies y = \lambda - \frac{1}{2} \\ 2z-1 = 3\lambda &\implies z = \frac{3\lambda+1}{2}. \end{aligned}$$

Now substitute $x = \frac{\lambda}{2}$, $y = \lambda - \frac{1}{2}$, and $z = \frac{3\lambda+1}{2}$ into the constraint $x + 2y + 3z = 4$:

$$\begin{aligned} \frac{\lambda}{2} + 2\left(\lambda - \frac{1}{2}\right) + 3\left(\frac{3\lambda+1}{2}\right) &= 4 \\ \implies \frac{\lambda}{2} + 2\lambda - 1 + \frac{9\lambda+3}{2} &= 4 \\ \implies \frac{10\lambda+3}{2} + 2\lambda - 1 &= 4 \\ \implies 10\lambda + 3 + 4\lambda - 2 &= 8 \\ \implies 14\lambda + 1 &= 8 \\ \implies \lambda &= \frac{1}{2}. \end{aligned}$$

Now that we have λ, plug back in to get (x, y, z):

$$x = \frac{\lambda}{2} = \frac{1}{4},$$

$$y = \lambda - \frac{1}{2} = \frac{1}{2} - \frac{1}{2} = 0,$$

$$z = \frac{3\lambda + 1}{2} = \frac{3\left(\frac{1}{2}\right) + 1}{2} = \frac{\frac{3}{2} + 1}{2} = \frac{5}{4}.$$

This gives a single LM-critical point $(x, y, z) = \left(\frac{1}{4}, 0, \frac{5}{4}\right)$, where $f\left(\frac{1}{4}, 0, \frac{5}{4}\right) = \frac{3}{8}$.

2. There are no boundary cases to consider.
3. The limit cases are if one of the variables goes to $\pm\infty$. However, in such a situation f obviously becomes large, so there are no minimums in the situation.

In conclusion, the global minimum is $f\left(\frac{1}{4}, 0, \frac{5}{4}\right) = \frac{3}{8}$. □

Digression: Clever geometric approach for observant students

We show a way you can skip all the calculus steps in the previous problem if you can see how to rewrite the question as a geometry one. This approach is so clever that you don't even need to find $\left(\frac{1}{4}, 0, \frac{5}{4}\right)$; it will directly tell you the minimum value. Don't try this on an exam unless you really know what you're doing.

Let $P := (x, y, z)$ be a point. Let $\mathcal{H}$ denote the plane $x + 2y + 3z = 4$. The magic trick is to rewrite

$$f = x^2 + y^2 + z^2 + y - z = x^2 + \left(y + \frac{1}{2}\right)^2 + \left(z - \frac{1}{2}\right)^2 - \frac{1}{2} = PQ^2 - \frac{1}{2}$$

where $Q := \left(0, -\frac{1}{2}, \frac{1}{2}\right)$. In other words, $f(P)$ is a shift of the squared distance of P from Q. So actually, the global minimum we found $\left(\frac{1}{4}, 0, \frac{5}{4}\right)$ is the point on the plane $\mathcal{H}$ closest to $Q = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$.

But if all you care about is the distance from $Q = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$ to the $\mathcal{H}$, then using calculus is overkill: instead use the recipe from way back in Section 5.6:

$$\min(PQ) = \frac{\left|1 \cdot 0 + 2 \cdot -\frac{1}{2} + 3 \cdot \frac{1}{2} - 4\right|}{\sqrt{1^2 + 2^2 + 3^2}} = \frac{7}{2\sqrt{14}}.$$

So the minimum is

$$\min(f) = \min(PQ)^2 - \frac{1}{2} = \left(\frac{7}{2\sqrt{14}}\right)^2 - \frac{1}{2} = \frac{3}{8}.$$

Thus, an extremely clever student could have bypassed the entire problem by translating it into a geometry question. Don't worry, you won't be expected to come up with something like this in 18.02.

To top all that of, here is a Lagrange multipliers example that requires considering tons of cases. This is probably too lengthy of a calculation for 18.02 because of the amount of arithmetic required; it's here just to illustrate.

Sample Question

Compute the minimum and maximum possible of $x^3 + 3y^3 + 4z^3$ subject to $x^4 + y^4 + z^4 = 2$.

Solution. The region has no boundary nor limit cases. So, we will only focus on calculating the LM-critical points.

The gradients of $f(x,y,z) = x^3 + 3y^3 + 4z^3$ and $g(x,y,z) = x^4 + y^4 + z^4$ are given by

$$\nabla f = \begin{pmatrix} 3x^2 \\ 9y^2 \\ 12z^2 \end{pmatrix}$$

$$\nabla g = \begin{pmatrix} 4x^3 \\ 4y^3 \\ 4z^3 \end{pmatrix}.$$

We begin by simplifying each equation:

- $3x^2 - 4\lambda x^3 = 0$ becomes: $x^2(3 - 4\lambda x) = 0$. This gives two possibilities:
  - $x = 0$, or
  - $\lambda = \frac{3}{4x}$ (assuming $x \neq 0$).
- $9y^2 - 4\lambda y^3 = 0$ becomes: $y^2(9 - 4\lambda y) = 0$. This gives two possibilities:
  - $y = 0$, or
  - $\lambda = \frac{9}{4y}$ (assuming $y \neq 0$).
- $12z^2 - 4\lambda z^3 = 0$ becomes: $z^2(12 - 4\lambda z) = 0$. This gives two possibilities:
  - $z = 0$, or
  - $\lambda = \frac{3}{z}$ (assuming $z \neq 0$).

This gives a total of eight cases! We will go through them all individually.

Case 1 where $x = 0$, $y = 0$, $z \neq 0$ From the constraint $z^4 = 2$, we get:

$$z = \pm\sqrt[4]{2}.$$

Thus, $\lambda = \frac{3}{z} = \pm\frac{3}{\sqrt[4]{2}}$.

Case 2 where $x = 0$, $y \neq 0$, $z = 0$ From the constraint $y^4 = 2$, we get:

$$y = \pm\sqrt[4]{2}.$$

Thus, $\lambda = \frac{9}{4y} = \pm\frac{9}{4\sqrt[4]{2}}$.

Case 3 where $x \neq 0$, $y = 0$, $z = 0$ From the constraint $x^4 = 2$, we get:

$$x = \pm\sqrt[4]{2}.$$

Thus, $\lambda = \frac{3}{4x} = \pm\frac{3}{4\sqrt[4]{2}}$.

Case 4 where $x = 0$, $y \neq 0$, $z \neq 0$ From the constraint $y^4 + z^4 = 2$, we solve using $\lambda = \frac{9}{4y}$ and $\lambda = \frac{3}{z}$. Equating these, we get:

$$\frac{9}{4y} = \frac{3}{z} \implies z = \frac{4}{3}y.$$

Substitute into the constraint:

$$y^4 + \left(\frac{4}{3}y\right)^4 = 2 \implies y^4 + \frac{256}{81}y^4 = 2.$$

This simplifies to:

$$\frac{337}{81}y^4 = 2 \implies y^4 = \frac{162}{337} \implies y = \pm\sqrt[4]{\frac{162}{337}}.$$

Thus, $z = \pm\frac{4}{3}\sqrt[4]{\frac{162}{337}}$.

Case 5 where $x \neq 0$, $y = 0$, $z \neq 0$ From the constraint $x^4 + z^4 = 2$, we solve using $\lambda = \frac{3}{4x}$ and $\lambda = \frac{3}{z}$. Equating these, we get:

$$\frac{3}{4x} = \frac{3}{z} \implies z = 4x.$$

Substitute into the constraint:

$$x^4 + (4x)^4 = 2 \implies x^4 + 256x^4 = 2 \implies 257x^4 = 2.$$

Thus, $x^4 = \frac{2}{257}$, and:

$$x = \pm\sqrt[4]{\frac{2}{257}}, \quad z = \pm 4\sqrt[4]{\frac{2}{257}}.$$

Case 6 where $x \neq 0$, $y \neq 0$, $z = 0$ From the constraint $x^4 + y^4 = 2$, we solve using $\lambda = \frac{3}{4x}$ and $\lambda = \frac{9}{4y}$. Equating these, we get:

$$\frac{3}{4x} = \frac{9}{4y} \implies y = 3x.$$

Substitute into the constraint:

$$x^4 + (3x)^4 = 2 \implies x^4 + 81x^4 = 2 \implies 82x^4 = 2.$$

Thus, $x^4 = \frac{1}{41}$, and:

$$x = \pm\sqrt[4]{\frac{1}{41}}, \quad y = \pm 3\sqrt[4]{\frac{1}{41}}.$$

Case 7 where $x \neq 0$, $y \neq 0$, $z \neq 0$ Equating the three expressions for λ:

$$\frac{3}{4x} = \frac{9}{4y} = \frac{3}{z}.$$

From $\frac{3}{4x} = \frac{9}{4y}$, we get $y = 3x$, and from $\frac{3}{4x} = \frac{3}{z}$, we get $z = 4x$. Substitute into the constraint:

$$x^4 + (3x)^4 + (4x)^4 = 2 \implies x^4 + 81x^4 + 256x^4 = 2 \implies 338x^4 = 2.$$

Thus, $x^4 = \frac{2}{338}$, and:

$$x = \pm\sqrt[4]{\frac{1}{169}}, \quad y = \pm 3\sqrt[4]{\frac{1}{169}}, \quad z = \pm 4\sqrt[4]{\frac{1}{169}}.$$

Since $169 = 13^2$ is a square, this could be written more simply as

$$x = \pm\frac{1}{\sqrt{13}}, \quad y = \pm\frac{3}{\sqrt{13}}, \quad z = \pm\frac{4}{\sqrt{13}}.$$

Case 8 where $x = 0$, $y = 0$, $z = 0$ This doesn't yield a valid solution because it doesn't like on the constraint $x^4 + y^4 + z^4 = 2$.

Hence there are a whopping total of 26 LM-critical points. They are:

- $x = 0$, $y = 0$, $z = \pm\sqrt[4]{2}$,
- $x = 0$, $y = \pm\sqrt[4]{2}$, $z = 0$,
- $x = \pm\sqrt[4]{2}$, $y = 0$, $z = 0$,
- $x = 0$, $y = \pm\sqrt[4]{\frac{162}{337}}$, $z = \pm\frac{4}{3}\sqrt[4]{\frac{162}{337}}$.
- $x = \pm\sqrt[4]{\frac{2}{257}}$, $z = \pm 4\sqrt[4]{\frac{2}{257}}$, $y = 0$,
- $x = \pm\sqrt[4]{\frac{1}{41}}$, $y = \pm 3\sqrt[4]{\frac{1}{41}}$, $z = 0$,
- $x = \pm\frac{1}{\sqrt{13}}$, $y = \pm\frac{3}{\sqrt{13}}$, $z = \pm\frac{4}{\sqrt{13}}$.

When searching for the maximum, we should always take $+$ for $\pm$ to maximize $f(x, y, z)$; similarly, the minimum uses only $-$ for $\pm$. Note also that plugging in all $-$'s is the negative of plugging in all $+$'s. So this reduces us from 26 cases to just 7. If we actually try all seven, we find that the last one is the optimal one; that is, the maximum and minimums are

$$f\left(\frac{1}{\sqrt{13}}, \frac{3}{\sqrt{13}}, \frac{4}{\sqrt{13}}\right) = 2\sqrt{13}$$
$$f\left(-\frac{1}{\sqrt{13}}, -\frac{3}{\sqrt{13}}, -\frac{4}{\sqrt{13}}\right) = -2\sqrt{13}.$$ □

§19.6 [TEXT] Example of easy case with a "common-sense" boundary

As we alluded to earlier, the boundary of the so-called "easy case" can have infinitely many points, so you cannot just plug them in one by one to inspect them all. In some situations, it will still be doable just by inspection, because the function is really easy to describe on the boundary. We give one such example below.

Sample Question

Compute the minimum possible value of

$$f(x, y) = x^2 + y^2 + xy - 6y$$

in the first quadrant $x, y \geq 0$.

Solution. Follow the recipe:

0. The first quadrant has limit cases if either $x \to +\infty$ or $y \to +\infty$. The boundary consists of two rays: the positive x axis (from $(0, 0)$ due east) and the positive y axis (from $(0, 0)$ due north).
1. We calculate the critical points. The gradient is given by

$$\nabla f = \begin{pmatrix} 4x + y \\ 2y + x - 6 \end{pmatrix}.$$

Setting the gradient to the zero vector, we need to solve the system

$$4x + y = 0$$
$$x + 2y - 6 = 0$$

which is an easy linear system. If we let $y = -4x$ from the first equation, we get

$$2(-4x) + x - 6 = 0 \Longrightarrow -8x + x - 6 = 0 \Longrightarrow -7x = 6 \Longrightarrow x = -\frac{6}{7}.$$

Since we are working in the first quadrant where $x \geq 0$, this value is not valid. Therefore, there are no critical points in our region $\mathcal{R}$.

2. As we saw, we have two boundaries. There are infinitely many points on the positive x and y axis, so we cannot just plug them all in one by one. In principle, you could redo the entire easy-case recipe for each boundary parts, and it will work.

   However, luckily, the function is quite easy to analyze on each part, and we can do it with just high school algebra, no calculus needed. This way we don't need to go through the whole recipe again.

   The boundary of $x = 0$ and $y \geq 0$ Substitute $x = 0$ into the function: $f(0, y) = 2(0)^2 + y^2 + 0 \cdot y - 6y = y^2 - 6y$. In other words, we need to see what the smallest value of

   $$f(0, y) = y^2 - 6y \quad \text{across all} \quad y \geq 0$$

   could be. You *could* use the derivative of y, but I think the fastest thing to do is actually complete the square: the function

   $$f(0, y) = (y - 3)^2 - 9$$

   obviously has the smallest value at $f(0, 3) = -9$. (And $f(0, 0) = 0$ and $f(0, +\infty) = +\infty$ are both worse.)

   The boundary of $y = 0$ and $x \geq 0$ Substitute $y = 0$ into the function:

   $$f(x, 0) = 2x^2 + 0^2 + x(0) - 6(0) = 2x^2.$$

   It's obvious that $f(x, 0)$ is minimized at $x = 0$. (Note that $f(+\infty, 0) = +\infty$ is worse.)

3. In the limit case where either $x \to +\infty$ and $y \to +\infty$ it's clear that $f \to +\infty$.

In conclusion, the best value is actually the one from Step 2: we have $f(0, 3) = -9$ being the smallest possible value. □

§19.7 [TEXT] Example of easy case that ends up using Lagrange multipliers for the boundary

Now here's a case where the boundary requires Lagrange multipliers. So, it's really a hard-case optimization problem *within* an easy-case optimization problem. If you've seen the movie *Inception*, yes, one of those.

Sample Question

Compute the minimum and maximum possible value, if they exist, of

$$f(x, y, z) = x^4 + y^4 + z^4$$

over the region $x^2 + y^2 + z^2 \leq 1$.

Solution. At first glance, this seems like it should be in the easy case! The region $\mathcal{R}$ consisting of the closed ball $x^2 + y^2 + z^2 \leq 1$ is indeed three-dimensional. But the reason this sample question is in this section is because we will find that checking the boundary case requires another application of Lagrange multipliers.

Let's carry out the easy case recipe.

0. There are no limit cases, but a boundary $x^2 + y^2 + z^2 = 1$, the unit sphere (not filled).
1. First let's find the critical points of $f(x, y, z) = x^4 + y^4 + z^4$. Write

$$\nabla f = \begin{pmatrix} 4x^3 \\ 4y^3 \\ 4z^3 \end{pmatrix}.$$

   Solving the insultingly easy system of equations $4x^3 = 4y^3 = 4z^3 = 0$ we see the only critical point is apparently $x = y = z = 0$. The value there is $f(0, 0, 0) = 0$.
2. The boundary of $\mathcal{R}$ is $x^2 + y^2 + z^2 = 1$, the unit sphere; we denote this sphere by $\mathcal{S}$. So now we have to evaluate f on this boundary. The issue is that there are too many points on this unit sphere! We can't just check them one by one. And unlike the previous example, the function is not simple enough that we can use common sense to deal with it.

   Therefore, we will use Lagrange multipliers with the constraint function $g(x, y, z) = x^2 + y^2 + z^2$ to find the minimum possible value of f on this new region $\mathcal{S}$.

   0. The new region $\mathcal{S}$ has no boundary and no limit cases.
   1. Let's find the LM-critical points for f on $\mathcal{S}$. Take the gradient of g to get

$$\nabla g = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}.$$

      The only point at which $\nabla g = \mathbf{0}$ is $x = y = z = 0$ which isn't on the sphere $\mathcal{S}$, so we don't have to worry about $\nabla g = \mathbf{0}$ the case. Now we instead solve

$$\begin{pmatrix} 4x^3 \\ 4y^3 \\ 4z^3 \end{pmatrix} = \lambda \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}.$$

      This requires some manual labor to solve, because there are lots of cases. The equation for x says that

$$4x^3 = \lambda \cdot 2x \iff x = 0 \quad \text{or} \quad x = \pm\sqrt{\frac{\lambda}{2}}$$

      and similarly for y and z:

$$4y^3 = \lambda \cdot 2y \iff y = 0 \quad \text{or} \quad y = \pm\sqrt{\frac{\lambda}{2}}$$
$$4z^3 = \lambda \cdot 2z \iff z = 0 \quad \text{or} \quad z = \pm\sqrt{\frac{\lambda}{2}}$$

      A priori, this seems like it will require us to take a lot of cases. However, we can take advantage of symmetry to reduce the amount of work we have to do. The trick is to get rid of λ as follows:

      Observation: All the nonzero variables x, y, z should have the same absolute value.

      To spell it out:

- If all three variables are nonzero, then $|x| = |y| = |z| = \frac{1}{\sqrt{3}}$ (because $x^2 + y^2 + z^2 = 1$ as well).
- If two variables are nonzero, then their absolute values are $\frac{1}{\sqrt{2}}$ by the same token.
- And if only one variable is nonzero, it is ± 1.

(Note of course that $(0,0,0)$ does not lie on $\mathcal{S}$.) Think about why this works.

So there are **26 LM-critical points** given by the following list:
- $\left(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right)$; there are 8 points in this case. The f-values are all $\frac{1}{3}$.
- $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, 0\right)$; there are 4 points in this case. The f-values are all $\frac{1}{2}$.
- $\left(\pm\frac{1}{\sqrt{2}}, 0, \pm\frac{1}{\sqrt{2}}\right)$; there are 4 points in this case. The f-values are all $\frac{1}{2}$.
- $\left(0, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$; there are 4 points in this case. The f-values are all $\frac{1}{2}$.
- $(\pm 1, 0, 0)$; there are 2 points in this case. The f-values are all 1.
- $(0, \pm 1, 0)$; there are 2 points in this case. The f-values are all 1.
- $(0, 0 \pm 1)$; there are 2 points in this case. The f-values are all 1.

Phew! Okay. The other cases are much shorter:

2. $\mathcal{S}$ has no boundary to consider.

3. $\mathcal{S}$ has no limit cases to consider.

3. $\mathcal{R}$ has no limit cases to consider.

Okay, marathon done. Collate everything together. The values of f we saw were 0, $\frac{1}{3}$, $\frac{1}{2}$ and 1, and there were no limit cases of any sort. So:

- $f(0,0,0) = 0$ is the global minimum.
- $f(\pm 1, 0, 0) = f(0, \pm 1, 0) = f(0, 0, \pm 1) = 1$ are the global maximums. $\square$

§19.8 [SIDENOTE] A little common sense can you save you a lot of work

If you step back and think a bit before you try to dive into calculus, you might find that having a bit of "common sense" might save you a lot of work. What I mean is, imagine you gave the question to your high school self before you learned *any* calculus at all. Would they be able to say anything about what properties the answer could have? The answer is, yes, pretty often.

Let's take the example we just did: we asked for the minimum and maximum of

$$f(x, y, z) = x^4 + y^4 + z^4$$

over the region $x^2 + y^2 + z^2 \le 1$. To show the recipe, I turned off my brain and jumped straight into a really long calculation. But it turns out you can cut out a lot of the steps if you just use some common sense, not involving any calculus:

- The *minimum* is actually obvious: it's just 0, because fourth powers are always nonnegative! So $f \ge 0$ is obvious even to a high schooler, and $f(0,0,0) = 0$.
- For the *maximum* you can actually see *a priori* that it must occur on the boundary $x^2 + y^2 + z^2 = 1$. Why is this? Suppose you had a point in the strict interior $P = (0.1, 0.2, 0.3)$ with $f > 0$. Then $f(P) = f(0.1, 0.2, 0.3)$ is some number. But you could obviously increase the value of f just by scaling the absolute value of things in P! For example, if I double all the coordinates of P to get $Q = (0.2, 0.4, 0.6)$, then $f(Q) = 16 f(P)$. As long as Q stays within the sphere, this will be a better value.

  So any point in the interior is obviously not a maximum: if you have a point P strictly the interior, you could increase $f(P)$ by moving P farther from the origin.

That means if we had used a bit of common sense, we could have gotten the minimum with no work at all, and we could have skipped straight to the LM step for the maximum. So if you aren't too overwhelmed by material already in this class, look for shortcuts like this when you can.

Digression: Even faster solution for observant students

We saw the minimum is obviously 0 if you just thought about it. In fact, you can *also* similarly find the maximum with no calculus at all if you realize the answer should be $f(1,0,0) = 1$.

Here's how. What we're trying to prove is that

$$x^4 + y^4 + z^4 \leq 1$$

whenever $x^2 + y^2 + z^2 \leq 1$. Because y^2 and z^2 are nonnegative, it's obvious that $x^2 \leq 1$. But in fact it's easy to see that

$$x^4 \leq x^2 \quad \text{is true whenever} \quad x^2 \leq 1.$$

Similarly, $y^4 \leq y^2$ and $z^4 \leq z^2$. Thus

$$x^4 + y^4 + z^4 \leq x^2 + y^2 + z^2 = 1.$$

§19.9 [SIDENOTE] Compactness as a way to check your work

This is an optional section containing a nice theorem from 18.100 that could help you check your work, but isn't necessary in theory if you never make mistakes. (But in practice…)

I need a new word called "compact", and like before, it's beyond the scope of 18.02 to give a proper definition. However, I will hazard the following one: for 18.02 examples, $\mathcal{R}$ **is compact if there are no limit cases**. That is,

- All the constraints are $=$, $\leq$, or $\geq$; no $<$ or $>$,
- None of the variables can go to $\pm\infty$.

Then the theorem I promised you is:

Tip: Compact optimization theorem

If $\mathcal{R}$ is a compact region, and f is a function to optimize on the region which is continuously defined everywhere, then there must be at least one global minimum, and at least one global maximum.

This works in both the easy case (no Lagrange multipliers) and the hard case (with Lagrange multipliers).

Here's some examples of how this theorem can help you:

Example

- Suppose you are asked to optimize a continuous function $f(x, y)$ on the square $-1 \leq x \leq 1$, $-1 \leq y \leq 1$. We saw this square has no limit cases. Then the compact optimization theorem promises you that the answer "no global minimum" or "no global maximum" will never occur.
- Suppose you are asked to optimize a continuous function $f(x, y, z)$ on the sphere $x^2 + y^2 + z^2 = 1$ (which means you are probably going to use Lagrange multipliers). We saw this sphere has no limit cases (and not even a boundary). Then the compact optimization theorem promises you that the answer "no global minimum" or "no global maximum" will never occur.
- Suppose you are asked to optimize a continuous function $f(x, y, z)$ on the closed ball $x^2 + y^2 + z^2 \leq 1$, like in the last example. This closed ball also has no limit cases, so the compact optimization theorem promises you that the answer "no global minimum" or "no global maximum" will never occur.

§19.10 [RECAP] Recap of Part Foxtrot on Optimization

- We introduced the notion of critical points as points where $\nabla f = \mathbf{0}$.
  - We saw that critical points could be local minimums, local maximums, or saddle points.
  - We introduced the second derivative test as a way to tell some of these cases apart, although the second derivative test can be inconclusive.
- We talked about how regions have dimensions, limit cases, and boundaries. Although we didn't give a proper definition, we explain rules of thumb that work in 18.02.
- For optimization problems with no $=$ constraints, we check the critical points, limit cases, and boundaries.
- For optimization problems with one $=$ constraint, we check the LM-critical points, limit cases, and boundaries.
- Finding either critical points or LM-critical points involves solving systems of equations. There is no fixed recipe for this, but we gave some possible strategies that you can try depending on the exact shape of the system you get.

§19.11 [EXER] Exercises

Exercise 19.1. Let ABC be the triangle in the xy-plane with vertices $A = (0, 12)$, $B = (-5, 0)$, $C = (9, 0)$. For what point P in the plane is the sum

$$PA^2 + PB^2 + PC^2$$

as small as possible?

Exercise 19.2. Compute the minimum possible value of $x + y$ given that $\sin(x) + \sin(y) = 1$ and $x, y \geq 0$.

Exercise 19.3 (Suggested by Ting-Wei Chao)**.** Compute the global minimum of the function

$$f(x, y) = |x^2 + y^2 - 25| - 3x - 4y.$$

Chapter 20. Practice midterm for Parts Delta, Echo, Foxtrot

This is a practice midterm that was given on October 21, 2024,[18] covering topics in Parts Delta, Echo, Foxtrot. Solutions are in Chapter 47.

Exercise 20.1. A butterfly is fluttering in the xy plane with position given by $\mathbf{r}(t) = \langle \cos(t), \cos(t) \rangle$, starting from time $t = 0$ at $\mathbf{r}(0) = \langle 1, 1 \rangle$.

- Compute the speed of the butterfly at $t = \frac{\pi}{3}$.
- Compute the arc length of the butterfly's trajectory from $t = 0$ to $t = 2\pi$.
- Sketch the butterfly's trajectory from $t = 0$ to $t = 2\pi$ in the xy plane.

Exercise 20.2. Let $k > 0$ be a fixed real number and let $f(x, y) = x^3 + ky^2$. Assume that the level curve of f for the value 21 passes through the point $P = (1, 2)$. Compute the equation of the tangent line to this level curve at the point P.

Exercise 20.3. Let $f(x, y) = x^{5y}$ for $x, y > 0$. Use linear approximation to estimate $f(1.001, 3.001)$ starting from the point $(1, 3)$.

Exercise 20.4. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = \cos(\pi x) + \frac{y^4}{4} - \frac{y^3}{3} - y^2.$$

- Compute all the critical points and classify them as saddle point, local minimum, or local maximum.
- Compute the global minimums and global maximums of f, if they exist.

Exercise 20.5. Compute the minimum and maximum possible value of $x + 2y + 2z$ over real numbers x, y, z satisfying $x^2 + y^2 + z^2 \le 100$.

Exercise 20.6. Consider the level surface of $f(x, y, z) = (x - 1)^2 + (y - 1)^3 + (z - 1)^4$ that passes through the origin $O = (0, 0, 0)$. Let $\mathcal{H}$ denote the tangent plane to this surface at O. Give an example of two nonzero tangent vectors to this surface at O whose span is $\mathcal{H}$.

[18] A cute quote from an anonymous student: "lowk i feel like ur so cracked every review thing u give is like 3x harder than the actual thing 😭 but thank you for hosting these ur amazing".

Part Golf: 2D integrals of scalar functions

For comparison, Part Golf corresponds to §13.1-§13.5 of Poonen's notes.

Chapter 21. A zoomed out pep talk of Part Golf

This whole chapter is a pep talk. We'll get to recipes and details in later sections.

§21.1 [TEXT] The big table of integrals

The rest of 18.02 is going to cover a bunch of different integrals. If you've been following my advice to pay attention to type safety so far, it'll help you here. I'll freely admit that I (Evan) often make type-errors in this part of 18.02 as well, so don't let your guard down.

Remember that:

> **Idea**
>
> Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is given, and $0 \le d \le n$. The goal of a d-dimensional integral of f is to add up all the values of f among some d-dimensional object living in $\mathbb{R}^n$.

For example, this idea even makes sense for $d = 0$! In 18.02, a 0-dimensional object is a point (or a bunch of points), and you can evaluate f at a point by just plugging it in. So philosophically, a 0-dimensional integral is just a finite sum of f at some points. This might seem stupid that I bring up this degenerate case, but it turns out later when we cover div/grad/curl the 0-dimensional case is relevant.

Here's a giant chart in Figure 41. (The chart is so big it doesn't quite fit in the page, but you can download a large PDF version).

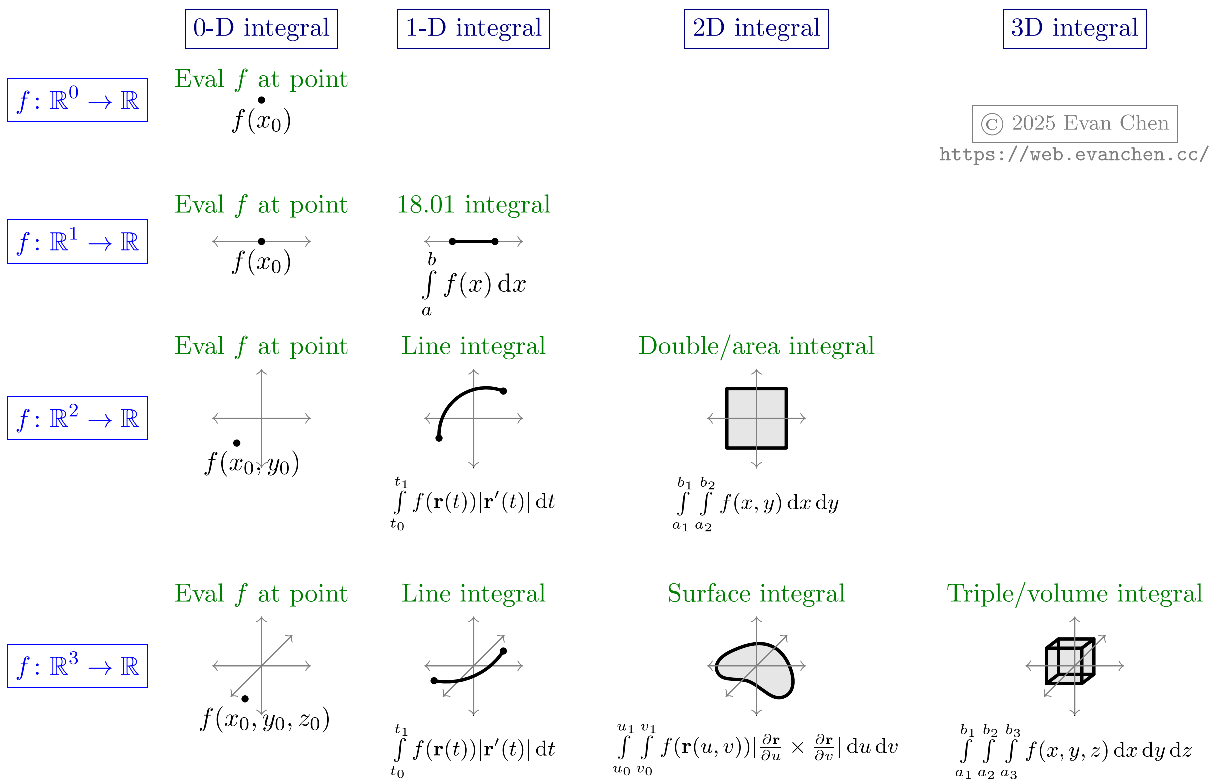

Figure 41: For each $0 \le d \le n \le 3$, we draw the kind of integral and give it a name. Download at https://web.evanchen.cc/textbooks/poster-ints.pdf.

This chart has ten different kinds of integrals, one for each (d, n) with $0 \le d \le n \le 3$. Here's a rundown of the things in the chart.

- The case $d = 0$ is stupid, as I just said, and it's only here because I'll reference it later.
- The case $d = 1$ and $n = 1$ was covered in 18.01. Good old single-variable integral computed using the antiderivative, via the fundamental theorem of calculus.
- After that, the conceptually simplest cases are actually $d = n = 2$ and $d = n = 3$ — the ones on the diagonal. In general, these might be called **double/area integrals** for $n = 2$ and **triple/volume integrals** for $n = 3$. We'll say a bit in a moment about how to compute these in practice, but the good news is that often you can just chain together old 18.01 integrals; you don't even need a parametrization some of the time.
- When $d = 1$ and $2 \le n \le 3$, what you get are **line integrals**. The idea is that you have a trajectory in $\mathbb{R}^n$ which is defined by some parametric equation $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$. You also have a function $f : \mathbb{R}^n \to \mathbb{R}$. The line integral lets you add up the values of f along the trajectory.

  This is just turns out to be a *single* 18.01 integral. Usually your path is parametrized by a single variable t. So even though the expression inside the integral

  $$\int_{t_0}^{t_1} f(\mathbf{r}(\mathbf{t}))\ |\mathbf{r}'(t)|\,\mathrm{d}t$$

  inside the integral might look intimidating, if you are really given a concrete f and $\mathbf{r}(t)$, then what you *really* have is

  $$\int_{t_0}^{t_1} [\text{expression involving only}\quad t]\,\mathrm{d}t$$

  which is an 18.01 integral! And so that's something you already know how to do.

  In other words, if you have $d = 1$ and $n > 1$, you basically replace it right away with a single integral over the parametrizing line segment. In other words **line integrals translate directly into single 18.01 integrals**.
- When $d = 2$ and $n = 3$, we have the **surface integral**. To compute these, you usually have to parametrize a *surface*; but since a surface is two-dimensional, rather than $\mathbf{r}(t)$ for a time parameter t you have $\mathbf{r}(u, v)$ for two parameters u and v to describe the surface. That makes these a little more annoying.

  But like the line integral, after you work out the parametrization stuff, the surface integral will transform into a 2-variable area integral. In other words **surface integrals translate directly into area integral**.

So the bottom trio — 2D/3D line integral and surface integral — end up being special instances of the single and double integrals. We'll see some examples of this later; but it'll actually be the *last* thing we cover in part Golf. Most of part Golf is dedicated towards double and triple integrals instead.

§21.2 [TEXT] Warning about the bottom trio

The integrals in Figure 41 would be better called **scalar-field line integral** and **scalar-field surface integral** to emphasize that this is integration for a *scalar* function $f : \mathbb{R}^2 \to \mathbb{R}$ or $f : \mathbb{R}^3 \to \mathbb{R}$. The reason you won't see this term much is the following important caveat **the line integral and surface integral in Figure 41 are used rarely**.

In 18.02 the only cases where we use these are:

- Arc length (for $d = 1$) and surface area (for $d = 2$)
- Questions related to mass, or center of mass, of objects like wires (for $d = 1$) or metal plates (for $d = 2$).

In this book, this will happen in Chapter 29, and then after that the line and surface integrals will always be the **vector field** variant instead (in Part India and Juliett). If you want to flip ahead, take a glance at Figure 75 and Table 15.

§21.3 [TEXT] Idea of how these are computed when $d = n$ and $n \geq 2$

So as I just said, focus for now on $d = n = 2$ or $d = n = 3$ (the double and triple integral cells in chart Figure 41).

The easiest cases are when the region you're integrating is a rectangle or prism. Despite looking scary because of the number of integral signs, they are actually considered the "easy case" to think about for practical calculations:

- A double integral over a rectangle is two 18.01 integrals followed one after another.
- A triple integral over a rectangular prism really is three 18.01 integrals followed one after another.

Then there are cases where $d = n = 2$ or $d = n = 3$ but the region is not rectangular. For example, maybe in $\mathbb{R}^2$ you are trying to do an **area integral** over the disk $x^2 + y^2 \leq 1$ or you are trying to do a **volume integral** over the ball $x^2 + y^2 + z^2 \leq 1$ for example.

- Even in this case, sometimes you could still set up a double integral or triple integral without having to change variables. For example, an integral over the disk

  $$\iint_{x^2+y^2\leq 1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

  might actually be rewritten a double integral

  $$\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

  Although it looks more frightening because the limits of integration are expressions and not numbers, it doesn't require any new techniques. It really is just two 18.01 integrals, one after another.

- If rewriting as a double or triple integral fails, then the strategy is instead to **change variables**. This method will be covered extensively in Chapter 23.

So to summarize

Idea

Whenever you try to compute a multivariable integral in Figure 41, your goal is to translate it into a rectangular-looking single/double/triple integral, then evaluate by using your old 18.01 methods many times.

This is actually really, really good news! You might have remembered from 18.01 that computing integrals of single-variable functions like $\int e^x \sin(x)$ was, well, hard![19] Computing antiderivatives was

[19] It's $\frac{1}{2}e^x(\sin(x) - \cos(x)) + C$, by the way.

not easy at all; in fact, it's so nontrivial that MIT students made an event called the integration bee that's like the spelling bee but for integrals (I'm not kidding). You might have feared that in 18.02, you might need to learn something even more horrifying.

But no, you don't! It's a lot like how you might be scared of multivariate differentiation at first, with the symbols ∇f or partial derivatives, until you realize that calculating partial derivatives is something you *actually already know how to do* from 18.01.

The same will be true for multivariable integrals. The challenge won't actually be the anti-derivatives, which are unchanged from 18.01. The hard part will actually be figuring out the *limits* of integration!

Chapter 22. Double integrals

One common theme from 18.02 that you might have noticed in part Foxtrot is that, unlike in 18.01 where you were hyper-focused on the function f you were optimizing, in 18.02 the *region* you're working with deserves a lot of attention. This will be true for the material in this chapter too — you ought to paying most attention to the region before you even look at the function f that's being integrated.

§22.1 [RECIPE] Integrating over rectangles

If you want to integrate over a rectangle, this is super easy. It's basically like partial derivatives, where you pretend some variables are constant and only one variable is going to vary at once. It's easier to see an example before the recipe.

Warning: Some sources might not write the variables in the $\int$'s for you

Rather than writing just $\int_a^b f(t)\,\mathrm{d}t$, I will usually prefer to write $\int_{t=a}^b f(t)\,\mathrm{d}t$, to make it easier to see which variable is integrated over. Not all sources will be nice enough to do this and will actually make you read the $\mathrm{d}x$ and $\mathrm{d}y$ backwards; e.g. if you see

$$\int_0^6 \int_0^1 xy^2 \,\mathrm{d}x\,\mathrm{d}y$$

then this actually means

$$\int_0^6 \left(\int_0^1 xy^2 \,\mathrm{d}x \right) \mathrm{d}y$$

so $0 \le x \le 1$ and $0 \le y \le 6$. For me reading backwards like this is annoying as hell, so I think it's just much easier to write

$$\int_{y=0}^6 \int_{x=0}^1 xy^2 \,\mathrm{d}x\,\mathrm{d}y$$

and I recommend you use that notation instead. The advantage is that then you pretty much don't have to look at the $\mathrm{d}x\,\mathrm{d}y$ at the far right anymore; the information you need is all in one place at the far left.

Sample Question

Integrate $\int_{y=0}^6 \int_{x=0}^1 xy^2 \,\mathrm{d}x\,\mathrm{d}y$.

Solution.

1. The first step is to compute the inner integral with respect to x, treating y as a constant.

   The inner integral is:

$$\int_{x=0}^1 xy^2 \,\mathrm{d}x.$$

   Since y^2 is treated as a constant with respect to x, we can factor it out of the integral:

$$y^2 \int_{x=0}^{1} x \, \mathrm{d}x.$$

Now, compute $\int_{x=0}^{1} x \, \mathrm{d}x$:

$$\int_{x=0}^{1} x \, \mathrm{d}x = \left[\frac{x^2}{2}\right]_0^1 = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2}.$$

Thus, the result of the inner integral is:

$$y^2 \cdot \frac{1}{2} = \frac{y^2}{2}.$$

2. Now, substitute the result of the inner integral into the outer integral:

$$\int_{y=0}^{6} \frac{y^2}{2} \, \mathrm{d}y = \frac{1}{2} \int_{y=0}^{6} y^2 \, \mathrm{d}y = \frac{1}{2} \left[\frac{y^3}{3}\right]_0^6 = \frac{1}{2} \left(\frac{6^3}{3} - \frac{0^3}{3}\right) = \boxed{36}.$$

□

Easy, right? The general recipe is the same.

Recipe for integrating over a rectangle

To integrate something of the form $\int \left(\int \mathrm{d}y\right) \mathrm{d}x$:

1. Evaluate the inner integral as in 18.01, treating x as constant. This should give you some expression in x with no y's left.
2. Replace the inner integral with the result from the previous step to get an 18.01 integral with only x in it. Integrate that.

Here's another example.

Sample Question

Evaluate the double integral:

$$\int_{x=0}^{\pi} \int_{y=0}^{1} x \cos(xy) \, \mathrm{d}y \, \mathrm{d}x.$$

Solution.

1. The first step is to compute the inner integral with respect to y, treating x as a constant. The inner integral is:

$$\int_{y=0}^{1} x \cos(xy) \, \mathrm{d}y.$$

Since x is treated as a constant with respect to y, we can factor x out of the integral:

$$x \int_{y=0}^{1} \cos(xy) \, \mathrm{d}y.$$

Now, we compute $\int_{y=0}^{1} \cos(xy)\,\mathrm{d}y$.

$$\int_{y=0}^{1} \cos(xy)\,\mathrm{d}u = \left[\frac{1}{x}\sin(xy)\right]_0^1 = \frac{\sin(x)}{x}.$$

Thus, the result of the inner integral is:

$$x \cdot \frac{\sin(x)}{x} = \sin(x).$$

2. Now, substitute the result of the inner integral into the outer integral:

$$\int_{x=0}^{\pi} \sin(x)\,\mathrm{d}x.$$

We know that $\int \sin(x)\,\mathrm{d}x = -\cos(x)$. Therefore:

$$\int_{x=0}^{\pi} \sin(x)\,\mathrm{d}x = [-\cos(x)]_0^{\pi} = -\cos(\pi) + \cos(0).$$

Using $\cos(\pi) = -1$ and $\cos(0) = 1$, we get:

$$-(-1) + 1 = 1 + 1 = \boxed{2}.$$

□

§22.2 [RECIPE] Doing xy-integration without a rectangle

In general, a lot of 2D regions $\mathcal{R}$ can still be done with xy integration, even when they aren't rectangles. In that case, the integral is notated

$$\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y := \text{integral of } \ f \ \text{ over } \ \mathcal{R}$$

for whatever function f you're integrating. If the region is given by a few inequalities you can also write the region directly in, i.e. $\iint_{x^2+y^2\le 1} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ would mean the integral of f over the unit disk.

Here's how you do it.

Remark

A lot of other sources might write this as $\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}A$ instead, which is shorter; it's understood that the area element $\mathrm{d}A$ is shorthand for $\mathrm{d}x\,\mathrm{d}y$.

However, when you're starting off I will still explicitly write $\mathrm{d}x\,\mathrm{d}y$, because I don't want to hide the integration variables — training wheels, I guess. That said, if you know what you're doing and want to write $\mathrm{d}A$ to save time, go for it!

We'll talk more about the weirder d symbols later.

Recipe for converting to xy-integration

1. Draw a picture of the region as best you can.
2. Write the region as a list of inequalities.[20]
3. Pick *one* of x and y, and use your picture to describe all the values it could take.
4. Solve for the *other* variable in all the inequalities.

Remark: This recipe works fine for rectangles, too!

You can do this recipe even with a rectangle. If you do, what the recipe tells you that for a rectangle you can integrate in either order: given the rectangle of points (x, y) with $a \le x \le b$ and $c < y \le d$, we have

$$\int_{x=a}^{b} \int_{y=c}^{d} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_{y=c}^{d} \int_{x=a}^{b} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Sometimes this will be easier. One shape of exam question will to be choose f such that the left-hand side is annoying to calculate directly but the right-hand side is easy to calculate, and ask for the left-hand side. So this is meant to test your ability to recognize when the other order is better.

For example, let's take the region in Poonen's example 13.1:

Sample Question

Show both ways of setting up an integral of a function $f(x, y)$ over the region bounded by $y - x = 2$ and $y = x^2$.

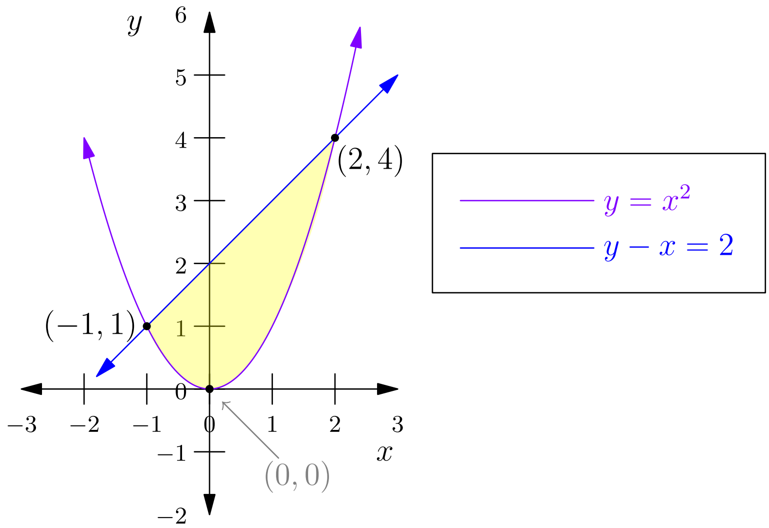

Figure 42: The region between $y = x^2$ and $y - x = 2$.

Solution. See Figure 42. There are two intersection points that it's pretty clear we'll want to know, so we can solve for those intersection points by solving the system and add them to our picture:

[20] I don't think other sources always write the inequalities the way I do. But I think this will help you a lot with making sure bounds go the right way.

$$\begin{cases} y - x = 2 \\ y = x^2 \end{cases} \implies x + 2 = x^2 \implies x = -1 \quad \text{or} \quad x = 2$$
$$\implies (x, y) = (-1, 1) \quad \text{or} \quad (x, y) = (2, 4).$$

I'll also mark $(0, 0)$, the bottom of the parabola.

So we want the part of the plane that lies *above* the parabola $y = x^2$ but *below* the line $y - x = 2$. So I think you'll find things easier to think about if you consider the region as the system of inequalities

$$\begin{aligned} y &\ge x^2 \\ y - x &\le 2. \end{aligned}$$

Now there are two ways to do the slicing, depending on which of x and y you want outside.

If x is outer First, let's imagine we let x be the outer integral. Then from the picture, you can see $-1 \le x \le 2$. If we solve for y, we find its region is

$$x^2 \le y \le x + 2.$$

See Figure 43.

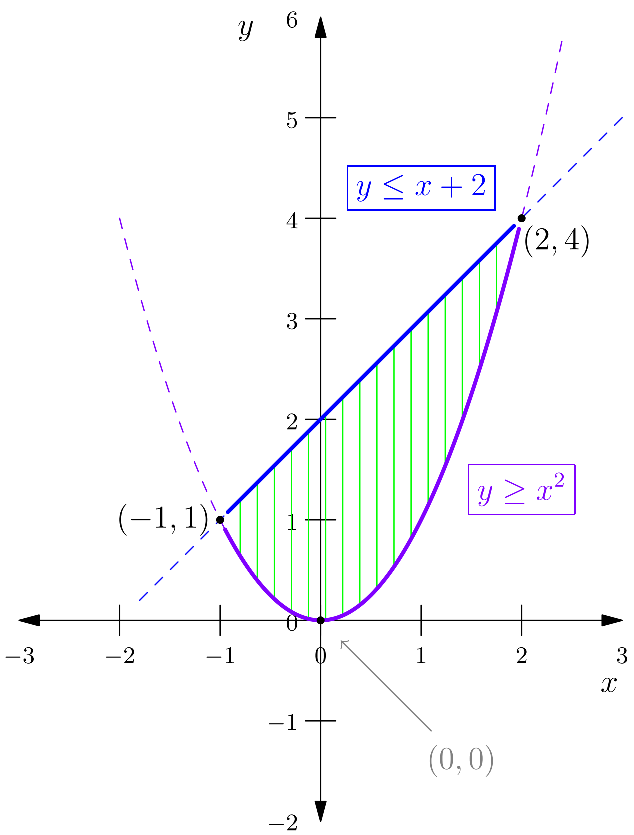

Figure 43: Dissecting Figure 42 vertically, which is pretty nice. There's a single top lid (blue) and a bottom lip (purple) so that for each given x the slice of y (drawn in green) is easy to describe.

Hence, we get the double integral as

$$\int_{x=-1}^{2} \int_{y=x^2}^{x+2} f(x, y) \, \mathrm{d}y \, \mathrm{d}x.$$

If y is outer On the other hand, let's imagine we used y first. From the picture, we see that y ranges from 0 all the way up to 4. (So in what follows I'll write $y \geq 0$ to make notation better.)

But x is gnarly. The issue is that when you solve for x you get *three* inequalities:

- $y \leq x^2$ solves to $-\sqrt{y} \leq x \leq \sqrt{y}$
- $y - x \leq 2$ solves to $y - 2 \leq x$.

See Figure 44.

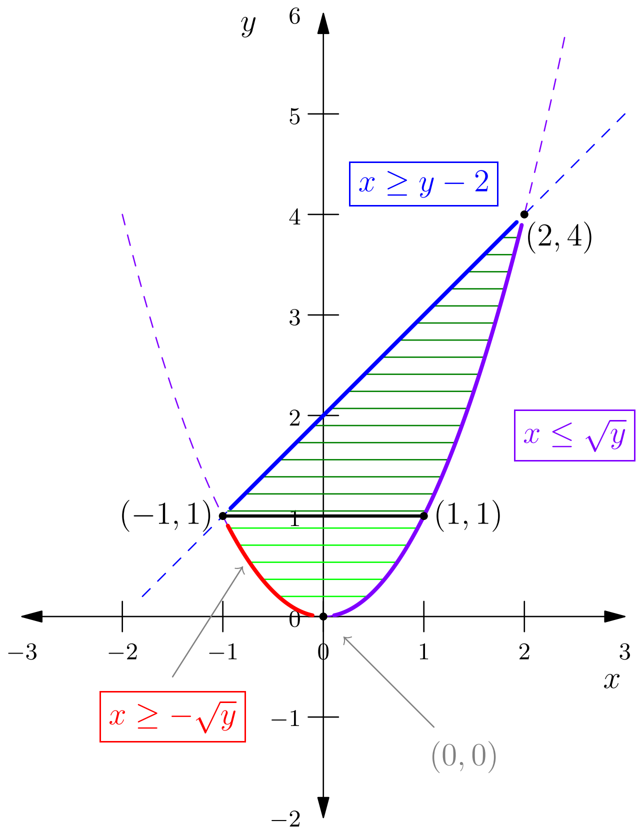

Figure 44: Dissecting Figure 42 horizontally, which is less nice: there are cases. Above the line $y = 1$, you have a blue wall to the left and a curved purple arc to the right. But below $y = 1$, you instead have a red arc of the parabola to the left, and a purple arc of the parabola to the right.

If you know how the max function works, you could even write this as

$$\max(y - 2, -\sqrt{y}) \leq x \leq \sqrt{y}.$$

The main issue is that the lower endpoint for x behaves differently with cases. For $y \leq 1$, the bound of $-\sqrt{y}$ triumphs over the bound of $y - 2$. But for $y \geq 1$, the bound of $y - 2$ is the more informative inequality. So if we wanted to write this as a double integral, we would actually have to split into two:

$$\int_{y=0}^{1} \int_{x=-\sqrt{y}}^{\sqrt{y}} f(x, y) \, \mathrm{d}x \, \mathrm{d}y + \int_{y=1}^{4} \int_{x=y-2}^{\sqrt{y}} f(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

$\square$

§22.3 [TEXT] Example with a concrete function f

In the previous example we showed how one would integrate a random function f over the region between the line $y - x = 2$ and the parabola $y = x^2$. Again, this process is only based on the region — it doesn't depend on f.

To flesh things out, let's pick an example function

$$f(x, y) = 2x + 4y$$

as Poonen did, and show how we would find the integral.

> **Sample Question**
>
> Consider the region $\mathcal{R}$ we just described, the set of points between bounded between $y - x = 2$ and $y = x^2$. Integrate $\iint_{\mathcal{R}} (2x + 4y) \, \mathrm{d}x \, \mathrm{d}y$ over this region.

Solution. As we saw, there are two different ways to set it up. We'll do the one that's nice (and show the worse one afterwards for comparison), where we have x on the outside.

We are given the integral

$$\int_{x=-1}^{2} \int_{y=x^2}^{x+2} (2x + 4y) \, \mathrm{d}y \, \mathrm{d}x.$$

1. The first step is to compute the inner integral with respect to y, treating x as a constant.

   The inner integral is:

   $$\int_{y=x^2}^{x+2} (2x + 4y) \, \mathrm{d}y.$$

   We can split this integral into two parts:

   $$\int_{y=x^2}^{x+2} 2x \, \mathrm{d}y + \int_{y=x^2}^{x+2} 4y \, \mathrm{d}y.$$

   - The first term is:

     $$2x \int_{y=x^2}^{x+2} 1 \, \mathrm{d}y = 2x[y]_{y=x^2}^{y=x+2} = 2x((x + 2) - x^2).$$

   - The second term is:

     $$4 \int_{y=x^2}^{x+2} y \, \mathrm{d}y = 4 \left[\frac{y^2}{2} \right]_{y=x^2}^{y=x+2} = 4 \left(\frac{(x + 2)^2}{2} - \frac{(x^2)^2}{2} \right) = 2(x^2 + 4x + 4 - x^4).$$

   Thus, the inner integral is:

   $$2x(x + 2 - x^2) + 2(x^2 + 4x + 4 - x^4) = -2x^4 - 2x^3 + 4x^2 + 12x + 8.$$

2. Now, we compute the outer integral:

$$\begin{aligned}
&\int_{x=-1}^{2} \left(-2x^4 - 2x^3 + 4x^2 + 12x + 8\right) \mathrm{d}x \\
&= \left[-2\frac{x^5}{5} - 2 \cdot \frac{x^4}{4} + 4\frac{x^3}{3} + 12 \cdot \frac{x^2}{2} + 8x \right]_{x=-1}^{2}.
\end{aligned}$$

This is a lot of arithmetic, sorry. One way is to work term by term:

$$\begin{aligned}
-2\left[\frac{x^5}{5}\right]_{x=-1}^{x=2} &= -2\left(\frac{32}{5} - \frac{(-1)^5}{5}\right) = -2 \cdot \frac{33}{5} = -\frac{66}{5} \\
-2\left[\frac{x^4}{4}\right]_{x=-1}^{x=2} &= -2\left(\frac{16}{4} - \frac{1}{4}\right) = -2 \cdot \frac{15}{4} = -\frac{15}{2} \\
4\left[\frac{x^3}{3}\right]_{x=-1}^{x=2} &= 4\left(\frac{8}{3} - \frac{(-1)^3}{3}\right) = 4 \cdot \frac{9}{3} = 12 \\
12\left[\frac{x^2}{2}\right]_{x=-1}^{x=2} &= 12 \cdot \frac{3}{2} = 18 \\
8 \cdot (2 - (-1)) &= 8 \cdot 3 = 24.
\end{aligned}$$

Add these to get the answer:

$$-\frac{66}{5} - \frac{15}{2} + 12 + 18 + 24 = \boxed{\frac{333}{10}}.$$

□

§22.4 [SIDENOTE] What it looks like if you integrate the hard way

In the previous sample question, we picked x to be the outer integral so that we didn't have to do cases or deal with square roots. This was pretty clearly a good choice.

For contrast, I'll show you what happens if you have y in the outer integral instead — just to make a point that things can get ugly. (You can read it if you want the practice with iterated integrals, or skip it if you believe me.) To reiterate, we will directly calculate

$$\int_{y=0}^{1} \int_{x=-\sqrt{y}}^{\sqrt{y}} (2x + 4y) \,\mathrm{d}x \,\mathrm{d}y + \int_{y=1}^{4} \int_{x=y-2}^{\sqrt{y}} (2x + 4y) \,\mathrm{d}x \,\mathrm{d}y.$$

- We calculate the first hunk

$$\int_{y=0}^{1} \int_{x=-\sqrt{y}}^{\sqrt{y}} (2x + 4y) \,\mathrm{d}x \,\mathrm{d}y.$$

1. The first step is to compute the inner integral with respect to x, treating y as a constant. The inner integral is:

$$\int_{x=-\sqrt{y}}^{\sqrt{y}} (2x + 4y) \,\mathrm{d}x.$$

We can split this into two integrals:

$$\int_{x=-\sqrt{y}}^{\sqrt{y}} 2x \, \mathrm{d}x + \int_{x=-\sqrt{y}}^{\sqrt{y}} 4y \, \mathrm{d}x.$$

- The first term is:

$$2\int_{x=-\sqrt{y}}^{\sqrt{y}} x \, \mathrm{d}x = 2\left[\frac{x^2}{2}\right]_{x=-\sqrt{y}}^{x=\sqrt{y}} = 2 \cdot \left(\frac{\left(\sqrt{y}\right)^2}{2} - \frac{\left(-\sqrt{y}\right)^2}{2}\right).$$

- The second term is:

$$4y\int_{x=-\sqrt{y}}^{\sqrt{y}} 1 \, \mathrm{d}x = 4y[x]_{x=-\sqrt{y}}^{x=\sqrt{y}} = 4y(\sqrt{y} - (-\sqrt{y})) = 4y \cdot 2\sqrt{y} = 8y^{\frac{3}{2}}.$$

Thus, the inner integral is:

$$0 + 8y^{\frac{3}{2}} = 8y^{\frac{3}{2}}.$$

2. Now, we compute the outer integral:

$$\int_{y=0}^{1} 8y^{3/2} \, \mathrm{d}y.$$

We use the power rule for integration:

$$\int y^{3/2} \, \mathrm{d}y = \frac{\frac{y^{5/2}}{5}}{2} = \frac{2}{5}y^{5/2}.$$

Thus, the outer integral becomes:

$$8\int_{y=0}^{1} y^{\frac{3}{2}} \, \mathrm{d}y = 8 \cdot \frac{2}{5}\left[y^{\frac{5}{2}}\right]_{y=0}^{y=1} = 8 \cdot \frac{2}{5} \cdot \left(1^{\frac{5}{2}} - 0^{\frac{5}{2}}\right) = 8 \cdot \frac{2}{5} = \frac{16}{5}.$$

Hence the first hunk is

$$\int_{y=0}^{1}\int_{x=-\sqrt{y}}^{\sqrt{y}} (2x + 4y) \, \mathrm{d}x \, \mathrm{d}y = \frac{16}{5} = 3.2.$$

- We calculate the second hunk

$$\int_{y=0}^{1}\int_{x=y-2}^{\sqrt{y}} (2x + 4y) \, \mathrm{d}x \, \mathrm{d}y.$$

1. The first step is to compute the inner integral with respect to x, treating y as a constant. The inner integral is:

$$\int_{x=y-2}^{\sqrt{y}} (2x + 4y) \, \mathrm{d}x.$$

We can split this into two integrals:

$$\int_{x=y-2}^{\sqrt{y}} 2x \,\mathrm{d}x + \int_{x=y-2}^{\sqrt{y}} 4y \,\mathrm{d}x.$$

- The first term is:

$$\int_{x=y-2}^{\sqrt{y}} 2x \,\mathrm{d}x = 2\left[\frac{x^2}{2}\right]_{x=y-2}^{x=\sqrt{y}} = \left((\sqrt{y})^2 - (y-2)^2\right).$$

  Simplifying:

$$\left(y - (y^2 - 4y + 4)\right) = y - \left(y^2 - 4y + 4\right) = y - y^2 + 4y - 4 = -y^2 + 5y - 4.$$

- The second term is:

$$4y \int_{x=y-2}^{\sqrt{y}} 1 \,\mathrm{d}x = 4y(\sqrt{y} - (y-2)) = 4y(\sqrt{y} - y + 2) = 4y(\sqrt{y} - y + 2).$$

  Thus, the inner integral is:

$$\left(-y^2 + 5y - 4\right) + 4y(\sqrt{y} - y + 2) = -y^2 + 5y - 4 + 4y\sqrt{y} - 4y^2 + 8y.$$

Simplifying we get the inner integral to be

$$-5y^2 + 13y + 4y\sqrt{y} - 4.$$

2. Now, we compute the outer integral:

$$\int_{y=1}^{4} \left(-5y^2 + 13y + 4y\sqrt{y} - 4\right) \mathrm{d}y.$$

To keep things organized, we integrate each term individually:

$$\int_{y=1}^{4} -5y^2 \,\mathrm{d}y = -5\left[\frac{y^3}{3}\right]_{y=1}^{y=4} = -5 \cdot \left(\frac{64}{3} - \frac{1}{3}\right) = -5 \cdot \frac{63}{3} = -105$$

$$\int_{y=1}^{4} 13y \,\mathrm{d}y = 13\left[\frac{y^2}{2}\right]_{y=1}^{y=4} = 13 \cdot \left(\frac{16}{2} - \frac{1}{2}\right) = 13 \cdot \frac{15}{2} = 97.5$$

$$\int_{y=1}^{4} 4y\sqrt{y} \,\mathrm{d}y = 4\int_{y=1}^{4} y^{3/2} \,\mathrm{d}y = 4 \cdot \left[\frac{2}{5}y^{5/2}\right]_{y=1}^{y=4} = 4 \cdot \frac{2}{5}(32 - 1) = \frac{248}{5} = 49.6$$

$$\int_{y=1}^{4} -4 \,\mathrm{d}y = -4[y]_{y=1}^{y=4} = -4(4-1) = -12.$$

Now, add up the integrals:

$$-105 + 97.5 + 49.6 - 12 = 30.1.$$

- The final answer is $3.2 + 30.1 = 33.3$ as expected.

So we got the same answer, no surprise, but it took a lot more work to get it.

§22.5 [TEXT] A few physical interpretations of integrals

Depending on what function f is chosen, the integral may have some physical meaning. We give a few examples here:

§22.5.1 Area

If you choose $f = 1$ you get area.

Recipe for area

To find the area of a region $\mathcal{R}$, use

$$\operatorname{Area}(\mathcal{R}) = \iint_{\mathcal{R}} 1 \, dx \, dy.$$

Digression: This is the definition of area

Sometimes people ask me why we choose to integrate 1 as opposed to some other function. The answer might be a bit surprising: you can actually take the above integral as the *definition* of area. (If you think back carefully to what you learned in high school, you might realize that nobody actually ever gave you a precise definition of the word "area", and that was for good reason.)

Tip

You can and will use the recipe the other way too: suppose you're doing some problem and you end up with $\iint_{\mathcal{R}} dx \, dy$ where $\mathcal{R}$ is the circle $x^2 + y^2 \le 1$. Don't go through the trouble of actually calculating the integral: it's the area of a circle with radius 1, which is just π!

Sample Question

Consider the region $\mathcal{R}$ we just described, the set of points between bounded between $y - x = 2$ and $y = x^2$. Compute its area.

Solution. We'll write this as

$$\int_{x=-1}^{2} \int_{y=x^2}^{x+2} 1 \, dy \, dx.$$

The inner integral is easy $\int_{y=x^2}^{x+2} dy = (x+2) - x^2$. So the answer is

$$\int_{x=-1}^{2} (x + 2 - x^2) \, dx = \left[\frac{x^2}{2} + 2x - \frac{x^3}{3} \right]_{x=-1}^{x=2} = \left(2 + 4 - \frac{8}{3}\right) - \left(\frac{1}{2} - 2 + \frac{1}{3}\right) = \boxed{\frac{9}{2}}. \qquad \square$$

Sample Question

Compute the area of the region $\mathcal{R}$ where $0 \le x \le 10$ and $0 \le y \le x^2$.

Solution. Write

$$\int_{x=0}^{10} \int_{y=0}^{x^2} 1 \, \mathrm{d}y \, \mathrm{d}x = \int_{x=0}^{10} x^2 \, \mathrm{d}x = \left[\frac{x^3}{3}\right]_{x=0}^{10} = \boxed{\frac{1000}{3}}. \qquad \square$$

Actually this is just a rephrasing of the "area under the curve" you learned in 18.01, when you would write $\int_{x=0}^{10} x^2 \, \mathrm{d}x = \left[\frac{x^3}{3}\right]_{x=0}^{10} = \frac{1000}{3}$ and were told "this is the area under the curve $y = x^2$", as in Figure 45. But the 18.02 definition is more versatile, because it lets us give a definition of area for *any* integrable region in the xy-plane, not just those under a curve of the form $y = f(x)$.

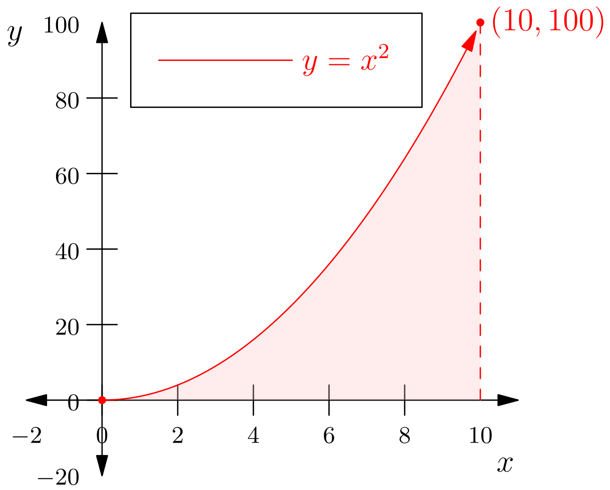

Figure 45: The area $\int_{x=0}^{10} \int_{y=0}^{x^2} 1 \, \mathrm{d}y \, \mathrm{d}x$ in 18.02 language matches what you expect from the 18.01 integral $\int_{x=0}^{10} x^2 \, \mathrm{d}x$.

§22.5.2 Mass and center of mass

If you imagine your region $\mathcal{R}$ as a blob of some substance (concrete, wood, water, etc.), then you could also imagine it has a *density* at each point in the region (say, in kilograms per square meter). In 18.02 we usually denote the density by ρ, which is a function taking each point P in the region $\mathcal{R}$ and outputting its density.

In that case, the total mass of $\mathcal{R}$ is the integral of the densities:

$$\operatorname{mass}(\mathcal{R}) = \iint_{\mathcal{R}} \rho(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

Given a region you can also consider the *center of mass*. The idea/definition is that the x-coordinate of the center of mass should be the weighted average of the x-coordinates of the points in the region, and is usually denoted $\bar{x}$:

$$\bar{x} := x\text{-coord of the center of mass} = \frac{1}{\operatorname{mass}(\mathcal{R})} \iint_{\mathcal{R}} x \cdot \rho(x, y) \, \mathrm{d}x \, \mathrm{d}y.$$

And the same for the others. Let's repeat this in recipe form.[21]

[21]It took considerable self-restraint to not title the recipe "Mass Tech".

Recipe for total mass and center of mass

Suppose $\mathcal{R}$ is a region and ρ is a density function for the region.

1. The total mass is given by $\operatorname{mass}(\mathcal{R}) = \iint_{\mathcal{R}} \rho(x,y)\,dx\,dy$.
2. The center of mass is the point $(\bar{x}, \bar{y})$ defined by

$$(\bar{x}, \bar{y}) := \left(\frac{\iint_{\mathcal{R}} x \cdot \rho(x,y)\,dx\,dy}{\operatorname{mass}(\mathcal{R})}, \frac{\iint_{\mathcal{R}} y \cdot \rho(x,y)\,dx\,dy}{\operatorname{mass}(\mathcal{R})} \right).$$

Type signature

If $\mathcal{R}$ is a region in $\mathbb{R}^2$,

- Then a density function $\rho : \mathcal{R} \to \mathbb{R}_{\geq 0}$ should take on nonnegative values. (For physicists: in SI units, you might imagine it as kilograms per square meter.)
- The mass is a nonnegative real number (kilograms).
- The center of mass is also a *point* inside $\mathcal{R}$. (Draw this as a dot, not an arrow.)

Sample Question

Compute the center of mass of the square with vertices $(5,5)$, $(5,9)$, $(9,9)$ and $(9,5)$, assuming a constant density $\rho = 1$.

Solution. Of course, by symmetry we expect the answer to be $\boxed{(7,7)}$. Let's see this in full. The mass of $\mathcal{R}$ is given by

$$\operatorname{mass}(\mathcal{R}) = \iint_{\mathcal{R}} 1\,dx\,dy = \int_{x=5}^{9} \int_{y=5}^{9} 1\,dy\,dx = (9-5) \cdot (9-5) = 16.$$

The x-coordinate of the center of mass is

$$\begin{aligned} \bar{x} &= \frac{1}{\operatorname{mass}(\mathcal{R})} \iint_{\mathcal{R}} x \cdot 1\,dx\,dy = \frac{1}{16} \int_{x=5}^{9} \int_{y=5}^{9} x\,dy\,dx \\ &= \frac{1}{16} \int_{y=5}^{9} [x \cdot (9-5)]\,dx = \frac{1}{16} \cdot 4 \int_{x=5}^{9} x\,dx = \frac{1}{4} \left[\frac{x^2}{2} \right]_{x=5}^{9} = \frac{1}{4} \left(\frac{81}{2} - \frac{25}{2} \right) = 7. \end{aligned}$$

The calculation for $\bar{y}$ is exactly the same, and we get $(7,7)$ as we hoped. □

Remark

Unsurprisingly if $\rho = 1$ is constant (imagine 1 kilogram per square meter), then the mass of the region $\mathcal{R}$ is just $\iint_{\mathcal{R}} dx\,dy$, i.e. the area. (So a region whose area is 16 square meters and where the density is 1 kilogram per square meter in the whole substance should be 16 kilograms.)

Sample Question

Compute the center of mass of the square with vertices $(5,5)$, $(5,9)$, $(9,9)$ and $(9,5)$, assuming a density function of $\rho(x,y) = x+y$.

Solution. First compute the mass:

$$\begin{aligned}
\operatorname{mass}(\mathcal{R}) &= \iint_{\mathcal{R}} (x+y)\,\mathrm{d}x\,\mathrm{d}y = \int_{x=5}^{9}\int_{y=5}^{9} (x+y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=5}^{9} \left[xy + \frac{y^2}{2}\right]_{y=5}^{9} \mathrm{d}x = \int_{x=5}^{9} \left(x(9-5) + \frac{81}{2} - \frac{25}{2}\right)\mathrm{d}x = \int_{x=5}^{9} (4x+28)\,\mathrm{d}x \\
&= 4\left[\frac{x^2}{2}\right]_{x=5}^{9} + 28[x]_{x=5}^{9} = 2(81-25) + 28\cdot 4 = 224.
\end{aligned}$$

Then the x-coordinate of the center of mass is

$$\begin{aligned}
\bar{x} &= \frac{1}{\operatorname{mass}(\mathcal{R})} \iint_{\mathcal{R}} x(x+y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{224}\int_{x=5}^{9}\int_{y=5}^{9} (x^2+xy)\,\mathrm{d}y\,\mathrm{d}x \\
&= \frac{1}{224}\int_{x=5}^{9} \left[x^2y + \frac{xy^2}{2}\right]_{y=5}^{9} \mathrm{d}x = \frac{1}{224}\int_{x=5}^{9}\left(x^2(9-5) + \frac{x(81-25)}{2}\right)\mathrm{d}x \\
&= \frac{1}{224}\int_{x=5}^{9} (4x^2+28x)\,\mathrm{d}x = \frac{1}{224}\left[\frac{4x^3}{3} + 14x^2\right]_{x=5}^{9} \\
&= \frac{1}{224}\left(\frac{4(729)-4(125)}{3} + 14(81-25)\right) = \frac{149}{21}.
\end{aligned}$$

And $\bar{y} = \frac{149}{21}$ in exactly the same way. Hence the answer $\boxed{\left(\frac{149}{21}, \frac{149}{21}\right) \approx (7.095, 7.095)}$.

(This passes a sanity check: our new square is a bit denser near $(9,9)$ than $(5,5)$. So we expect the center of mass to move in that direction slightly. We still have symmetry across the line $y=x$.) $\square$

§22.6 [SIDENOTE] What's the analogy to "area under the curve" from 18.01?

In 18.01, you were told that the integral $\int_{x=a}^{b} f(x)\,\mathrm{d}x$ denotes the area under the curve $y=f(x)$ from $x=a$ to $x=b$.

In 18.02, if you have $\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$, and you want to interpret it analogously, what you would do is look at the surface $z=f(x,y)$ in an xyz-plane, where you imagine the xy-plane and the region $\mathcal{R}$ at the bottom, and z being a height. Then the double integral analogously calculates the volume underneath the surface.

However, we won't actually use this interpretation much in 18.02. As I said before, in 18.02 we usually prefer to draw pictures where all the axis variables are treated with equal respect. (Whereas the 18.01 picture I just mentioned uses x as input and y as output; the two axes don't play the same role.) So picturing the double integral with things like mass or center of mass is more in line with the 18.02 spirit, even though there is no 18.01 analog.

§22.7 [RECIPE] Swapping the order of integration

If you're an instructor teaching multivariable calculus, one trope for generating exam questions is to take some region $\mathcal{R}$ that can be sliced both horizontally and vertically, but for which one way is much easier to integrate than the other. For the problem statement, you give the student the integral written in the "bad" order. The solution is to convert *back* into a region $\mathcal{R}$, and then use this to recover the "good" order. Put in recipe form:

Recipe for swapping the order of integration

If you are given $\int_{x=?}^{?} \int_{y=?}^{?} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$ and you wish to switch the order of integration the other way:

1. Convert the limits of integration back into inequality/region format, getting some region $\mathcal{R}$.
2. Re-apply the recipe from Section 22.2 using the other variable as the outer one now.

Sample Question

Evaluate the double integral:

$$\int_{x=0}^{2} \int_{y=\frac{x}{2}}^{1} e^{y^2}\,\mathrm{d}y\,\mathrm{d}x.$$

Solution. To evaluate this integral, note that integrating e^{y^2} directly with respect to y is not feasible using standard methods from 18.01. Thus, we need to swap the order of integration.

First convert this back into region format:

$$\mathcal{R} = \begin{cases} 0 \le x \le 2 \\ \frac{x}{2} \le y \le 1 \end{cases}.$$

We see that y goes in the range $0 \le y \le 1$. The region being integrated is drawn in Figure 46.

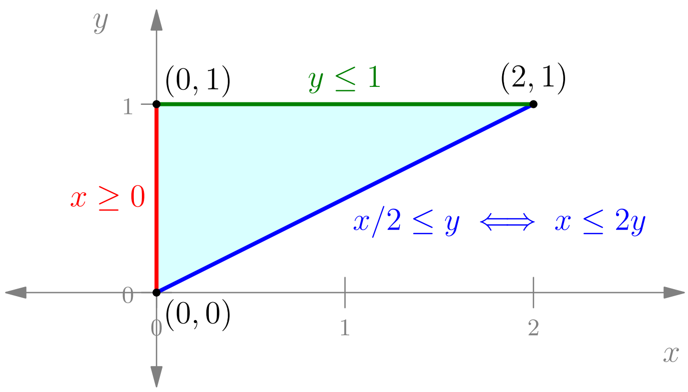

Figure 46: The region $0 \le x \le 2$ and $\frac{x}{2} \le y \le 1$. Note that the function you are integrating, in this case e^{y^2}, is irrelevant to the region being integrated over!

Solving for x in terms of y gives three conditions: in addition to $0 \le x \le 2$ we need $x \le 2y$. Since $y \le 1$, we can ignore the condition $x \le 2$, and the region can be rewritten to

$$\mathcal{R} = \begin{cases} 0 \le y \le 1 \\ 0 \le x \le 2y \end{cases}.$$

Turning this *back* into a double integral gives

$$\int_{y=0}^{1} \int_{x=0}^{2y} e^{y^2} \,\mathrm{d}y \,\mathrm{d}x.$$

The inner integral is with respect to x, but the integrand e^{y^2} is independent of x. Therefore, the inner integral becomes:

$$\int_{x=0}^{2y} e^{y^2} \,\mathrm{d}x = 2y \cdot e^{y^2}.$$

Thus, it remains to calculate

$$\int_{y=0}^{1} \left(2y \cdot e^{y^2}\right) \mathrm{d}y \,\mathrm{d}x.$$

And now things are different: $2y \cdot e^{y^2}$ *does* have a valid anti-derivative. If you use the 18.01 method or even just are good at guessing, you can find the indefinite 18.01 integral

$$\int 2ye^{y^2} \,\mathrm{d}y = e^{y^2} + C.$$

So the final answer to the problem is

$$\int_{y=0}^{1} 2ye^{y^2} \,\mathrm{d}y \,\mathrm{d}x = \left[e^{y^2}\right]_{y=0}^{y=1} = \boxed{e-1}.$$

□

Sample Question

Let

$$k = \sqrt[5]{\frac{37}{3}\pi} \approx 2.078.$$

Evaluate the double integral:

$$\int_{y=0}^{k^2} \int_{x=\sqrt{y}}^{k} y \sin(x^5) \,\mathrm{d}x \,\mathrm{d}y$$

Solution. Integrating $\sin(x^5)$ is not reasonable, so we swap the order of integration and pray. The region being integrated is

$$\mathcal{R} = \begin{cases} 0 \le y \le k^2 \\ \sqrt{y} \le x \le k \end{cases}.$$

The values of x range all the way from 0 to k. We draw the region in Figure 47.

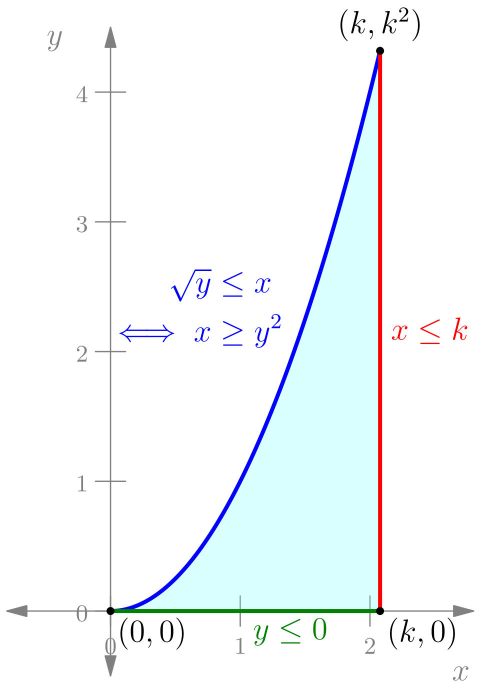

Figure 47: The region $0 \le y \le k^2$ and $\sqrt{y} \le x \le k$.

Solving for y, we see that we have three constraints, $0 \le y$, $y \le x^2$ and $y \le k^2$. But since $x \le k$, the condition $y \le k^2$ is redundant. The region can be rewritten as simply

$$\mathcal{R} = \begin{cases} 0 \le x \le k \\ 0 \le y \le x^2 \end{cases}.$$

Convert back into a double integral:

$$\int_{x=0}^{k} \int_{y=0}^{x^2} y \sin(x^5) \, \mathrm{d}y \, \mathrm{d}x.$$

We now compute the inner integral with respect to y:

$$\int_{y=0}^{x^2} y \sin(x^5) \, \mathrm{d}y.$$

Since $\sin(x^5)$ is independent of y, we can factor it out of the integral:

$$\sin(x^5) \int_{y=0}^{x^2} y \, \mathrm{d}y = \sin(x^5) \left[\frac{y^2}{2} \right]_{y=0}^{y=x^2}.$$

Substituting the limits of integration:

$$\sin(x^5) \cdot \frac{x^4}{2}.$$

Now substitute this result into the outer integral:

$$\int_{x=0}^{k} \frac{x^4}{2} \sin(x^5) \, \mathrm{d}x.$$

Let's perform the 18.01 u-substitution $u = x^5$, so $\mathrm{d}u = 5x^4\,\mathrm{d}x$, or $\mathrm{d}x = \frac{\mathrm{d}u}{5x^4}$. The limits of integration change as follows:

- When $x = 0$, $u = 0$.
- When $x = k$, $u = \frac{37}{3}\pi$.

Thus, knowing that $\int \sin(u) = -\cos(u) + C$, the integral becomes:

$$\begin{aligned} \frac{1}{2}\int_{u=0}^{\frac{37}{3}\pi} \frac{\sin(u)}{5}\,\mathrm{d}u &= \frac{1}{10}\int_{u=0}^{\frac{37}{3}\pi} \sin(u)\,\mathrm{d}u \\ &= \frac{1}{10}\left(-\cos\left(\frac{37}{3}\pi\right) + \cos(0)\right). \end{aligned}$$

Using $\cos\left(\frac{37}{3}\pi\right) = \frac{1}{2}$ and $\cos(0) = 1$, we get:

$$\frac{1}{10}\left(-\frac{1}{2} + 1\right) = \frac{1}{10} \cdot \frac{1}{2} = \boxed{\frac{1}{20}}. \qquad \square$$

§22.8 [EXER] Exercises

Exercise 22.1. Let $\mathcal{R}$ be the region between the curves $y = \sqrt{x}$ and $y = x^3$. Compute $\iint_{\mathcal{R}} x^{100} y^{200}\,\mathrm{d}x\,\mathrm{d}y$ using both horizontal and vertical slicing.

Exercise 22.2. Let $\mathcal{R}$ be the region between the curves $y = \sqrt{x}$ and $y = x^2$. Assume $\mathcal{R}$ has constant density. Calculate its center of mass.

Exercise 22.3. Evaluate the double integral:

$$\int_{y=0}^{1} \int_{x=y}^{\sqrt[5]{y}} \frac{xy^2}{1 - x^{12}}\,\mathrm{d}x\,\mathrm{d}y.$$

Exercise 22.4 (*). Prove that

$$\int_{x=0}^{999^5} \sqrt[3]{\sqrt[5]{x} + 1}$$

is a rational number.

Chapter 23. Change of variables

We'll do just two variables for now; the 3D situation is exactly the same and we cover it later.

§23.1 [TEXT] Interval notation

One quick notational thing if you haven't seen this before:

Definition: Interval notation

Suppose $[a, b]$ and $[c, d]$ are closed intervals in $\mathbb{R}$ (so $a \le b$ and $c \le d$). By $[a, b] \times [c, d]$ we mean the rectangle consisting of points (x, y) such that $a \le x \le b$ and $c \le y \le d$. (So the four corners of the rectangle are (a, c), (a, d), (b, c), (b, d).)

Example

For example $[0, 1] \times [0, 1]$ would be a unit square whose southwest corner is at the origin. Similarly, $[0, 5] \times [0, 3]$ would be a rectangle of width 5 and height 3.

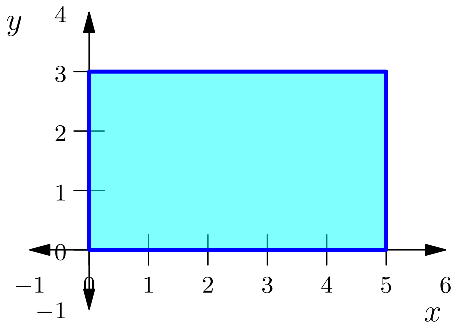

Figure 48: A picture of $[0, 5] \times [0, 3]$. This is just the set of points where x is in the interval $[0, 5]$ and y is in the closed interval $[0, 3]$.

§23.2 [TEXT] Transition maps

As it turns out xy-integration (or xyz-integration in 3D) isn't always going to be nice, even if you try both horizontal and vertical slicing.

The standard example that's given looks something like this: suppose you want to integrate the region bounded by the four lines

$$xy = \frac{16}{9}, \quad xy = \frac{16}{25}, \quad x = 4y, \quad y = 4x.$$

This region is sketched in Figure 49. As I promised you, I think it's better for your thinking if you write these as inequalities:

$$\begin{aligned} \frac{16}{25} &\le xy \le \frac{16}{9} \\ \frac{1}{4} &\le \frac{y}{x} \le 4. \end{aligned} \tag{11}$$

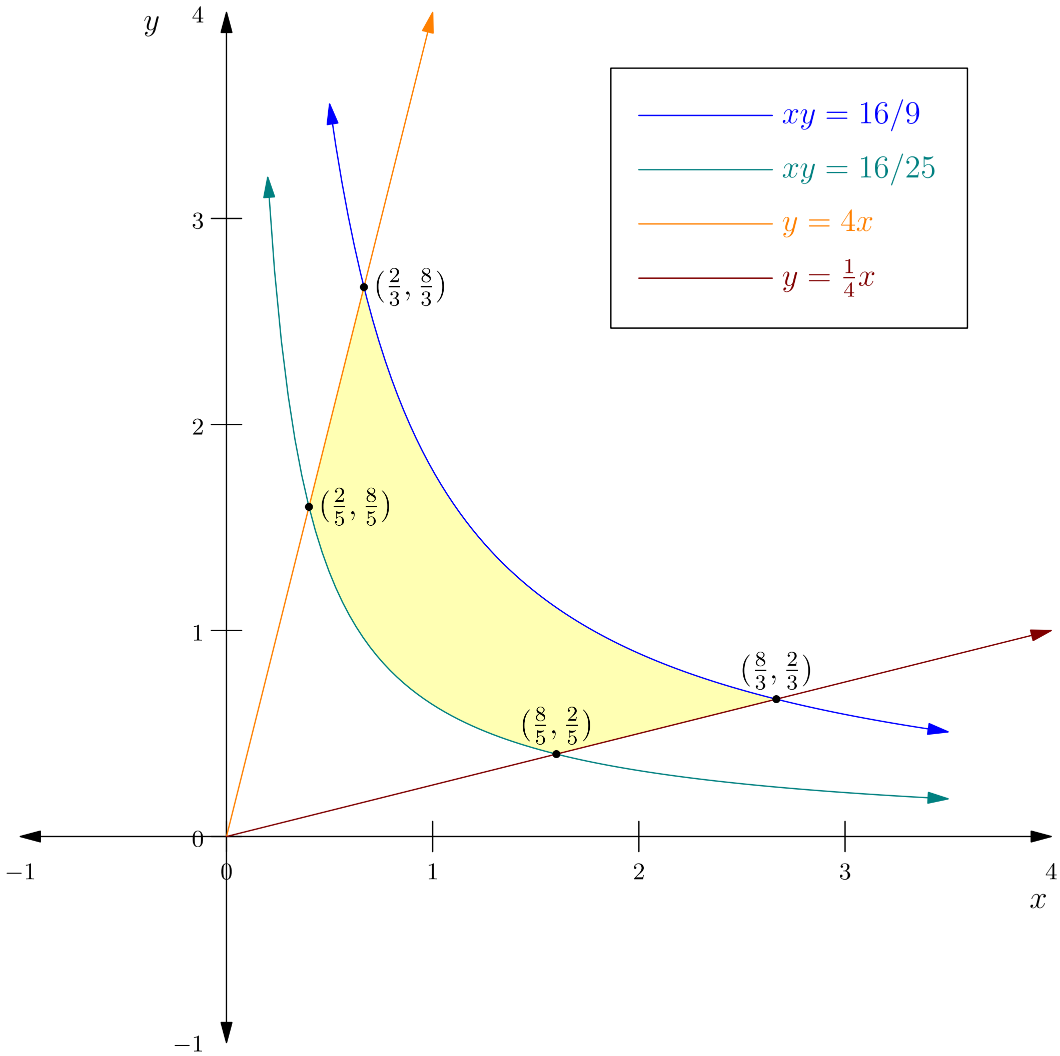

Figure 49: A cursed region bounded by four curves. Trying to do xy-integration in either direction will be annoying as heck.

This chapter introduces a technique called "change of variables" that will allow us to handle this annoying-looking yellow region for when we don't want to do xy-integration. The idea is to make a new map of the yellow region with a different coordinate system. To do this, I need to tell you a new term:

Definition of transition map

Suppose $\mathcal{R}$ is a region. Let $\mathcal{R}_{\text{new}}$ be another region, often a rectangle. A **transition map** for $\mathcal{R}$ is a function $\mathbf{T} : \mathcal{R}_{\text{new}} \to \mathcal{R}$ that transforms $\mathcal{R}_{\text{new}}$ to $\mathcal{R}$.

In 18.02 we always require that all the points except possibly the boundaries of $\mathcal{R}_{\text{new}}$ get mapped to different points in $\mathcal{R}$. Thus, writing the inverse $\mathbf{T}^{-1}$ usually also makes sense.

If $\mathcal{R}_{\text{new}}$ is a rectangle — and again, that's quite common — then sometimes $\mathbf{T}$ is also called a cell (e.g. my Napkin does this when discussing differential forms).

Remark: An analogy to the world map

Cartography or geography enthusiasts will find that the word "map" gives them the right instincts. If you print a world map on an 8.5×11 or A4 sheet of paper, it gives you a coordinate system for the world with longitude and latitude. So $\mathcal{R}$ can be thought of as the surface of the Earth, while $\mathcal{R}_{\text{new}}$ is the rectangular sheet of paper. (I'm lying a little bit because the Earth lives in 3D space but not 2D space, but bear with me.)

The map is always distorted in some places, because the Earth is bent: the north and south pole will often get stretched a ton, for example. But that's okay — **as long as each longitude and latitude gives you a different point on Earth, we're satisfied**. Technically there are exceptions at the north and south poles, but those are on the boundary and we let it slide.

This corresponds to the idea that a cell can capture a complicated area with two coordinates. See Figure 50.

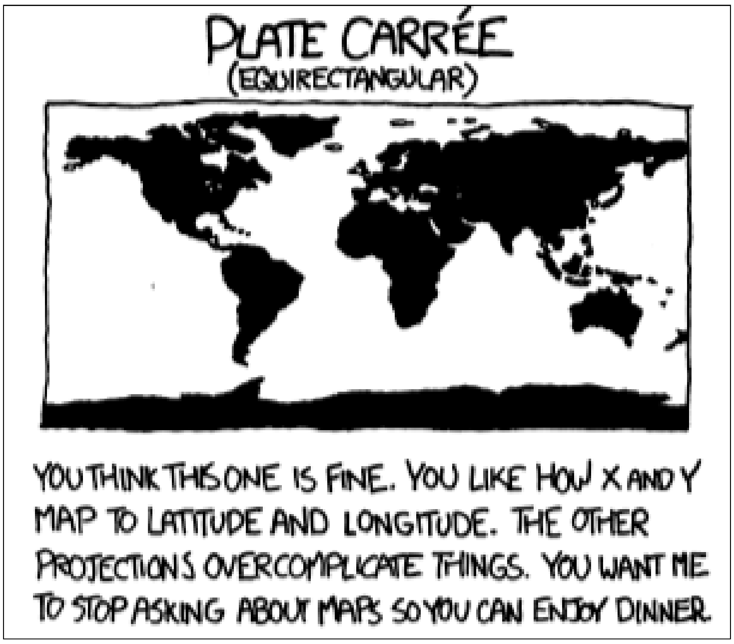

Figure 50: One of the map projections from XKCD 977, a chart titled *What your favorite map projection says about you*. There's several more if you're curious.

Now how does a transition map help us? Well, first, let's show how we can do cartography on the region we just saw, and then worry about the integration later.

The key idea is that we need to rig our transition function such that

$$u = \frac{y}{x}, \qquad v = xy$$

so that our two inequalities we saw earlier are just

$$\frac{1}{4} \le v \le 4, \qquad \frac{16}{25} \le v \le \frac{16}{9}.$$

This lets us make a portrayal of the yellow region as a sheet of paper. See Figure 51, which is really important to us!

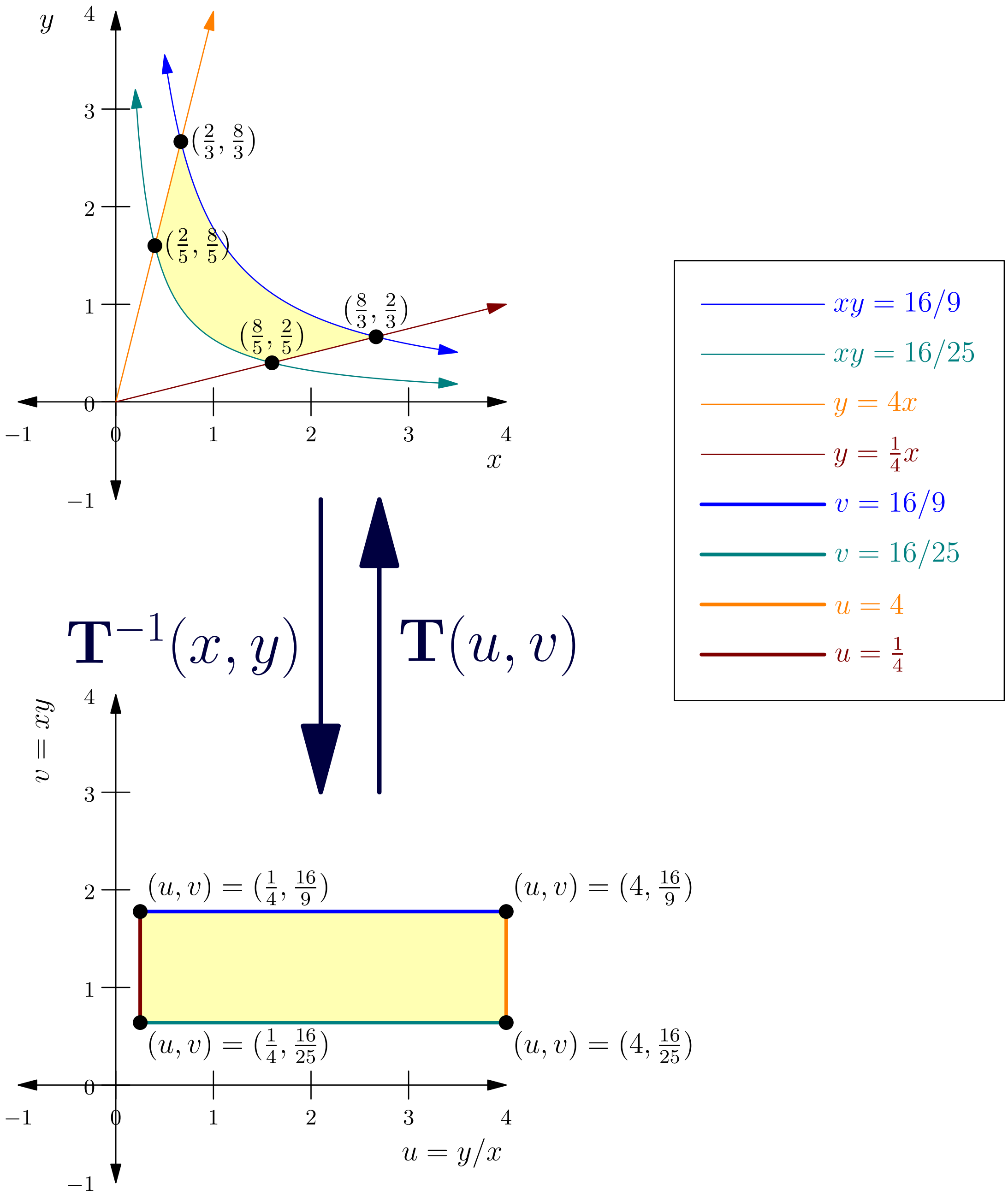

Figure 51: We use a rectangle with (u, v) as our new region $\mathcal{R}_{\text{new}}$. The transition map lets us do cartography on the region $\mathcal{R}$.

How do we actually express the transition map of $\mathbf{T}$? It's actually easier to write the *inverse*; in this context it's actually more natural to write

$$\mathbf{T}^{-1}(x, y) = \left(\frac{y}{x}, xy\right).$$

If you really need $\mathbf{T}$ itself, you would solve for u and v in terms of x and y to get

$$x = \sqrt{\frac{v}{u}}$$
$$y = \sqrt{uv}$$

so that our transition map is given exactly by

$$\mathbf{T}(u,v) = \left(\sqrt{\frac{v}{u}}, \sqrt{uv}\right).$$

However, actually for integration purposes (we'll see this next section) you can actually get away with only the formula for $\mathbf{T}^{-1}$ instead.

§23.3 [TEXT] Integration once you have a transition map

If you remember change-of-variables from 18.01, the 18.02 is the grown-up version where you have a transition map instead.

Definition: Jacobian determinant

Let $\mathbf{T}$ be a transition map defined from a region in $\mathbb{R}^n$ to $\mathbb{R}^n$. The **Jacobian matrix** is the matrix whose rows are the gradients of each component written as row vectors; the **Jacobian determinant** is its determinant. In these notes we denote the matrix by $J_{\mathbf{T}}$ (and the determinant by $\det J_{\mathbf{T}}$).

For example in a 2×2 case, if the transition map $\mathbf{T}(u,v)$ is written as $\mathbf{T}(u,v) = (p(u,v), q(u,v))$, then

$$J_{\mathbf{T}} = \begin{pmatrix} \frac{\partial p}{\partial u} & \frac{\partial p}{\partial v} \\ \frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} \end{pmatrix}.$$

Example

Let's consider the transition map $\mathbf{T}(u,v)$ we saw earlier, that is

$$\mathbf{T}(u,v) = \left(\sqrt{\frac{v}{u}}, \sqrt{uv}\right).$$

We compute the gradient of $(u,v) \mapsto \sqrt{\frac{v}{u}}$ by taking the two partials:

$$\frac{\partial}{\partial u}\sqrt{\frac{v}{u}} = -\frac{1}{2}u^{-\frac{3}{2}}v^{\frac{1}{2}}$$
$$\frac{\partial}{\partial v}\sqrt{\frac{v}{u}} = \frac{1}{2}u^{-\frac{1}{2}}v^{-\frac{1}{2}}.$$

The other component $(u,v) \mapsto \sqrt{uv}$ has the following gradient:

$$\frac{\partial}{\partial u}\sqrt{uv} = \frac{1}{2}u^{-\frac{1}{2}}v^{\frac{1}{2}}$$
$$\frac{\partial}{\partial v}\sqrt{uv} = \frac{1}{2}u^{\frac{1}{2}}v^{-\frac{1}{2}}.$$

So the Jacobian matrix for $\mathbf{T}$ is

$$J_{\mathbf{T}} = \begin{pmatrix} \frac{1}{2}u^{-\frac{3}{2}}v^{\frac{1}{2}} & \frac{1}{2}u^{-\frac{1}{2}}v^{-\frac{1}{2}} \\ \frac{1}{2}u^{-\frac{1}{2}}v^{\frac{1}{2}} & \frac{1}{2}u^{\frac{1}{2}}v^{-\frac{1}{2}} \end{pmatrix}.$$

Example

We can also find the Jacobian matrix of the *inverse* map too, that is, the transition map $\mathbf{T}^{-1} : \mathcal{R} \to \mathcal{R}_{\text{new}}$ defined by

$$\mathbf{T}^{-1}(x, y) = \left(\frac{y}{x}, xy\right).$$

In other words, this is the map that transforms (x, y) into (u, v). This is actually less painful because you don't have to deal with the square roots everywhere.

$$\begin{aligned} \frac{\partial}{\partial x}\left(\frac{y}{x}\right) &= -\frac{y}{x^2}, & \frac{\partial}{\partial y}\left(\frac{y}{x}\right) &= \frac{1}{x} \\ \frac{\partial}{\partial x}(xy) &= y, & \frac{\partial}{\partial y}(xy) &= x. \end{aligned}$$

So the Jacobian matrix for $\mathbf{T}^{-1}$ is

$$J_{\mathbf{T}^{-1}} = \begin{pmatrix} -\frac{y}{x^2} & \frac{1}{x} \\ y & x \end{pmatrix}.$$

Okay, now for the result. We'll postpone giving a justification for this theorem until Section 24.6, since I want to have done a few concrete examples before drawing the right picture.

Memorize: Change of variables

Suppose you need to integrate $\iint_{\mathcal{R}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y$ and you have a transition map $\mathbf{T}(u, v) : \mathcal{R}_{\text{new}} \to \mathcal{R}$. Then the transition map lets you change the integral as follows:

$$\iint_{\mathcal{R}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{\mathcal{R}_{\text{new}}} f(u, v)|\det J_{\mathbf{T}}|\,\mathrm{d}u\,\mathrm{d}v$$

Alternatively, if it's easier to compute $J_{\mathbf{T}^{-1}}$, the following formula also works:

$$\iint_{\mathcal{R}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_{\mathcal{R}_{\text{new}}} \frac{f(u, v)}{|\det J_{\mathbf{T}^{-1}}|}\,\mathrm{d}u\,\mathrm{d}v$$

However, in the latter case your Jacobian determinant will have x and y in it that you need to translate back into u and v.

Here $|\det J_{\mathbf{T}}|$ is called the **area scaling factor**: it's the absolute value of the determinant of the Jacobian matrix. It's indeed true that

$$\det J_{\mathbf{T}^{-1}} = \frac{1}{\det(J_{\mathbf{T}})}$$

which means that if your transition map has a nicer inverse than the original, you might prefer to use that instead.

Type signature

The area scaling factor is always a *nonnegative* real number.

Tip

You might find it easier to remember both formulas if you write

$$\mathrm{d}u\,\mathrm{d}v = |\det J_{\mathbf{T}}|\,\mathrm{d}x\,\mathrm{d}y$$

so it looks more like $\mathrm{d}u = \frac{\partial u}{\partial x}\,\mathrm{d}x$ from 18.01. (Indeed the 18.01 formula is the special case of a 1×1 matrix!)

Digression on what $\mathrm{d}u\,\mathrm{d}v$ means

For 18.02, the equation $\mathrm{d}u\,\mathrm{d}v = |\det J_{\mathbf{T}}|\,\mathrm{d}x\,\mathrm{d}y$ is more of a mnemonic right now than an actual equation; that's because in 18.02 we don't give a definition of what $\mathrm{d}x$ or $\mathrm{d}y$ mean. It can be made into a precise statement using something called a *differential form*. This is out of scope for 18.02, which has the unfortunate consequence that I can't give a formal explanation why the change-of-variable formula works. That said, see Section 24.6 later for an informal explanation.

This is the analog in 18.01 when you did change of variables from x to u (called u-substitution sometimes), and you changed $\mathrm{d}x$ to $\frac{\mathrm{d}x}{\mathrm{d}u}\,\mathrm{d}u$. In 18.02, the derivative from 18.01 is replaced by the enormous Jacobian determinant.

Let's see an example of how to carry out this integration.

Sample Question

Compute the area of the region $\mathcal{R}$ bounded by the curves

$$xy = \frac{16}{9}, \quad xy = \frac{16}{25}, \quad x = 4y, \quad y = 4x.$$

Solution. In the previous sections we introduced variables $u = \frac{y}{x}$ and $v = xy$, and considered the region

$$\mathcal{R}_{\text{new}} = \left[\frac{1}{4}, 4\right] \times \left[\frac{16}{25}, \frac{16}{9}\right]$$

which were the pairs of points (u, v) in that rectangle we described earlier. We made a transition map $\mathbf{T} : \mathcal{R}_{\text{new}} \to \mathcal{R}$ written as either

$$\mathbf{T}(u, v) = \left(\sqrt{\frac{v}{u}}, \sqrt{uv}\right)$$

$$\mathbf{T}^{-1}(x, y) = \left(\frac{y}{x}, xy\right).$$

We don't like square roots, so we'll use the determinant of the Jacobian matrix for $\mathbf{T}^{-1}$, which is

$$\det(J_{\mathbf{T}^{-1}}) = \begin{vmatrix} -\frac{y}{x^2} & \frac{1}{x} \\ y & x \end{vmatrix} = \left(-\frac{y}{x^2}\right) \cdot x - \frac{1}{x} \cdot y = -\frac{y}{x} - \frac{y}{x} = -\frac{2y}{x}.$$

Since we used the upside-down version of the formula, we need to translate this back into u and v through the given formula. In this case since $u = \frac{y}{x}$, you can do it just by looking:

$$\det(J_{\mathbf{T}^{-1}}) = -2u.$$

$$\begin{aligned} \operatorname{Area}(\mathcal{R}) &= \int_{u=\frac14}^{4} \int_{v=\frac{16}{25}}^{\frac{16}{9}} \frac{1}{|\det(J_{\mathbf{T}^{-1}})|} \, dv \, du \\ &= \int_{u=\frac14}^{4} \int_{v=\frac{16}{25}}^{\frac{16}{9}} \frac{1}{2u} \, dv \, du \\ &= \int_{u=\frac14}^{4} \frac{1}{2u} \cdot \left(\frac{16}{9} - \frac{16}{25}\right) du \\ &= \frac{128}{225} \int_{u=\frac14}^{4} \frac{1}{u} \, du \\ &= \frac{128}{225}\left(\log 4 - \log\left(\frac14\right)\right) = \boxed{\frac{512\log 2}{225}}. \end{aligned}$$

□

§23.4 [TEXT] Another example: the area of a unit disk

Sample Question

Show that the area of the unit disk $x^2 + y^2 \le 1$ is π.

Solution. For reasons that will soon be obvious, we use the letters r and θ rather than u and v for this problem. This time our cartographer's transition map is going to be given by

$$\mathbf{T} : [0,1] \times [0, 2\pi] \to \mathbb{R}^2$$
$$\mathbf{T}(r, \theta) := (r\cos\theta, r\sin\theta).$$

You might recognize this as polar coordinates. This gives us a way to plot the unit disk as a rectangular map; see the figure.

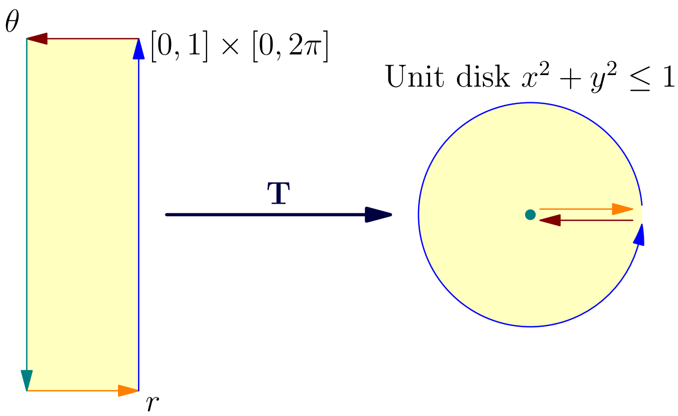

Figure 52: The map $\mathbf{T} : [0,1] \times [0,2\pi] \to \mathbb{R}^2$ whose image is the unit disk.

(Careful students might notice that the points on the line segment from $(0,0)$ to $(1,0)$ are repeated more than once under the transition map; again, in 18.02 we allow this repetition on the boundary.)

We calculate the Jacobian of $\mathbf{T}$:

$$J_{\mathbf{T}} = \begin{pmatrix} \frac{\partial}{\partial r}(r\cos\theta) & \frac{\partial}{\partial \theta}(r\cos\theta) \\ \frac{\partial}{\partial r}(r\sin\theta) & \frac{\partial}{\partial \theta}(r\sin\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

The area scaling factor is then

$$|\det J_{\mathbf{T}}| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta - (-r\sin^2\theta) = r(\cos^2\theta + \sin^2\theta) = r.$$

Hence, the transition map gives us the following change of variables:

$$\iint_{x^2+y^2=1} 1\,\mathrm{d}x\,\mathrm{d}y = \int_{r=0}^{1}\int_{\theta=0}^{2\pi} r\,\mathrm{d}\theta\,\mathrm{d}r.$$

This is easy to integrate:

$$\begin{aligned} \int_{r=0}^{1}\left(\int_{\theta=0}^{2\pi} r\,\mathrm{d}\theta\right)\mathrm{d}r &= \int_{r=0}^{1}(2\pi r)\,\mathrm{d}r \\ &= 2\pi\int_{r=0}^{1}(r)\,\mathrm{d}r \\ &= 2\pi\left[\frac{r^2}{2}\right]_{r=0}^{r=1} = \boxed{\pi}. \end{aligned}$$

□

Tip: remembering forwards vs backwards

If you have trouble remembering which way is "forwards" (meaning you use $|\det J|$) versus which way is "backwards" (meaning you use $\frac{1}{|\det J|}$), it might help to look at Table 12 to see these side by side. Just remember: the *polar* one is forwards, so we get $\mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta$.

Example	Forwards	Backwards
Example setup	$x = r\cos\theta$ $y = r\sin\theta$	$u = y/x$ $v = xy$
Example Jacobian matrix	$\begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}$	$\begin{pmatrix} -y/x^2 & 1/x \\ y & x \end{pmatrix}$
Example Jacobian determinant	r	$-2y/x = -2u$
Example change of variables	$\mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta$	$\mathrm{d}u\,\mathrm{d}v = 2u\,\mathrm{d}x\,\mathrm{d}y$ $\mathrm{d}x\,\mathrm{d}y = \frac{1}{2u}\,\mathrm{d}u\,\mathrm{d}v$

Table 12: Side-by-side comparison of forwards and backwards changes of variables for the two examples we just did.

§23.5 [TEXT] Example: the area of an ellipse

Once we know the area of a circle, we can also compute the area of an ellipse by *reducing* to the area of a circle, as follows.

Sample Question

Let $a, b > 0$ be positive real numbers Compute that the area inside the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Solution. Let $\mathcal{R}$ be the interior of the ellipse. Consider the transformation $\mathbf{T}$ defined by

$$\begin{aligned} x &= au \\ y &= bv. \end{aligned}$$

That is, $\mathbf{T}(u, v) = (au, bv)$. The Jacobian of this matrix is easy to calculate:

$$J_{\mathbf{T}} = \begin{pmatrix} \frac{\partial}{\partial u}(au) & \frac{\partial}{\partial v}(au) \\ \frac{\partial}{\partial u}(bu) & \frac{\partial}{\partial v}(bu) \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

which has determinant ab.

This transformation provides a mapping between the regions

$$\mathbf{T} : \{u^2 + v^2 \le 1\} \to \mathcal{R} = \left\{ \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1 \right\}.$$

Hence, via change of variables the area of $\mathcal{R}$ is related by

$$\begin{aligned} \operatorname{Area}(\mathcal{R}) &= \iint_{\mathcal{R}} 1 \, \mathrm{d}x \, \mathrm{d}y \\ &= \iint_{u^2+v^2 \le 1} \det J_{\mathbf{T}} \, \mathrm{d}u \, \mathrm{d}v \\ &= ab \iint_{u^2+v^2 \le 1} \mathrm{d}u \, \mathrm{d}v. \\ &= ab \operatorname{Area}(\{u^2 + v^2 \le 1\}) = ab\pi. \end{aligned}$$

□

To put this example into a picture, the idea is that we use a change of variables to map the ellipse into a *circle*, where the Jacobian determinant is the constant function ab. That determinant factors out, and we get the result above.

§23.6 [SIDENOTE] Tip: Factoring integrals over rectangles

Especially with polar coordinates, you will often find you get an integral of the shape

$$\int_{u=\text{number}}^{\text{number}} \int_{v=\text{number}}^{\text{number}} f(u) g(v) \, \mathrm{d}v \, \mathrm{d}u$$

that is, the part inside splits cleanly as the product of stuff involving u and stuff involving v. In that case, if you imagine actually doing the integration, you'll find that this actually just equals

$$\left(\int_{u=\text{number}}^{\text{number}} f(u)v \, \mathrm{d}u\right)\left(\int_{v=\text{number}}^{\text{number}} g(v) \, \mathrm{d}v\right).$$

For example, consider the following easy question and solution.

Sample Question

Evaluate

$$\int_{x=0}^{1} \int_{y=0}^{\pi} e^x \sin(y) \, \mathrm{d}y \, \mathrm{d}x.$$

Solution. The integral can be written as:

$$\int_{x=0}^{1} e^x \left(\int_{y=0}^{\pi} \sin(y) \, \mathrm{d}y\right) \mathrm{d}x.$$

The inner integral is

$$\int_{y=0}^{\pi} \sin(y) \, \mathrm{d}y = [-\cos(y)]_{y=0}^{\pi} = (-\cos(\pi)) - (-\cos(0)) = (-(-1)) - (-1) = 1 + 1 = 2.$$

Substitute the result back into the integral:

$$\int_{x=0}^{1} e^x \cdot 2 \, \mathrm{d}x = 2 \int_{x=0}^{1} e^x \, \mathrm{d}x = 2[e^x]_{x=0}^{1} = 2(e^1 - e^0) = \boxed{2e - 2}. \qquad \square$$

If you pay attention to the solution above, you'll notice that in fact $\int_{y=0}^{\pi} \sin(y) \, \mathrm{d}y = 2$ is just a number, and it gets pulled out of the integral right away. So in effect, we actually have

$$\int_{x=0}^{1} \int_{y=0}^{\pi} e^x \sin(y) \, \mathrm{d}y \, \mathrm{d}x = \left(\int_{x=0}^{1} e^x \, \mathrm{d}x\right)\left(\int_{y=0}^{\pi} \sin(y) \, \mathrm{d}y\right).$$

This is a bit of a convenience feature that might save a bit of headspace. It's a tiny optimization, but it's worth pointing out.

Tip

Look for the common pattern

$$\int_{u=\text{number}}^{\text{number}} \int_{v=\text{number}}^{\text{number}} f(u)g(v) \, \mathrm{d}v \, \mathrm{d}u = \left(\int_{u=\text{number}}^{\text{number}} f(u)v \, \mathrm{d}u\right)\left(\int_{v=\text{number}}^{\text{number}} g(v) \, \mathrm{d}v\right).$$

Remember, this doesn't work if either the integrand doesn't factor, or the limits of integration aren't just numbers (i.e. the limit of v depends on u).

As another example of a use case, in the polar integration we just did, we have

$$\int_{r=0}^{1} \int_{\theta=0}^{2\pi} r \, \mathrm{d}\theta \, \mathrm{d}r = \left(\int_{r=0}^{1} r \, \mathrm{d}r\right)\left(\int_{\theta=0}^{2\pi} \theta \, \mathrm{d}\theta\right) = \left[\frac{r^2}{2}\right]_{r=0}^{r=1} \cdot (2\pi) = \pi.$$

(Polar coordinates, covered next chapter, have this particular pattern a lot. Often the thing you're integrating has no θ dependence at all.)

§23.7 [EXER] Exercises

Exercise 23.1. Let $\mathcal{R}$ be all the points on or inside the triangle with vertices $(0,0)$, $(1,2)$ and $(2,1)$. Compute

$$\iint_{\mathcal{R}} \frac{(x+y)^2}{xy} \, \mathrm{d}x \, \mathrm{d}y.$$

(Recommended approach: use change of variables with $u = x + y$ and $v = \frac{x}{y}$.)

Chapter 24. Polar coordinates

§24.1 [TEXT] Polar coordinates are a special case of change of variables

Last chapter one of the transition maps we used was

$$\mathbf{T}_{\text{polar}}(r, \theta) = (r\cos\theta, r\sin\theta).$$

This particular change is so common that you should actually memorize its Jacobian determinant and area scaling factor. Remember from last chapter we computed

$$J_{\mathbf{T}} = \begin{pmatrix} \frac{\partial}{\partial r}(r\cos\theta) & \frac{\partial}{\partial \theta}(r\cos\theta) \\ \frac{\partial}{\partial r}(r\sin\theta) & \frac{\partial}{\partial \theta}(r\sin\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

$$|\det J_{\mathbf{T}}| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta - (-r\sin^2\theta) = r(\cos^2\theta + \sin^2\theta) = r.$$

You should actually just remember the final result of this calculation so you don't have to work it out again. Colloquially, it can be written like so:

> **! Memorize: The Jacobian for polar coordinates**
>
> When converting from Cartesian coordinates (x, y) to polar coordinates (r, θ) you should replace $\mathrm{d}x\,\mathrm{d}y$ to $r\,\mathrm{d}r\,\mathrm{d}\theta$. Colloquially, we write
>
> $$\mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta.$$

Many other sources will write $\mathrm{d}A$ as a shorthand for *both*: so if you have (x, y) coordinates then $\mathrm{d}A = \mathrm{d}x\,\mathrm{d}y$, while if you have polar coordinates then $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$. (They're equal, after all.) Again, this can be made more precise later on in life once you have access to a new object called a *differential form*, but for now just treat it as a mnemonic for one really common change of variables, rather than a formal statement.

As training wheels, I'm still going to avoid writing $\mathrm{d}A$ for one more chapter, so that when you see $\mathrm{d}x\,\mathrm{d}y$ or $\mathrm{d}y\,\mathrm{d}x$ you know you're *supposed* to make a change of variables (and won't accidentally write $\mathrm{d}r\,\mathrm{d}\theta$ with the factor of r missing).

§24.2 [TEXT] Polar coordinates can be thought of as a coordinate system

In what follows, true to the name "polar *coordinates*", I'll write just

$$(r, \theta)_{\text{pol}} := (r\cos\theta, r\sin\theta)$$

so I don't have to keep dragging the $\mathbf{T}_{\text{polar}}$ everywhere. (However, other places will just write (r, θ) everywhere, since it's unlikely to be confused with xy-coordinates due to the letter change.)

The upshot is that in practice:

> **Idea**
>
> Once you remember that $\mathrm{d}x\,\mathrm{d}y$ turns into $r\,\mathrm{d}r\,\mathrm{d}\theta$, you can jump into problems directly in polar coordinates, skipping the x and y altogether.

For example, if you want to find the area of the unit disk, you know in polar coordinates the unit disk is $0 \le r \le 1$ and $0 \le \theta \le 2\pi$, so you can just directly think via the integral

$$\int_{r=0}^{1} \int_{\theta=0}^{2\pi} r \, \mathrm{d}\theta \, \mathrm{d}r$$

and not even bother writing the xy-version $\iint_{x^2+y^2 \le 1} \mathrm{d}x \, \mathrm{d}y$. Compared to back in Section 23.4, it's the same thing; it's just a shift in mindset where you go from "take an xy-picture and translate into polar coordinates" to "take a picture and write it directly in polar coordinates".

§24.3 [TEXT] Famous example: the offset circle

There's one particularly famous exercise that's often used when teaching this stuff. I'm actually going to split it into two parts.

Sample Question

Let $\mathcal{R}$ denote the disk of radius 1 centered at $(1, 0)$. Express the region $\mathcal{R}$ in polar coordinates.

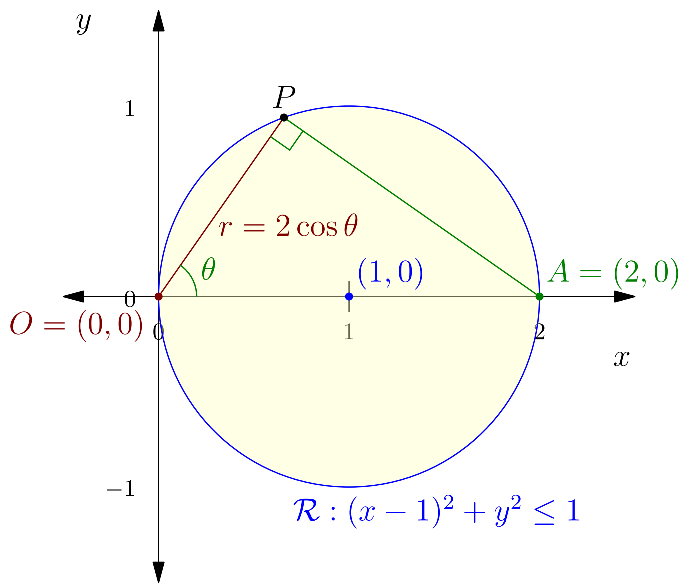

Figure 53: A sketch of $(x-1)^2 + y^2 \le 1$. One might expect this to be nasty when converted to polar, but it turns out to be $r \le 2\cos\theta$ for $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$, which is much better than expected.

Solution. Here are two ways to proceed.

Geometric approach See Figure 53. It's clear from the figure we want $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$. If one lets $O = (0, 0)$, P be a point on the boundary of the circle, and then let $A = (2, 0)$ (so segment AO is a diameter of the circle). Then we in fact get a right triangle $\triangle OPA$ with

$$\angle P = 90^\circ, \quad OA = 2, \quad \text{and } \angle AOP = \theta.$$

Hence, the boundary of our offset circle is given by $r = 2\cos\theta$. The disk (i.e. the part inside the boundary) is then

$$0 \le r \le 2\cos\theta.$$

Algebra approach Initially we have (x, y) such that

$$(x-1)^2 + y^2 \le 1.$$

Expanding this equation gives

$$x^2 - 2x + 1 + y^2 \le 1 \implies x^2 + y^2 \le 2x.$$

In polar coordinates, $x = r\cos\theta$ and $y = r\sin\theta$, so we substitute to get

$$r^2 \le 2r\cos\theta.$$

We need $\cos\theta \ge 0$ to be nonnegative for this to be feasible, and we take $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$ as a result. In that case the condition ends up becoming

$$0 \le r \le 2\cos\theta.$$

In conclusion, the answer is $\mathcal{R}$ in polar coordinates is exactly

$$\boxed{-\frac{\pi}{2} \le \theta \le \frac{\pi}{2} \quad \text{and} \quad r \le 2\cos\theta}. \qquad \square$$

In other words, the graph of $r = 2\cos\theta$ is actually just an offset circle. This is a bit of a surprise to people who are seeing it for the first time, and even a bit to me now, but the geometry argument should justify why.

Okay, here's the famous exercise I promised you.

Sample Question

Let $\mathcal{R}$ denote the disk of radius 1 centered at $(1,0)$. Calculate $\iint_{\mathcal{R}} \sqrt{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y$.

Solution. If you try to use xy integration, it's a disaster. But we just saw the polar coordinates for $\mathcal{R}$ are surprisingly good. So of course let

$$x = r\cos\theta, \quad y = r\sin\theta.$$

In these coordinates, $\sqrt{x^2+y^2} = r$, so the integrand becomes r.

As we just saw, the region $\mathcal{R}$ consists of $-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$ and $0 \le r \le 2\cos\theta$, so

$$\iint_{\mathcal{R}} \sqrt{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y = \int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{r=0}^{2\cos\theta} r \cdot \underbrace{r\,\mathrm{d}r\,\mathrm{d}\theta}_{=\mathrm{d}x\,\mathrm{d}y} = \int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{r=0}^{2\cos\theta} r^2\,\mathrm{d}r\,\mathrm{d}\theta.$$

First, integrate with respect to r:

$$\int_{r=0}^{2\cos\theta} r^2\,\mathrm{d}r = \left[\frac{r^3}{3}\right]_{r=0}^{2\cos\theta} = \frac{(2\cos\theta)^3}{3} = \frac{8\cos^3\theta}{3}.$$

Now, substitute this result into the outer integral:

$$\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{8\cos^3\theta}{3}\,\mathrm{d}\theta = \frac{8}{3}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3\theta\,\mathrm{d}\theta.$$

To evaluate $\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3\theta\,\mathrm{d}\theta$, we need to find an antiderivative of $\cos^3\theta$. That will require a bit of trigonometry acrobatics: the idea is to use

$$\cos^3\theta = \cos\theta(1-\sin^2\theta).$$

Then, set $u = \sin\theta$, so $\mathrm{d}u = \cos\theta\,\mathrm{d}\theta$:

$$\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3\theta\,\mathrm{d}\theta = \int_{u=-1}^{1} (1-u^2)\,\mathrm{d}u = \left[u - \frac{u^3}{3}\right]_{-1}^{1} = \left(1-\frac{1}{3}\right) - \left(-1+\frac{1}{3}\right) = \frac{4}{3}.$$

Substitute this result back into the integral:

$$\frac{8}{3}\int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3\theta\,\mathrm{d}\theta = \frac{8}{3}\cdot\frac{4}{3} = \boxed{\frac{32}{9}}.$$

□

§24.4 [TEXT] Example: ugly square roots or $x^2 + y^2$ everywhere

One other way to generate exam questions to force students to use polar coordinates is to dump a bunch of square roots everywhere that make xy-integration infeasible, thus requiring the use of polar coordinates instead. Here's an example of what that could look like.

Sample Question

Compute the double integral

$$\int_{y=0}^{3}\int_{x=-\sqrt{9-y^2}}^{x=\sqrt{9-y^2}} \left(x^2+y^2\right)^{\frac{5}{2}}\,\mathrm{d}x\,\mathrm{d}y.$$

Solution. Because of the presence of $\sqrt{9-y^2}$ and the $\left(x^2+y^2\right)^{\frac{5}{2}}$, we're practically forced to use polar coordinates. Indeed, the right way to think of the region we're integrating over is that it consists of

$$0 \le y \le 3 \quad\text{and}\quad x^2+y^2 \le 9$$

which is just the upper half a circle of radius 3. So using polar coordinates is just obviously the right thing to do, because the limits of integration are amazing:

- The region is defined by $0 \le r \le 3$ and $0 \le \theta \le \pi$.
- The integrand $x^2+y^2 = r^2$, so $\left(x^2+y^2\right)^{5/2} = r^5$.

Thus, the integral in polar coordinates becomes:

$$\int_{\theta=0}^{\pi}\int_{r=0}^{3} r^5 \cdot \underbrace{r\,\mathrm{d}r\,\mathrm{d}\theta}_{=\mathrm{d}x\,\mathrm{d}y} = \int_{\theta=0}^{\pi}\int_{r=0}^{3} r^6\,\mathrm{d}r\,\mathrm{d}\theta.$$

Now, integrate with respect to r:

$$\int_{r=0}^{3} r^6\,\mathrm{d}r = \left[\frac{r^7}{7}\right]_{r=0}^{3} = \frac{3^7}{7} = \frac{2187}{7}.$$

Now, integrate with respect to θ:

$$\int_{\theta=0}^{\pi} \frac{2187}{7}\,\mathrm{d}\theta = \frac{2187}{7}\cdot\pi = \frac{2187\pi}{7}.$$

□

You could easily imagine doing something similar with some different artificial function involving x^2+y^2 in some other way:

Sample Question

Compute the double integral

$$\int_{y=0}^{3} \int_{x=-\sqrt{9-y^2}}^{x=\sqrt{9-y^2}} \frac{1}{x^2+y^2+17} \, dx \, dy.$$

Solution. The region is the same, the only change is what to do with the thing inside:

$$\frac{1}{x^2+y^2+17} = \frac{1}{r^2+17}.$$

So, the overall integral becomes

$$\int_{\theta=0}^{\pi} \int_{r=0}^{3} \frac{1}{r^2+17} \cdot \underbrace{r \, dr \, d\theta}_{= dx \, dy}.$$

Now, we evaluate the inner integral:

$$\int_{r=0}^{3} \frac{r}{r^2+17} \, dr.$$

To integrate this, use the substitution $u = r^2 + 17$, so $du = 2r \, dr$ or $\frac{du}{2} = r \, dr$. When $r = 0$, $u = 17$; when $r = 3$, $u = 26$. The integral becomes:

$$\int_{r=0}^{3} \frac{r}{r^2+17} \, dr = \int_{u=17}^{26} \frac{1}{u} \cdot \frac{du}{2} = \frac{1}{2} \int_{u=17}^{26} \frac{1}{u} \, du.$$

Integrating with respect to u:

$$\frac{1}{2} \int_{u=17}^{26} \frac{1}{u} \, du = \frac{1}{2} [\log u]_{u=17}^{26} = \frac{1}{2}(\log 26 - \log 17) = \frac{1}{2} \log \frac{26}{17}.$$

Now, integrate with respect to θ:

$$\int_{\theta=0}^{\pi} \frac{1}{2} \log \frac{26}{17} \, d\theta = \boxed{\frac{\pi}{2} \log \frac{26}{17}}.$$

□

§24.5 [TEXT] Example: region described in circular terms

Another way you can force students to use polar coordinates is to give a region which is described as a circle to begin with. Again xy-coordinates are either infeasible or at least annoying.

Sample Question

The unit circle centered at $(0,0)$ is divided into four quarters by the x and y axes. Compute the center of mass of the quarter-circle in the first quadrant, assuming a uniform density distribution.

Solution. To find the center of mass of the quarter-circle in the first quadrant, we consider the region bounded by $0 \le r \le 1$ and $0 \le \theta \le \frac{\pi}{2}$ in polar coordinates. Since the density is uniform, we can use symmetry and polar coordinates to find the coordinates $(\bar{x}, \bar{y})$ of the center of mass.

Call the quarter-circle $\mathcal{R}$. The area of the quarter-circle is one-fourth of the area of the unit circle:

$$\operatorname{Area}(\mathcal{R}) = \frac{1}{4}\pi.$$

In polar coordinates, the coordinates x and y are given by:

$$x = r\cos\theta, \quad y = r\sin\theta.$$

The center of mass coordinates $(\bar{x}, \bar{y})$ are given by

$$\bar{x} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y, \quad \bar{y} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} y \, \mathrm{d}x \, \mathrm{d}y.$$

Since $\mathrm{d}x \, \mathrm{d}y = r \, \mathrm{d}r \, \mathrm{d}\theta$, we can express $\iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y$ and $\iint_{\mathcal{R}} y \, \mathrm{d}x \, \mathrm{d}y$ as follows:

For $\bar{x}$, we have:

$$\iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y = \int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{1} r\cos\theta \cdot r \, \mathrm{d}r \, \mathrm{d}\theta = \int_{\theta=0}^{\frac{\pi}{2}} \cos\theta \int_{r=0}^{1} r^2 \, \mathrm{d}r \, \mathrm{d}\theta.$$

First, integrate with respect to r:

$$\int_{r=0}^{1} r^2 \, \mathrm{d}r = \left[\frac{r^3}{3}\right]_{r=0}^{1} = \frac{1}{3}.$$

Thus,

$$\iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y = \int_{\theta=0}^{\frac{\pi}{2}} \cos\theta \cdot \frac{1}{3} \, \mathrm{d}\theta = \frac{1}{3} \int_{\theta=0}^{\frac{\pi}{2}} \cos\theta \, \mathrm{d}\theta.$$

Now, integrate with respect to θ:

$$\int_{\theta=0}^{\frac{\pi}{2}} \cos\theta \, \mathrm{d}\theta = [\sin\theta]_{\theta=0}^{\frac{\pi}{2}} = \sin\left(\frac{\pi}{2}\right) - \sin(0) = 1.$$

So,

$$\iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y = \frac{1}{3}.$$

Therefore,

$$\bar{x} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} x \, \mathrm{d}x \, \mathrm{d}y = \frac{1}{\frac{\pi}{4}} \cdot \frac{1}{3} = \frac{4}{3\pi}.$$

By symmetry, the calculation for $\bar{y}$ will be identical, since the quarter-circle region is symmetric about the line $y = x$:

$$\bar{y} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} y \, \mathrm{d}x \, \mathrm{d}y = \frac{4}{3\pi}.$$

Hence, the final answer is

$$(\bar{x}, \bar{y}) = \boxed{\left(\frac{4}{3\pi}, \frac{4}{3\pi}\right)}.$$

□

§24.6 [SIDENOTE] A picture of $r\,\mathrm{d}r\,\mathrm{d}\theta = (r\,\mathrm{d}\theta)\,\mathrm{d}r$

Here's a bit of a pictorial explanation of why the result $(r\,\mathrm{d}\theta)\,\mathrm{d}r$ makes sense. None of this is considered for exam, nor is it actually precise. But it should help with some convince you that $r\,\mathrm{d}r\,\mathrm{d}\theta$ is correct, and more generally that the Jacobian determinant is the right scaling factor.

The way that people typically draw a picture of $\mathrm{d}x\,\mathrm{d}y$ is to take some point $P = (x, y)$ in the plane and imagining looking at the arrow pointing from P to $(x + \mathrm{d}x, y)$ and $(x, y + \mathrm{d}y)$, where $\mathrm{d}x$ and $\mathrm{d}y$ are, loosely, "tiny displacements". Then $\mathrm{d}A$ is drawn as the area of the little rectangle you get. See Figure 54, where the two little arrows are drawn in red, and one gets the shaded blue region shown.

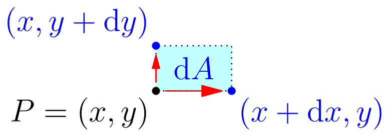

Figure 54: $\mathrm{d}A = \mathrm{d}x\,\mathrm{d}y$, drawn as a picture with small red arrows.

In the xy picture, the point P itself plays little role; the area of the little rectangle is always just $\mathrm{d}x\,\mathrm{d}y$, no matter what point P you pick.

However, when you change to polar coordinates, $\mathrm{d}A$ does actually depend on P: or rather, it doesn't care about θ, but it cares about r. If you have polar coordinates $P = (r, \theta)_{\text{pol}}$ for the starting point and draw two red arrows to $(r + \mathrm{d}r, \theta)_{\text{pol}}$ and $(r, \theta + \mathrm{d}\theta)_{\text{pol}}$, then the first red arrow still always has length $\mathrm{d}r$, but the second red arrow really has length $r\,\mathrm{d}\theta$ — it's close to $\mathrm{d}\theta$ arc of a circle of radius r. You can see this in Figure 55 for two points $P_1 = (r_1, \theta_1)_{\text{pol}}$, and $P_2 = \left(r_2, \theta_{2_{\text{pol}}}\right)$. The point P_1 is close to the origin, so both red arrows are small. But the point P_2 farther has a longer red arrow, because the small change $\mathrm{d}\theta$ is magnified by the radius of the circle. (Some students asked me whether I should be drawing the red arrow curved or straight. The answer is that I don't care — because we're thinking of all the displacements as "tiny", the difference between slightly curling the arrow and having it straight is considered negligible.)

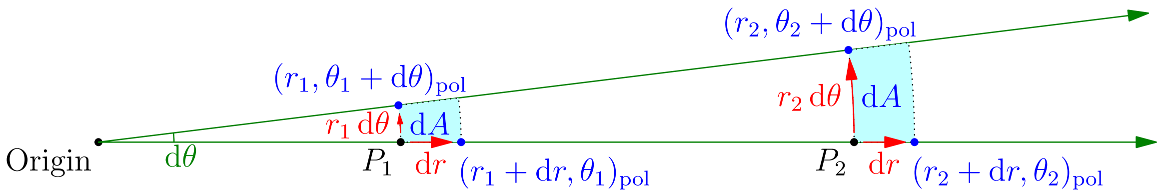

Figure 55: Illustration of $\mathrm{d}A = \mathrm{d}r(r\,\mathrm{d}\theta)$. Note that the $r\,\mathrm{d}\theta$ red arrow gets larger the farther from the origin you are.

The two arrows are almost perpendicular, so the area of the "rectangle"

$$\mathrm{d}A \approx \mathrm{d}r \cdot (r\,\mathrm{d}\theta)$$

which is what we expected.

So where does the Jacobian come in? Let's zoom in a lot on another random point P in polar coordinates, in Figure 56. This is similar to the last figure, but we've chosen a point P for which $\theta > 0$, so neither red arrow is parallel to the x-axis. The new feature is that the two red arrows now have their x and y coordinates written out:

- The first red arrow from P to $(r + \mathrm{d}r, \theta)_{\text{pol}}$ can be written with xy-components as

$$\mathbf{v}_1 := \begin{pmatrix} \cos\theta\,\mathrm{d}r \\ \sin\theta\,\mathrm{d}r \end{pmatrix}.$$

- The second red arrow from P to $(r, \theta + \mathrm{d}\theta)_{\text{pol}}$ can be written with xy-components as

$$\mathbf{v}_2 := \begin{pmatrix} -r \sin\theta \, \mathrm{d}\theta \\ r \cos\theta \, \mathrm{d}\theta \end{pmatrix}.$$

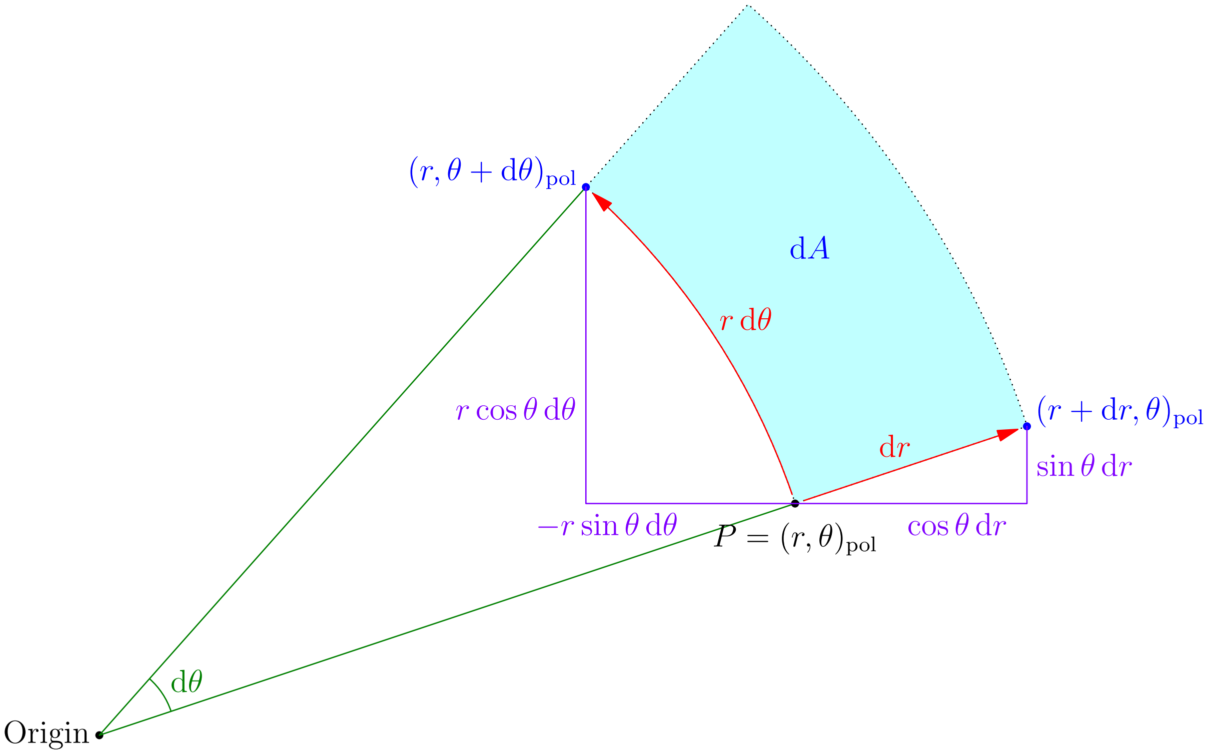

Figure 56: Showing where the polar Jacobian comes from in the change of variables formula. This picture has been exaggerated with a really large $\mathrm{d}\theta$ to make it legible, but in reality you should imagine $\mathrm{d}\theta$ is really tiny instead, so that the shaded region is basically a rectangle.

Now, if we are willing to approximate $\mathrm{d}A$ with the parallelogram spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$ — and we are willing to when $\mathrm{d}\theta$ and $\mathrm{d}r$ are really tiny (in contrast to this cartoon where $\mathrm{d}\theta$ has been drawn pretty big to make the picture legible) — then the approximation is given by the determinant from all the way back in Section 3.4:

$$\begin{aligned} \mathrm{d}A &\approx \left| \det \begin{pmatrix} \cos\theta \, \mathrm{d}r & -r \sin\theta \, \mathrm{d}\theta \\ \sin\theta \, \mathrm{d}r & r \cos\theta \, \mathrm{d}\theta \end{pmatrix} \right| \\ &= \left| \det \begin{pmatrix} \cos\theta & -r \sin\theta \\ \sin\theta & r \cos\theta \end{pmatrix} \right| \mathrm{d}r \, \mathrm{d}\theta \end{aligned}$$

and there's the Jacobian determinant we were waiting for! Ta-da.

Any change of variables can be drawn with a similar cartoon to Figure 56 to explain where the Jacobian comes from with precisely the same reasoning. To spell it out in the 2D case, suppose

$$\mathbf{T}(u, v) = (p(u, v), q(u, v))$$

is any general transition map (so the example we just did was $\mathbf{T}(r, \theta) = (r \cos\theta, r \sin\theta)$). One draws red arrows $\mathbf{T}(u, v)$ to $\mathbf{T}(u + \mathrm{d}u, v)$ and $\mathbf{T}(u, v + \mathrm{d}v)$ for "small" changes $\mathrm{d}u$ and $\mathrm{d}v$. These vectors will correspond approximately to the two vectors

$$\mathbf{v}_1 = \mathbf{T}(u + \mathrm{d}u, v) - \mathbf{T}(u, v) \approx \frac{\partial}{\partial u}\mathbf{T} = \frac{\partial p}{\partial u}\mathbf{e}_1 + \frac{\partial q}{\partial u}\mathbf{e}_2$$

and

$$\mathbf{v}_2 = \mathbf{T}(u, v + \mathrm{d}v) - \mathbf{T}(u, v) \approx \frac{\partial}{\partial v}\mathbf{T} = \frac{\partial p}{\partial v}\mathbf{e}_1 + \frac{\partial q}{\partial v}\mathbf{e}_2.$$

which each give a row of the Jacobian matrix; then the determinant gives the area of the parallelogram spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$; that coincides exactly with the (absolute value of the) Jacobian determinant.

The argument in 3D (and n dimensions in general) is the same, where the parallelogram is replaced by a parallelepiped, etc.

§24.7 [EXER] Exercises

Exercise 24.1. Compute

$$\int_{x=0}^{1} \int_{y=0}^{\sqrt{1-x^2}} xy \, \mathrm{d}y \, \mathrm{d}x.$$

Exercise 24.2. Compute

$$\iint_{(x-1)^2+y^2 \le 1} \frac{1}{\sqrt{x^2+y^2}} \, \mathrm{d}x \, \mathrm{d}y.$$

Exercise 24.3 (*). Compute

$$\iint_{x^2+y^2 \le 1} \sqrt{\left(x + \frac{3}{5}\right)^2 + \left(y + \frac{4}{5}\right)^2} \, \mathrm{d}x \, \mathrm{d}y.$$

Chapter 25. All the gazillion weird d shorthands

§25.1 [TEXT] The shorthand $\mathrm{d}A := \mathrm{d}x\,\mathrm{d}y$ for area

Up to here I've been pretty careful to always write $\iint_{\mathcal{R}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ to make it obvious what the integration variables are.

However, some of you are probably already starting to get tired of writing $\mathrm{d}x\,\mathrm{d}y$ and $\mathrm{d}y\,\mathrm{d}x$. In particular, I advised you earlier that you should prefer to write

$$\int_{x=0}^{5}\int_{y=0}^{3} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$$

rather than the harder-to-read $\int_0^5\int_0^3 f(x,y)\,\mathrm{d}y\,\mathrm{d}x$. Those of you who took my advice may not want to waste the time of remembering whether it's $\mathrm{d}y\,\mathrm{d}x$ or $\mathrm{d}x\,\mathrm{d}y$ at the end, since it doesn't matter for you anymore. For that reason, at this point I hereby bestow on you the following definition:

Definition of $\mathrm{d}A$

We let $\mathrm{d}A$ be a shorthand for either $\mathrm{d}x\,\mathrm{d}y$ or $\mathrm{d}y\,\mathrm{d}x$, whichever one is appropriate for the given context.

So now you can just write:

$$\int_{x=0}^{5}\int_{y=0}^{3} f(x,y)\,\mathrm{d}A.$$

I guess that saves two characters.

Tip: Variable names are often omitted too

In fact, when you use shorthand, you may even leave out x and y from f and just write

$$\int_{x=0}^{5}\int_{y=0}^{3} f\,\mathrm{d}A.$$

So any time shorthand is being used, don't be surprised if the variable names are missing altogether.

Be careful about overdoing this shorthand! For example, if you are working with *polar* coordinates, then in fact

$$\mathrm{d}A = \mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\theta$$

as we just saw. Note the extra factor of r! Seriously, $\mathrm{d}A \neq \mathrm{d}r\,\mathrm{d}\theta$!

If you trust yourself to not forget about the factor of r, or if you're doing a calculation for which the actual variables don't matter, you can also use $\mathrm{d}A$ here. For example, you might write

$$\iint_{\text{unit disk}} \mathrm{d}A = \pi$$

to say the area integral of the unit disk is π. (Pure mathematicians might appreciate how this does not commit to any choice of coordinates.)

But if you do this, be honest with yourself about whether you trust yourself with the shorthand:

> *A bad workman blames his tools.*

§25.2 [TEXT] … and six more shorthands

When we talk about vector fields or even just arc length, there are more new types of integrals. And people have all sorts of analogous shorthands. If you read enough different books, you'll probably eventually see all of $\mathrm{d}\mathbf{r}$, $\mathrm{d}A$, $\mathrm{d}V$, $\mathrm{d}s$, $\mathrm{d}S$, $\mathrm{d}\mathbf{S}$, in various online books. I can't imagine how annoying this is to someone learning the subject for the first time.

Symbol	Name	Used in	Abbreviation for
$\mathrm{d}A$	Area	Double/area integrals	$\mathrm{d}x\,\mathrm{d}y$ (in polar, replaced immediately with $r\,\mathrm{d}r\,\mathrm{d}\theta$)
$\mathrm{d}s$	Arc length	Scalar-field line integrals (in Chapter 29)	$\lvert\mathbf{r}'(t)\rvert\,\mathrm{d}t$ where $\mathbf{r}(t)$ parametrizes a path
$\mathrm{d}\mathbf{r}$	Line element	Vector-field line integrals (in Chapter 33)	$\mathbf{r}'(t)\,\mathrm{d}t$ where $\mathbf{r}(t)$ parametrizes a path
$\mathrm{d}S$	Surface area	Scalar-field surface integrals (in Chapter 29)	$\left\lvert\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right\rvert\,\mathrm{d}u\,\mathrm{d}v$ where $\mathbf{r}(u,v)$ parametrizes a surface
$\mathbf{n}\,\mathrm{d}S$ or $\mathrm{d}\mathbf{S}$	Surface normal	Vector-field surface integrals (in Chapter 38)	$\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v$ where $\mathbf{r}(u,v)$ parametrizes a surface
$\mathrm{d}V$	Volume	Triple/volume integrals (in Chapter 26)	$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$
$\mathbf{n}\,\mathrm{d}s$	Outward normal	Only 2D flux (in Chapter 35)	(90° clockwise rotation of $\mathbf{r}'(t)$) $\mathrm{d}t$

Table 13: A bunch of shorthands you'll meet later. Note that pretty much there is one shorthand for each kind of integral in Figure 41. In Part India when the upgraded poster Figure 75 is introduced, each of the new kinds of integrals (the purple pictures) also has a new shorthand.

So here's the deal.

- First, I'm going to make the following table of what all these shorthands mean. The result is Table 13. Feel free to print it and have it with you. Note that you haven't met most of these yet, so only the first row will make sense for now.
- Second, I'm going to avoid the shorthand when I first introduce things — for example, I deliberately avoided any shorthand on Figure 41 and the later Figure 75 — but later on I'll start to use it as you get more practice.
- Third, in each place where the shorthand *could* be used for the first time, I'll mention it. That is, I'll let you know every time a new row of the table is introduced.

But again, I think the thing to take away is that each of these is a shorthand. So if you don't like shorthands, you can just always replace it with the thing it stands for.

Digression on differential forms

Calling these a shorthand is a bit of a white lie, in that there is actually a rhyme and reason to all these d symbols. Most of them are what are called *differential forms* or *densities*, and you can make a precise definition of what these are. But this is so far beyond the scope of 18.02 I won't spend any more space on it.

Part Hotel: 3D integrals of scalar functions

For comparison, Part Hotel corresponds to §13.6-13.9 and §17.1-§17.6 of Poonen's notes.

Chapter 26. Triple integrals

We're going to now consider integrals with three variables, rather than two. If you understood double integrals, then triple integrals is more of the same:

> **Idea**
>
> All the two-variable stuff ports over to three-variable stuff in the obvious way.

§26.1 [RECAP] Recap of triple integrals

I'm cheekily calling this section a "recap" to emphasize that there's nothing new to learn here. Everything in the below list corresponds to a double integral thing you learned except with three variables rather than two.

One notational change: for 3D solids, I'll prefer to use the letter $\mathcal{T}$ instead of $\mathcal{R}$ for a 3D region moving forward. The reason is that much later on when we discuss the divergence theorem, we'll sometimes have both a 2D region and 3D region at the same time, so one needs different letters.

- Over a rectangular prism, we still integrate $\int_{x=a_1}^{b_1} \int_{y=a_2}^{b_2} \int_{z=a_3}^{b_3} f(x,y,z) \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}x$ one variable at a time.
- You can use
$$\mathrm{d}V := \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$$
as a shorthand if you want; this is the last row of Table 13.
- Instead of xy integration we have xyz-integration. Whereas for double integrations you had two choices (x outer and y inner vs. y outer and x inner), now you have $3! = 6$ choices for the order to do things in:
  - x outermost, y middle, z inner
  - x outermost, z middle, y inner
  - y outermost, x middle, z inner
  - y outermost, z middle, x inner
  - z outermost, x middle, y inner
  - z outermost, y middle, x inner.

  The idea is the same if you have a region that isn't a rectangular prism: write your region as inequalities.
- The change of variables formula is exactly the same, where the Jacobian is now a 3×3 matrix: if $\mathbf{T} : \mathcal{T}_{\text{new}} \to \mathcal{T}$ is a transition map of 3D regions, sending (u,v,w) to (x,y,z), then the Jacobian is
$$J_{\mathbf{T}} = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix}.$$
- Volume is
$$\operatorname{Vol}(\mathcal{T}) := \iiint_{\mathcal{T}} \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

You can take this as a *definition* of volume for this class.

- If $\delta : \mathcal{T} \to \mathbb{R}$ is a density function for a 3D space, then

$$\operatorname{Mass}(\mathcal{T}) := \iiint_{\mathcal{T}} \delta(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

is the total mass. The center of mass is given by three coordinates now:

$$\left(\frac{\iiint_{\mathcal{T}} x \cdot \delta(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\operatorname{Mass}(\mathcal{T})}, \frac{\iiint_{\mathcal{T}} y \cdot \delta(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\operatorname{Mass}(\mathcal{T})}, \frac{\iiint_{\mathcal{T}} z \cdot \delta(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\operatorname{Mass}(\mathcal{T})} \right).$$

(We use δ instead of ρ for 3D typically, because ρ gets used in spherical coordinates.)

§26.2 [TEXT] Examples of triple integrals

Sample Question

Compute the volume of the region bounded by $x^2 + y^2 \le 1$ and $x^2 + z^2 \le 1$.

Solution. Both inequalities must be satisfied simultaneously. Notice that for a fixed x, both y and z are bounded by:

$$y^2 \le 1 - x^2 \implies -\sqrt{1-x^2} \le y \le \sqrt{1-x^2},$$
$$z^2 \le 1 - x^2 \implies -\sqrt{1-x^2} \le z \le \sqrt{1-x^2}.$$

The variable x ranges from -1 to 1.

Hence, we will write this as a triple integral

$$\begin{aligned}
\operatorname{Vol}(\mathcal{T}) &= \int_{x=-1}^{1} \int_{y=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{z=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=-1}^{1} \int_{y=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 2\sqrt{1-x^2}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=-1}^{1} 2\sqrt{1-x^2} \int_{y=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=-1}^{1} 2\sqrt{1-x^2} \cdot 2\sqrt{1-x^2}\,\mathrm{d}x \\
&= \int_{x=-1}^{1} 4(1-x^2)\,\mathrm{d}x \\
&= 4\left[x - \frac{x^3}{3}\right]_{x=-1}^{1} = 4\left[\left(1-\frac{1}{3}\right) - \left(-1+\frac{1}{3}\right)\right] = \boxed{\frac{16}{3}}.
\end{aligned}$$

□

Digression on picture

If you draw a picture of the region, you get the intersection of these two cylinders which forms something apparently called a *Steinmetz solid.* (I say "apparently" because ChatGPT told me this name; I didn't know this had a name before either.)

Surprisingly, you actually *don't* want to use polar (or cylindrical) coordinates on this example. If you try to do so, I think you'll actually get stuck. Straight xyz-integration turns out to work because of the unexpectedly convenient fact that you get two square roots that miraculously cancel.

Sample Question

Compute the volume of the region bounded by the surfaces $z = 3(x^2 + y^2)$ and $z = 72 - 5(x^2 + y^2)$.

Solution. The given surfaces are both paraboloids:

1. $z = 3(x^2 + y^2)$ is an upward-opening paraboloid.
2. $z = 72 - 5(x^2 + y^2)$ is a downward-opening paraboloid.

Before diving in, let's figure out where these two intersect. If we set these equal we get

$$3(x^2 + y^2) = 72 - 5(x^2 + y^2) \Longrightarrow x^2 + y^2 = 9.$$

With that in mind, we can convert the region $\mathcal{T}$ to an inequality format: we write

$$3(x^2 + y^2) \le z \le 72 - 5(x^2 + y^2)$$

for the constraint on z and then

$$x^2 + y^2 \le 9$$

for the constraint on x and y.

Hence, the volume can be written as

$$\operatorname{Vol}(\mathcal{T}) = \iiint_{\substack{x^2+y^2 \le 9 \\ 3(x^2+y^2) \le z \le 72-5(x^2+y^2)}} \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

We'll separate the integral into an integral over the circle $x^2 + y^2 \le 9$ and then a single integral over the resulting z:

$$\begin{aligned} \operatorname{Vol}(\mathcal{T}) &= \iint_{x^2+y^2 \le 9} \left(\int_{z=3(x^2+y^2)}^{72-5(x^2+y^2)} \mathrm{d}z \right) \mathrm{d}x \, \mathrm{d}y \\ &= \iint_{x^2+y^2 \le 9} \left(72 - 8(x^2 + y^2)\right) \mathrm{d}x \, \mathrm{d}y. \end{aligned}$$

At *this* point we'll use polar coordinates: writing $x = r\cos\theta$, and $y = r\sin\theta$ as always, we have

$$\begin{aligned}
\operatorname{Vol}(\mathcal{T}) &= \iint_{x^2+y^2\le 9} \left(72 - 8(x^2+y^2)\right) \, dx \, dy \\
&= \int_{\theta=0}^{2\pi} \int_{r=0}^{3} (72 - 8r^2) \cdot (r \, dr \, d\theta) \\
&= \left(\int_{\theta=0}^{2\pi} 1 \, d\theta \right) \left(\int_{r=0}^{3} (72r - 8r^3) \, dr \right) \\
&= 2\pi \cdot \left[36r^2 - 2r^4 \right]_{r=0}^{3} = \boxed{324\pi}.
\end{aligned}$$

□

This previous example shows how, because of the way the problem was set, it was natural to do the integral for z separately but do x and y with polar coordinates. This technique is called *cylindrical coordinates*, a name that doesn't need to exist because it's just polar coordinates with z tacked on.

(Actually, I think the most surprising thing is the example we need with the Steinmetz solid earlier is *not* good to do with cylindrical coordinates, despite appearances.)

§26.3 [TEXT] Cylindrical coordinates (i.e. polar with z tacked on)

There's actually nothing new happening here — it's just polar coordinates with z tacked on.[22] If you were able to do the earlier example with $z = 3(x^2 + y^2)$ and $z = 72 - 5(x^2 + y^2)$ by yourself without reading the solution, then you can safely skip this entire section!

The transition map $(r, \theta, z) \mapsto (x, y, z)$ is given by

$$\begin{aligned}
x &= r \cos \theta \\
y &= r \sin \theta \\
z &= z.
\end{aligned}$$

This is illustrated in Figure 57.

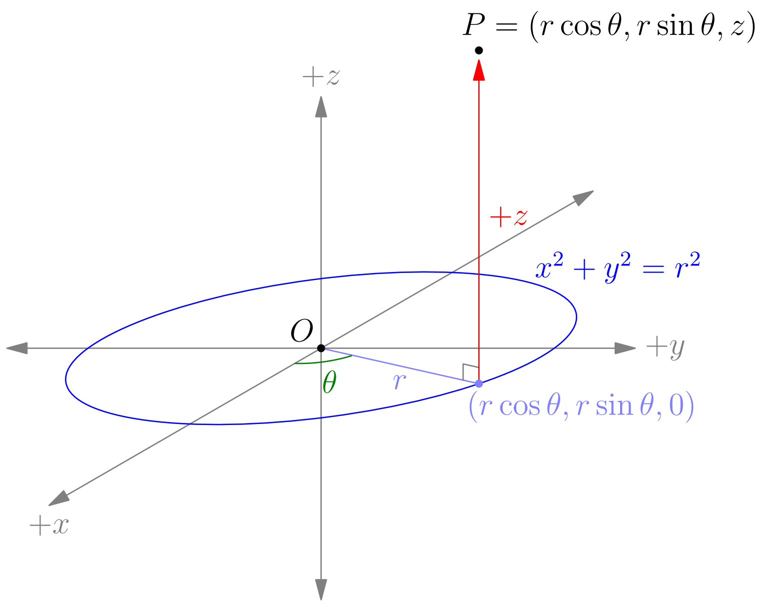

Figure 57: Picture of cylindrical coordinates. The xy-plane (now drawn as "flat" is just polar coordinates, as suggested by the blue circle. And then we tack on a height z.

[22] Technically, we maybe should use a different letter for the new z, but since they're equal we just use the same letter in both places. Also, in principle, I could also introduce a notation $(r, \theta, z)_{\text{cyl}}$ analogous to $(r, \theta)_{\text{pol}}$, but I don't think I'll have a need to do so.

The volume scaling factor is unsurprisingly the same as the one for 2D polar coordinates, and you may have used it implicitly on some previous problem sets already:

❗ Memorize: Scaling factor for cylindrical coordinates

$$\mathrm{d}V := \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z.$$

If you want to see this fully explicitly, you could compute the Jacobian

$$\begin{aligned}\det J_{\text{polar}} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= r.\end{aligned}$$

OK, let's run the example. Note that, as I said, we could have given this example *before* this section.

Sample Question

Compute the volume and center of mass of the cone defined by $9(x^2+y^2) \le z^2$ and $0 \le z \le 5$, assuming uniform density distribution $\delta = 1$.

Solution. The given inequalities describe a single cone with its apex at the origin, extending upwards to $z = 5$. To solve for both the volume and the center of mass, we'll employ cylindrical coordinates due to the symmetry of the cone. As always, $\mathcal{T}$ denotes the region (cone).

The values of z that appear at all are $z = 0$ to 5, and within them we have only the requirement that

$$9(x^2+y^2) \le z^2 \implies \sqrt{x^2+y^2} \le \frac{z}{3}.$$

In other words, we can write

$$\operatorname{Vol}(\mathcal{T}) = \int_{z=0}^{5} \iint_{\sqrt{x^2+y^2} \le \frac{z}{3}} 1\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

However, of course we should just change to cylindrical coordinates right away:

$$\begin{aligned}\operatorname{Vol}(\mathcal{T}) &= \int_{z=0}^{5} \int_{r=0}^{\frac{z}{3}} \int_{\theta=0}^{2\pi} r\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}z \\ &= 2\pi \int_{z=0}^{5} \int_{r=0}^{\frac{z}{3}} r\,\mathrm{d}r\,\mathrm{d}z \\ &= 2\pi \int_{z=0}^{5} \left[\frac{r^2}{2}\right]_{r=0}^{\frac{z}{3}} \mathrm{d}z \\ &= 2\pi \int_{z=0}^{5} \frac{z^2}{18}\,\mathrm{d}z = \frac{\pi}{9} \int_{z=0}^{5} z^2\,\mathrm{d}z = \frac{\pi}{9}\left[\frac{z^3}{3}\right]_{z=0}^{5} = \boxed{\frac{125\pi}{27}}.\end{aligned}$$

This gives us the volume of the cone. And since the density was constant, we also have $\operatorname{Mass}(\mathcal{T}) = \operatorname{Vol}(\mathcal{T}) = \frac{125\pi}{27}$.

As for the center of mass, nominally there are three integrals, but again we can shortcut the calculation by noting that by symmetry the center of mass $(\bar{x}, \bar{y}, \bar{z})$ should lie on the z-axis, meaning $\bar{x} = \bar{y} = 0$. Hence the only one we need to bother with is

$$\begin{aligned}
\bar{z} &= \frac{1}{\operatorname{Mass}(\mathcal{T})} \int_{z=0}^{5} \iint_{\sqrt{x^2+y^2} \le \frac{z}{3}} z \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\
&= \frac{1}{\operatorname{Mass}(\mathcal{T})} \int_{z=0}^{5} \int_{r=0}^{\frac{z}{3}} \int_{\theta=0}^{2\pi} rz \, \mathrm{d}\theta \, \mathrm{d}r \, \mathrm{d}z \\
&= \frac{1}{\operatorname{Mass}(\mathcal{T})} \int_{z=0}^{5} z \int_{r=0}^{\frac{z}{3}} \int_{\theta=0}^{2\pi} r \, \mathrm{d}\theta \, \mathrm{d}r \, \mathrm{d}z \\
&= \frac{2\pi}{\operatorname{Mass}(\mathcal{T})} \int_{z=0}^{5} z \cdot \frac{z^2}{18} \, \mathrm{d}z \quad \text{(repeating from earlier)} \\
&= \frac{2\pi}{\operatorname{Mass}(\mathcal{T})} \left[\frac{z^4}{72} \right]_{z=0}^{5} = \frac{5^4 \cdot \frac{\pi}{36}}{\operatorname{Mass}(\mathcal{T})} = \frac{5^4 \cdot \frac{\pi}{36}}{5^3 \cdot \frac{\pi}{27}} = \frac{15}{4}.
\end{aligned}$$

Hence the center of mass is $\boxed{\left(0, 0, \frac{15}{4}\right)}$. □

§26.4 [TEXT] Gravity

</> **Type signature**

Gravitational force is a vector.

Suppose a point of mass m is located at the origin $O = (0,0,0)$, In general, given a mass m at a point O and a point of mass M at a point P, Newton's law says the gravitational force exerted by P on O is

$$\mathbf{F}_{\text{gravity}} = \frac{G \cdot m \cdot M}{|OP|^2} \cdot \underbrace{\frac{\overrightarrow{OP}}{|OP|}}_{\text{unit vector from } O \text{ to } P}$$

where $G \approx 6.67408 \cdot 10^{-11} \cdot \mathrm{N} \cdot \mathrm{m}^2 \cdot \mathrm{kg}^{-2}$ is the gravitational constant.

But in real life, we usually want our mass M to take up a whole region $\mathcal{T}$, with some density δ. (Point masses don't occur in real life unless you count black holes.) So let's suppose we have a solid mass occupying region $\mathcal{T}$. In that case, each individual point $P = (x, y, z)$ in $\mathcal{T}$ can be thought of a *vector*

$$\text{Gravity exerted by } (x,y,z) \text{ on } (0,0,0) = \frac{Gm \cdot (\delta(x,y,z) \, \mathrm{d}V)}{x^2+y^2+z^2} \cdot \underbrace{\frac{\langle x, y, z \rangle}{\sqrt{x^2+y^2+z^2}}}_{\text{unit vector from } O \text{ to } P}.$$

The total gravitational force is then the integral of this over the entire mass $\mathcal{T}$.

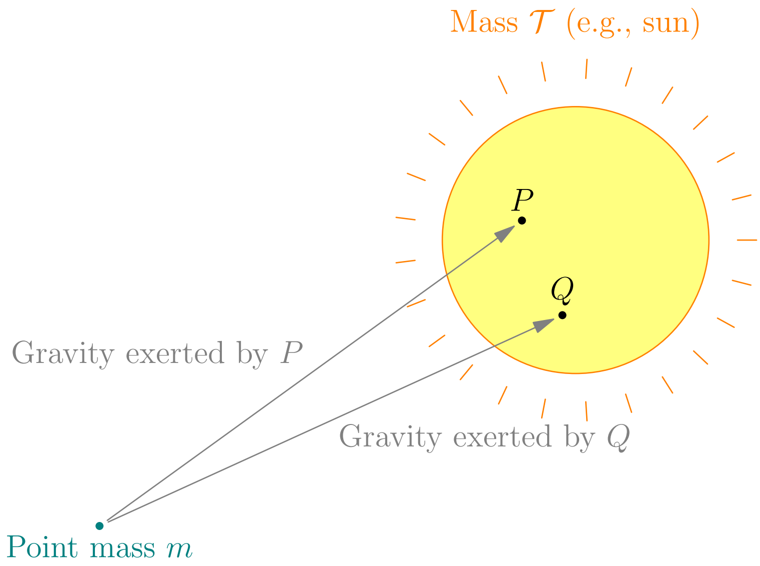

Figure 58: The force of gravity exerted by a large mass $\mathcal{T}$ such as the sun on a point mass of mass m. Each individual point like P or Q in the region $\mathcal{T}$ exerts a tiny force on the point mass of mass m. The total gravitational force is the sum (integral) across the whole region $\mathcal{T}$.

So the total gravitational force is nominally

$$\mathbf{G} = \int_{\mathcal{T}} \frac{Gm \cdot (\delta(x,y,z)\,\mathrm{d}V)}{x^2+y^2+z^2} \cdot \frac{\langle x,y,z\rangle}{\sqrt{x^2+y^2+z^2}}. \qquad (12)$$

Now, if you have been following my advice to always audit type safety, then you should stop me right here. This is the first time in the entire notes that I've had an integral where the integrand is a *vector* rather than the number. What's going on?

The general answer is that you should just do everything component wise. But to keep things simple for the course, I will never use Equation 12 in that form, so that our integrands always have type number rather than type vector. To do this, I'll rewrite Equation 12 as follows:

! Memorize: Gravitational attraction of a region on the origin

Suppose $\mathcal{T}$ is a region with density function δ. The gravitational vector $\mathbf{G} = \langle G_1, G_2, G_3\rangle$ on the origin is defined by

$$\begin{aligned} G_1 &:= Gm \iiint_{\mathcal{T}} \frac{x\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac32}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ G_2 &:= Gm \iiint_{\mathcal{T}} \frac{y\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac32}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ G_3 &:= Gm \iiint_{\mathcal{T}} \frac{z\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac32}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \end{aligned}$$

That is, $\mathbf{G} = G_1\mathbf{e}_1 + G_2\mathbf{e}_2 + G_3\mathbf{e}_3$.

Now G_1, G_2, G_3 are integrals of numbers again, so we're fine.

Because the $(x^2+y^2+z^2)^{\frac32}$ is so awkward to work with, you will commonly do a certain change-of-variables called *spherical coordinates*. So we'll punt all the examples to next chapter, Chapter 27.

§26.5 [EXER] Exercises

Exercise 26.1 (Napkin-ring problem)**.** Let $R > a > 0$ be given real numbers, and let $h := 2\sqrt{R^2 - a^2}$. A cylindrical hole of radius a is drilled through the center of a wooden ball of radius R to get a bead of height h, as shown in Figure 59. Compute the volume of the resulting bead as a function of h.

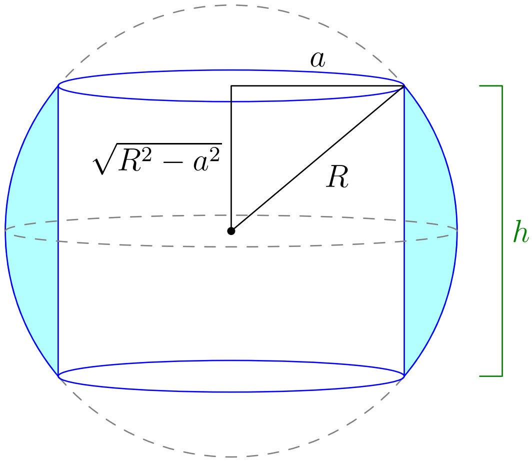

Figure 59: The napkin-ring problem. A bead is shown in blue, drilled out of a sphere of radius R, and with height $h = 2\sqrt{R^2 - a^2}$. One must determine the volume of the bead in terms of h.

Chapter 27. Spherical coordinates

In addition to cylindrical coordinates, there's one more system we'll use, called spherical coordinates. This chapter defines them and shows how to use them.

> ⚠️ **Warning: There are *eight* competing standards, check your book**
>
> Note that there are competing conventions! For us, the letter names are going to mean
>
> $$\begin{aligned} \rho &:= \text{distance to } (0,0,0) && \text{(spelled rho, pronounced like row)} \\ \varphi &:= \text{angle down } z \text{ axis} && \text{(spelled phi, pronounced like fee)} \\ \theta &:= \text{same as in polar coordinates} && \text{(spelled theta, prounced like thay-tah)} \end{aligned}$$
>
> and we write them in that order. However, depending on your book:
> - The names of θ and φ might be swapped. (Also note that the Greek letter φ may be written as ϕ in different fonts. If you use LaTeX, these are `\varphi` and `\phi`.)
> - The order of θ and φ might be swapped (regardless of whether the names change too).
> - ρ might be replaced by r instead.

§27.1 [TEXT] The definition of spherical coordinates

The idea behind spherical coordinates is that the projection of your point P onto the xy-plane will have polar coordinates $(r\cos\theta, r\sin\theta, 0)$ for some r. But then rather than using z to lift the point straight up, you rotate via some angle φ, getting a new distance ρ such that $r = \rho\sin\varphi$ which we'll use to replace r everywhere in what follows. See Figure 60 below.

$$\begin{aligned} P &= (\rho, \varphi, \theta)_{\text{sph}} \\ &= (r\cos\theta, r\sin\theta, \rho\cos\varphi) \\ &= (\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi) \end{aligned}$$

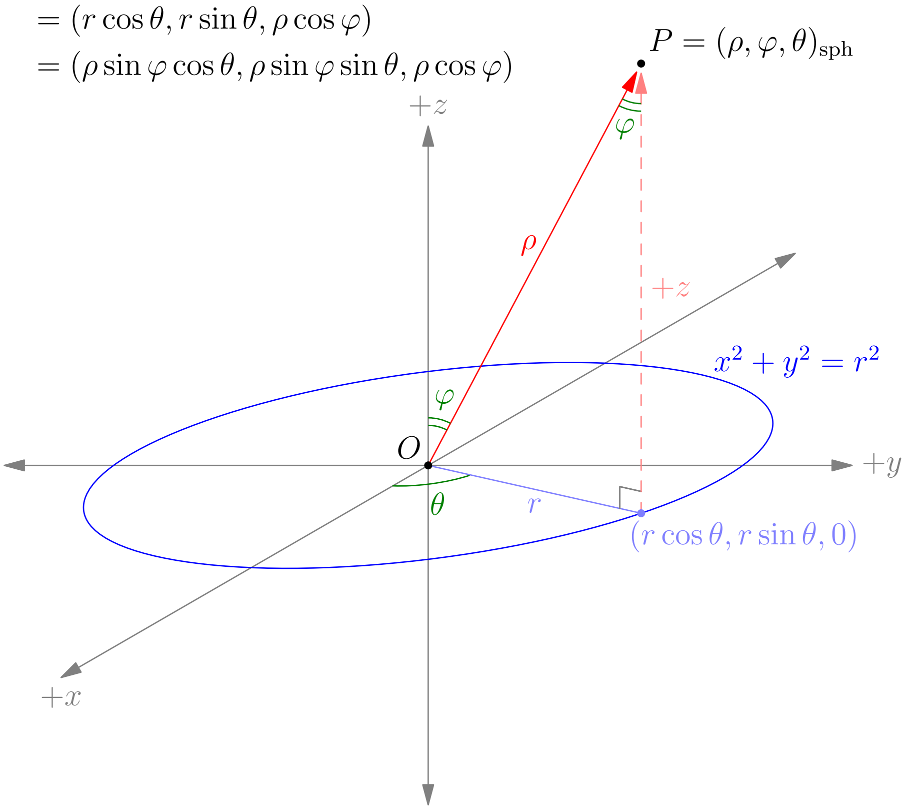

Figure 60: The diagram for spherical coordinates.

Because of the right triangle with angle φ, hypotenuse ρ, and legs r and z, we have

$$\begin{aligned} r &= \rho \sin \varphi \\ z &= \rho \cos \varphi. \end{aligned}$$

Unwinding everything to kill all the r's, the transition map $(\rho, \varphi, \theta) \mapsto (x, y, z)$ is given by

$$\begin{aligned} x &= \underbrace{\rho \sin \varphi}_{=r} \cos \theta \\ y &= \underbrace{\rho \sin \varphi}_{=r} \sin \theta \\ z &= \rho \cos \varphi. \end{aligned}$$

Just like how I wrote $(r, \theta)_{\text{pol}}$ for polar if I needed to be more concise, we'll have the analogous shorthand here:

Definition of spherical coordinates

We define spherical coordinates by

$$(\rho, \varphi, \theta)_{\text{sph}} := (\rho \sin \varphi \cos \theta, \rho \sin \varphi \sin \theta, \rho \cos \varphi).$$

Now, in order to integrate over this, there's supposed to be a change of variables with some Jacobian. To get the area scaling factor, we would compute the Jacobian

$$\det J_{\text{spherical}} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \theta} \end{vmatrix}.$$

This takes some effort, so you probably should only do this once in your life and then remember the result. It works out to

$$\begin{aligned} \det J_{\text{spherical}} &= \begin{vmatrix} \sin \varphi \cos \theta & \rho \cos \varphi \cos \theta & -\rho \sin \varphi \sin \theta \\ \sin \varphi \sin \theta & \rho \cos \varphi \sin \theta & \rho \sin \varphi \cos \theta \\ \cos \varphi & -\rho \sin \varphi & 0 \end{vmatrix} \\ &= \cos \varphi \begin{vmatrix} \rho \cos \varphi \cos \theta & -\rho \sin \varphi \sin \theta \\ \rho \cos \varphi \sin \theta & \rho \sin \varphi \cos \theta \end{vmatrix} + \rho \sin \varphi \begin{vmatrix} \sin \varphi \cos \theta & -\rho \sin \varphi \sin \theta \\ \sin \varphi \sin \theta & \rho \sin \varphi \cos \theta \end{vmatrix} \\ &= \cos \varphi (\rho^2 \cos \varphi \sin \varphi)(\cos^2 \theta + \sin^2 \theta) + \rho^2 \sin \varphi (\sin^2 \varphi)(\cos^2 \theta + \sin^2 \theta) \\ &= \rho^2 \sin \varphi (\cos^2 \varphi + \sin^2 \varphi) \\ &= \rho^2 \sin \varphi. \end{aligned}$$

I tried to do this calculation during recitation and got stuck at the board; not the kind of thing I'm good at. You really don't want to redo this calculation on an exam, so just remember the result.

Memorize: Scaling factor for spherical coordinates

$$\mathrm{d}V := \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \rho^2 \sin \varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta.$$

§27.2 [TEXT] The bounds of φ

Before talking about bounds for spherical coordinates, let me revisit polar coordinates for comparison.

§27.2.1 The bounds for polar coordinates

In polar (or cylindrical) coordinates, when we considered

$$(r, \theta)_{\text{pol}} = (r \cos \theta, r \sin \theta),$$

we usually choose the convention

$$r \geq 0 \quad \text{and} \quad 0 \leq \theta < 2\pi. \tag{13}$$

The thing I want to stress that some thought was put into choosing the interval for θ: the reason we use an interval of length 2π is because if you choose a value of θ bigger than 2π, then the point just "wraps around" to one you already knew; e.g.

$$(r, 2.7\pi)_{\text{pol}} = (r, 0.7\pi)_{\text{pol}}$$

denote the same point. More generally,

$$(r, \theta + 2\pi)_{\text{pol}} = (r, \theta)_{\text{pol}}.$$

That's why we adopt the convention Equation 13. When we define a coordinate system $(r, \theta)_{\text{pol}}$, we want to make sure that every (x, y) point is given by *exactly* one pair. That is, every point should have a coordinate, but different coordinates should occupy different points.

Hence, to avoid repeating the same point with the same coordinates, the usual convention is to choose $0 \leq \theta < 2\pi$, although the convention $-\pi < \theta \leq \pi$ works fine too, as does any interval of length 2π.

> **Digression on $r = 0$**
>
> The claim that Equation 13 lines up perfectly is a white lie: the origin $(0, 0)$ in xy-coordinates can be represented by $(0, \theta)_{\text{pol}}$ for every value of θ. So Equation 13 is almost right, except for the one special case $r = 0$ where θ is indeterminate. We will sweep this under the rug and not think about it.

§27.2.2 The bounds for spherical coordinates

Let's go back to spherical coordinates

$$(\rho, \varphi, \theta)_{\text{sph}} := (\rho \sin \varphi \cos \theta, \rho \sin \varphi \sin \theta, \rho \cos \varphi).$$

We want to choose a convention for values of (ρ, φ, θ) such that (except for a few degenerate cases that we'll ignore) every point has exactly one set of coordinates. The choice that we're going to use is:

> **! Memorize: The convention for spherical coordinate values**
>
> When we want to impose a range of values for spherical coordinates to avoid repeating points, we will choose the following convention:
>
> $$\begin{aligned} &\rho \geq 0 \\ &0 \leq \varphi \leq \pi \\ &0 \leq \theta < 2\pi \end{aligned} \tag{14}$$

That is, I claim that this choice of values Equation 14 will ensure every point is represented exactly once by $(\rho, \varphi, \theta)_{\text{sph}}$, with a small number of exceptions[23] we ignore.

The θ going from 0 to 2π is the same as in polar coordinates. However, the angle for φ might be a surprise to you; a common question asked on various forms is:

> **? Question**
>
> Why does φ only range from 0 to π? What happens if $-\pi < \varphi < 0$ or $\pi < \varphi < 2\pi$?

Well, if my claim about Equation 14 is true, that means if I plug in an "illegal" value of φ, into the formula then I should get a point that's already represented. This is a bit like how $(r, \theta + 2\pi)_{\text{pol}} = (r, \theta)_{\text{pol}}$, but the formula is a bit more complicated. So we'll illustrate two cases in full to show how to convert the illegal value into a legal one.

§27.2.3 First case: Illegal angle greater than π

In this case, I assert the following equation is true:

$$(\rho, \varphi + \pi, \theta)_{\text{sph}} = (\rho, \pi - \varphi, \theta \pm \pi)_{\text{sph}}. \tag{15}$$

Here the sign for $\theta \pm \pi$ is arbitrary, and it's chosen so that $0 \le \theta \pm \pi < 2\pi$ is a legal value.

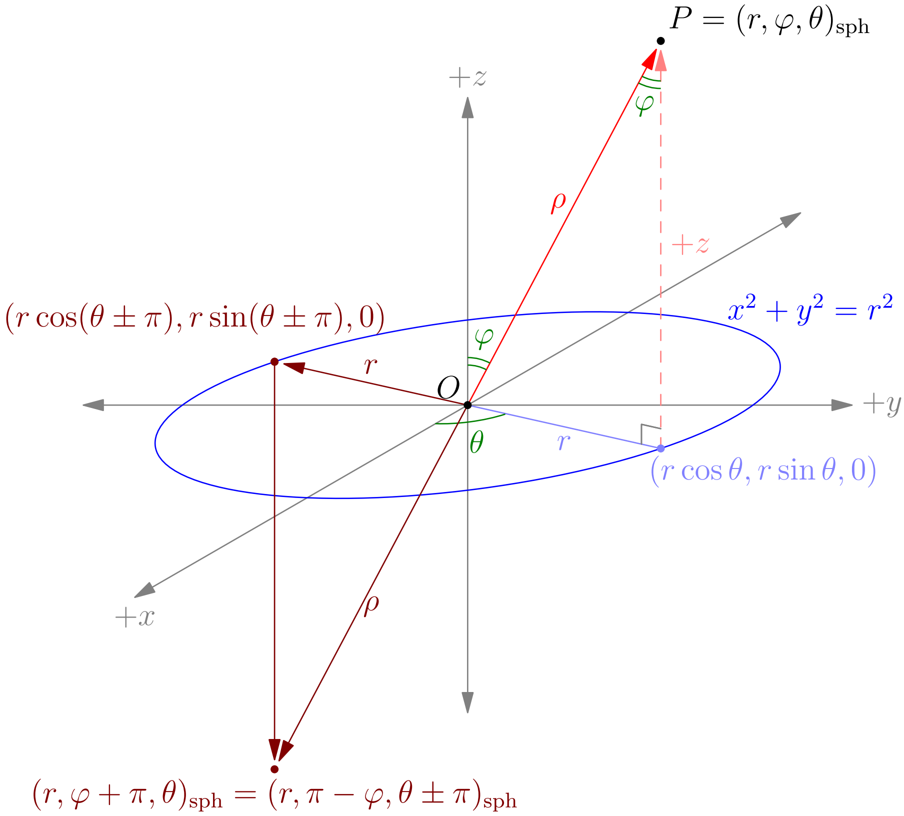

Figure 61: Illustrating Equation 15 in picture format.

Here are two ways to verify Equation 15 is true.

Algebraic proof (easier) We need to verify that the x, y, z coordinates on both sides are the same:

[23] If you want to know, the exceptions are exactly the z-axis, where θ can be arbitrary. Every other point should only appear once.

$$\begin{aligned}\rho\sin(\varphi+\pi)\cos\theta &= \rho\sin(\pi-\varphi)\cos(\theta\pm\pi)\\ \rho\sin(\varphi+\pi)\sin\theta &= \rho\sin(\pi-\varphi)\sin(\theta\pm\pi)\\ \rho\cos(\varphi+\pi) &= \rho\cos(\pi-\varphi).\end{aligned}$$

But $\sin(\varphi+\pi) = -\sin(\pi-\varphi)$, $\cos(\theta\pm\pi) = -\cos\theta$, $\sin(\theta\pm\pi) = -\sin(\theta)$, and $\cos(\varphi+\pi) = \cos(\pi-\varphi)$, so all the equations are true.

Geometric proof (more informative) Look at Figure 61. When the "illegal" value $\varphi+\pi$ is picked for the angle, the red arrow ends up going all the way through O. Hence, the projection of the new point onto the polar circle in blue ends up being the antipode $(r, \theta\pm\pi)_{\text{pol}}$ rather than $(r,\theta)_{\text{pol}}$. If we then think about the angle from the $+z$ axis to the brown radius ρ, it has changed to the angle $\pi-\varphi$ instead. This picture gives a geometric way of seeing why Equation 15 is true.

§27.2.4 Second case: Illegal angle less than zero

This time, I assert the following equation instead:

$$(\rho, -\varphi, \theta)_{\text{sph}} = (\rho, \varphi, \theta\pm\pi)_{\text{sph}}. \tag{16}$$

This is actually a bit easier to see than the last case.

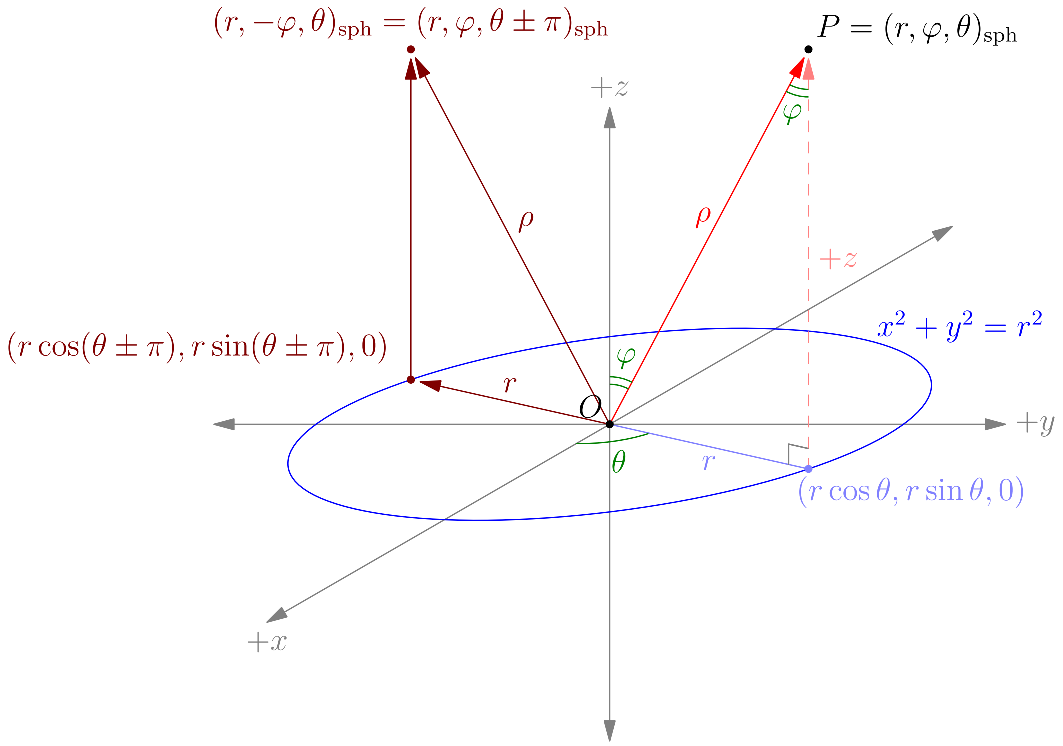

Figure 62: Illustrating Equation 16 in picture format.

Algebraic proof (easier) We need to verify that the x, y, z coordinates on both sides are the same:

$$\begin{aligned}\rho\sin(-\varphi)\cos\theta &= \rho\sin\varphi\cos(\theta\pm\varphi)\\ \rho\sin(-\varphi)\sin\theta &= \rho\sin\varphi\sin(\theta\pm\varphi)\\ \rho\cos(-\varphi) &= \rho\cos\varphi.\end{aligned}$$

But $\sin(-\varphi) = -\sin(\varphi)$, $\cos(\theta\pm\pi) = -\cos\theta$, $\sin(\theta\pm\pi) = -\sin(\theta)$, and $\cos(-\varphi) = \cos\varphi$, so all the equations are true.

Geometric proof (more informative) Look at Figure 62. This time, all that happens is we take the mirror image through the plane formed by the z-axis and the line OP.

§27.3 [TEXT] Examples of using spherical coordinates

Here are two cookie-cutter uses where we have a sphere centered at the origin and we just integrate over the entire sphere (so taking $0 \le \rho \le R$, $0 \le \varphi \le \pi$ and $0 \le \theta \le 2\pi$).

Sample Question

Consider a solid ball of radius R. Compute its volume.

Solution. Placing the ball $\mathcal{T}$ with its center at the origin:

$$\begin{aligned} \operatorname{Vol}(\mathcal{T}) &= \iiint_{\mathcal{T}} 1 \, \mathrm{d}V = \iiint_{\mathcal{T}} \rho^2 \sin\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta \\ &= \int_{\rho=0}^{R} \int_{\varphi=0}^{\pi} \int_{\theta=0}^{2\pi} \rho^2 \sin\varphi \, \mathrm{d}\theta \, \mathrm{d}\varphi \, \mathrm{d}\rho \\ &= \left(\int_{\rho=0}^{R} \rho^2 \, \mathrm{d}\rho \right) \left(\int_{\varphi=0}^{\pi} \sin\varphi \, \mathrm{d}\varphi \right) \left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta \right) \\ &= \frac{R^3}{3} \cdot 2 \cdot (2\pi) = \boxed{\frac{4}{3}\pi R^3}. \end{aligned}$$

□

Sample Question

Consider a solid ball of radius 1. Across all points P inside the ball, compute the average value of the distance from P to the center.

Here, the "average" value of a function f over a solid region $\mathcal{T}$ is defined as $\frac{1}{\operatorname{Vol}(\mathcal{T})} \iiint_{\mathcal{T}} f \, \mathrm{d}V$.

Solution. The sphere has volume $\frac{4}{3}\pi$ as we just saw. The only change to what we did before is that rather than integrating $1 \, \mathrm{d}V$, we replace 1 with the distance:

$$\begin{aligned} \iiint_{\mathcal{T}} (\text{distance to } (0,0,0)) \, \mathrm{d}V &= \iiint_{\mathcal{T}} \rho \, \mathrm{d}V = \iiint_{\mathcal{T}} \rho \cdot (\rho^2 \sin\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta) \\ &= \int_{\rho=0}^{1} \int_{\varphi=0}^{\pi} \int_{\theta=0}^{2\pi} \rho^3 \sin\varphi \, \mathrm{d}\theta \, \mathrm{d}\varphi \, \mathrm{d}\rho \\ &= \left(\int_{\rho=0}^{1} \rho^3 \, \mathrm{d}\rho \right) \left(\int_{\varphi=0}^{\pi} \sin\varphi \, \mathrm{d}\varphi \right) \left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta \right) \\ &= \frac{1}{4} \cdot 2 \cdot (2\pi) = \pi. \end{aligned}$$

So the average value is

$$\frac{\iiint_{\mathcal{T}} (\text{distance to } (0,0,0)) \, \mathrm{d}V}{\operatorname{Vol}(\mathcal{T})} = \frac{\pi}{\frac{4}{3}\pi} = \boxed{\frac{3}{4}}.$$

□

§27.4 [TEXT] Famous example: offset sphere

Recall the famous example in Section 24.3 where we showed that in polar coordinates, we could draw a circle passing through the origin; we called it an "offset circle". There's a 3D analog of this with an offset sphere where you have a sphere that's sitting on the xy-plane. It's actually pretty much exactly the same.

Sample Question

Let $\mathcal{T}$ denote the solid ball of radius 1 centered at $(0,0,1)$. Express the region $\mathcal{T}$ in spherical coordinates.

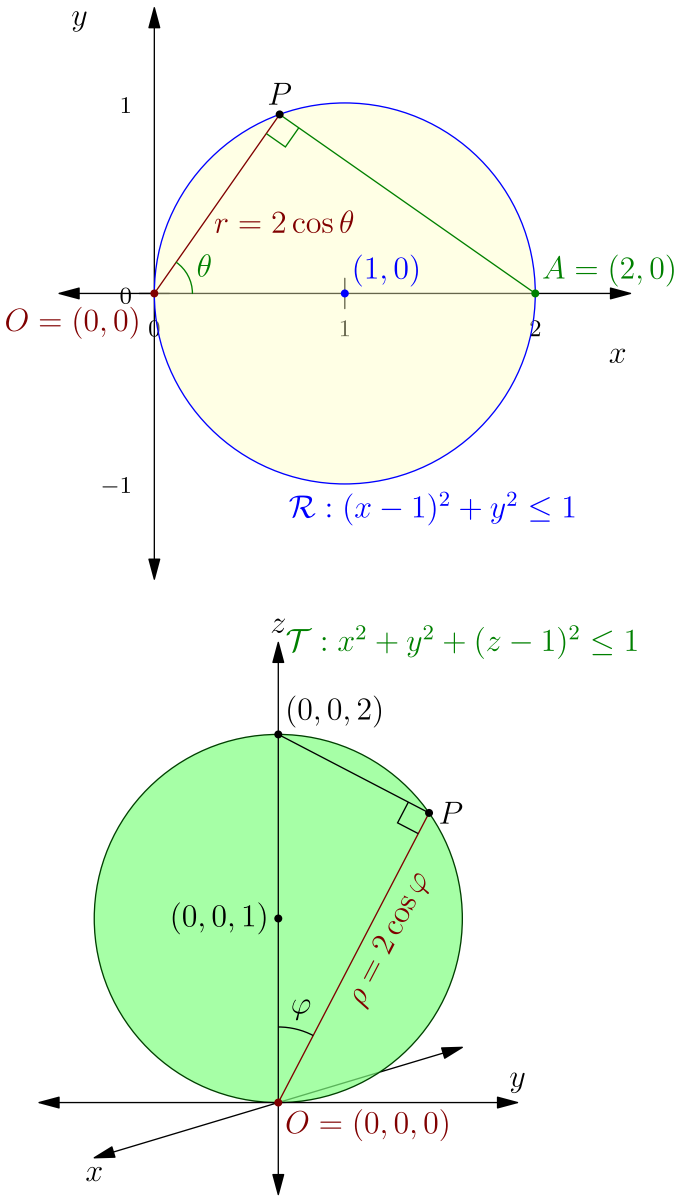

Figure 63: The sketch of $(x-1)^2 + y^2 \le 1$ from Figure 53 is drawn along with a 3D version: the solid ball $x^2 + y^2 + (z-1)^2 \le 1$, which is a unit ball centered at $(0,0,1)$ and lying above and tangent to the xy-plane at $(0,0,0)$. It corresponds to $\rho \le 2\cos\varphi$.

Solution. Like in Section 24.3, we can do either an algebra approach or a geometric one.

Geometric approach See Figure 63. All the points of the sphere lie in the half-space $z \geq 0$ which is described as requiring $0 \leq \varphi \leq \frac{\pi}{2}$. The value of θ is irrelevant by rotational symmetry, and can be anything from 0 to 2π. So we need to figure out how ρ relates to φ.

Let $O = (0,0,0)$ and $A = (0,0,2)$. Let P be a point on the surface of the sphere. Like before, we have

$$\angle P = 90^\circ, \quad OA = 2, \quad \text{and } \angle AOP = \varphi.$$

So the surface of the sphere are those points for which $\rho = 2\cos\varphi$. And the points *inside* the ball are $0 \leq \rho \leq 2\cos\varphi$, accordingly.

Algebraic approach The xyz coordinates of the ball are

$$x^2 + y^2 + (z-1)^2 \leq 1.$$

Recall the spherical coordinates transformation:

$$x = \rho \sin\varphi \cos\theta, \quad y = \rho \sin\varphi \sin\theta, \quad z = \rho\cos\varphi.$$

Substituting these into the sphere's equation:

$$(\rho\sin\varphi\cos\theta)^2 + (\rho\sin\varphi\sin\theta)^2 + (\rho\cos\varphi - 1)^2 \leq 1.$$

Expand and simplify:

$$\begin{aligned} 1 &\geq \rho^2\sin^2\varphi\cos^2\theta + \rho^2\sin^2\varphi\sin^2\theta + (\rho\cos\varphi - 1)^2 \\ &= \rho^2\sin^2\varphi(\cos^2\theta + \sin^2\theta) + (\rho\cos\varphi - 1)^2 \\ &= \rho^2\sin^2\varphi + (\rho\cos\varphi - 1)^2 \\ &= \rho^2\sin^2\varphi + \rho^2\cos^2\varphi - 2\rho\cos\varphi + 1 \\ &= \rho^2(\sin^2\varphi + \cos^2\varphi) - 2\rho\cos\varphi + 1 \\ &= \rho^2 - 2\rho\cos\varphi + 1. \end{aligned}$$

Rearranging, this gives

$$\begin{aligned} &0 \geq \rho^2 - 2\rho\cos\varphi = \rho(\rho - 2\cos\varphi) \\ \iff &0 \leq \rho \leq 2\cos\varphi. \end{aligned}$$

In particular this requires $\cos\varphi \geq 0$ i.e. $\varphi \leq \frac{\pi}{2}$.

In conclusion, the answer is $\mathcal{R}$ in polar coordinates is exactly

$$\boxed{0 \leq \theta < 2\pi \quad \text{and} \quad 0 \leq \varphi \leq \frac{\pi}{2} \quad \text{and} \quad \rho \leq 2\cos\varphi}. \qquad \square$$

The analogous famous exercise in 3D:

Sample Question

Let $\mathcal{T}$ denote the solid ball of radius 1 centered at $(0,0,1)$. Calculate

$$\iiint_{\mathcal{T}} \sqrt{x^2+y^2+z^2}\, dx\, dy\, dz.$$

Solution. As before, if we try to use xyz integration it's a disaster, but spherical coordinates are great because

$$\rho = \sqrt{x^2 + y^2 + z^2}.$$

We just saw that $\mathcal{T}$ is given in spherical coordinates according to $0 \le \varphi \le \frac{\pi}{2}$, $0 \le \theta < 2\pi$, $0 \le \rho \le 2\cos\varphi$. Thus, the integral becomes:

$$\begin{aligned}
\iiint_{\mathcal{T}} \rho \, \mathrm{d}V &= \iiint_{\mathcal{T}} \rho^3 \sin\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \int_{\rho=0}^{2\cos\varphi} \rho^3 \sin\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \sin\varphi \left[\frac{\rho^4}{4} \right]_{\rho=0}^{2\cos\varphi} \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \sin\varphi \cdot (4\cos^4\varphi) \, \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= 4 \int_{\theta=0}^{2\pi} \left[-\frac{1}{5}\cos^5\varphi \right]_{\varphi=0}^{\frac{\pi}{2}} \mathrm{d}\theta \\
&= 4 \int_{\theta=0}^{2\pi} \frac{1}{5} \, \mathrm{d}\theta \\
&= \boxed{\frac{8\pi}{5}}.
\end{aligned}$$

□

§27.5 [TEXT] Spherical coordinates for gravity

Let's go back to the equation for gravity where the components were given by

$$\begin{aligned}
G_1 &:= Gm \iiint_{\mathcal{T}} \frac{x\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac{3}{2}}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\
G_2 &:= Gm \iiint_{\mathcal{T}} \frac{y\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac{3}{2}}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\
G_3 &:= Gm \iiint_{\mathcal{T}} \frac{z\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac{3}{2}}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.
\end{aligned}$$

I didn't do any examples last section because using xyz coordinates when you have $(x^2+y^2+z^2)^{\frac{3}{2}}$ is just way too annoying. However, in spherical coordinates, the equations become much more manageable. For example, the one for G_3 reads:

$$\begin{aligned}
G_3 &= Gm \iiint_{\mathcal{T}} \frac{z\delta(x,y,z)}{(x^2+y^2+z^2)^{\frac{3}{2}}} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z \\
&= Gm \iiint_{\mathcal{T}} \frac{(\rho\cos\varphi)\delta(x,y,z)}{\rho^3} (\rho^2 \sin\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta) \qquad (17) \\
&= Gm \iiint_{\mathcal{T}} \delta(x,y,z) \sin\varphi \cos\varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta.
\end{aligned}$$

Let's see it in action with an offset sphere.

Sample Question

Suppose $\mathcal{T}$ is a metal ball of radius 1 of constant unit density, and P is a point of mass m on its surface. Calculate the magnitude of the force of gravity exerted on the point P.

Solution. We use the offset sphere again: we pick coordinates so that $P = (0,0,0)$ (so the origin is the point P, *not* the center of $\mathcal{T}$). The center of $\mathcal{T}$ will instead be at $(0,0,1)$. Then by symmetry, we have $G_1 = G_2 = 0$, and Equation 17 just says

$$G_3 = Gm \iiint_{\mathcal{T}} \sin\varphi \cos\varphi \,\mathrm{d}\rho \,\mathrm{d}\varphi \,\mathrm{d}\theta$$

after setting the density to 1.

Then we can put in the bounds of integration for the offset sphere:

$$\begin{aligned} G_3 &= Gm \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \int_{\rho=0}^{2\cos\varphi} \sin\varphi\cos\varphi \,\mathrm{d}\rho \,\mathrm{d}\varphi \,\mathrm{d}\theta \\ &= Gm \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} (2\cos\varphi) \cdot \sin\varphi\cos\varphi \,\mathrm{d}\varphi \,\mathrm{d}\theta \\ &= 2Gm \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \cos^2\varphi \sin\varphi \,\mathrm{d}\varphi \,\mathrm{d}\theta \\ &= 2Gm \int_{\theta=0}^{2\pi} \left[-\frac{1}{3}\cos^3\varphi \right]_{\varphi=0}^{\frac{\pi}{2}} \mathrm{d}\theta \\ &= 2Gm \int_{\theta=0}^{2\pi} \frac{1}{3} \,\mathrm{d}\theta \\ &= \frac{4\pi Gm}{3}. \end{aligned}$$

In other words, in the coordinate system we chose, gravity is given by

$$\mathbf{G} = \left\langle 0, 0, \frac{4\pi Gm}{3} \right\rangle.$$

The magnitude is $|\mathbf{G}| = \boxed{\frac{4\pi Gm}{3}}$. □

§27.6 [EXER] Exercises

Exercise 27.1. Consider a solid ball of radius 1 and a line ℓ through its center. Across all points P inside the ball, compute the average value of the distance from P to ℓ. (The average is defined as $\frac{1}{\mathrm{Vol}(\mathcal{T})} \iiint_{\mathcal{T}} d(P) \,\mathrm{d}V$, where $d(P)$ is the distance from P to ℓ.)

Exercise 27.2. Suppose $\mathcal{T}$ is a solid metal hemisphere of radius 1 of constant unit density, and P is a point of mass m at the center of the base of the hemisphere. Calculate the magnitude of the force of gravity exerted on the point P.

Chapter 28. Parametrizing surfaces

§28.1 [TEXT] Parametrizing surfaces

We now move on to parametrizing surfaces. This will require a bit more to get used to compared to parametrizing curves, because now there are two variables instead of one.

To draw a contrast, remember that back when we were parametrizing curves all the way back in Chapter 12, you wrote the notation $\mathbf{r}(t)$ and usually thought of the parameter t as a "time". So you could imagine that a curve in $\mathbb{R}^2$ or $\mathbb{R}^3$ lets you carve out a 1D curve $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ by considering a *timeline* as the input variable. See Figure 18 again.

In contrast, for 2D surfaces in $\mathbb{R}^3$, we are going to need two variables:

Idea

We will describe surfaces as images of some function $\mathbf{r}(u, v) : \mathcal{R} \to \mathbb{R}^3$ where $\mathcal{R}$ is some region in $\mathbb{R}^2$; see Figure 64.

The time analogy breaks down, so I'm going to use a different analogy: gridlines from a map, like longitude and latitude. This is actually going to be the same analogy we used in Section 23.2, when we presented transition maps from change of variables. The only difference is that in Section 23.2, we used 2D paper to plot out a weird region that also lived in 2D space. But when parametrizing a surface, we're going to use 2D paper, represented as region $\mathcal{R}$, to draw a 2D surface that lives in 3D space, which we denote by $\mathcal{S}$.

</> Type signature

To emphasize the types going on here, suppose $\mathcal{S}$ is a surface in 3D space. Then to parametrize a 2D surface you need to specify a 2D region $\mathcal{R}$ in $\mathbb{R}^2$ and then write down a function $\mathbf{r} : \mathcal{R} \to \mathbb{R}^3$ in two variables $\mathbf{r}(u, v)$ for (u, v) in the region $\mathcal{R}$ which covers all the points in $\mathcal{S}$.

Warning

Here $\mathcal{R}$ is a region in $\mathbb{R}^2$ used for the parametrization, often a rectangle. It is *not* the surface $\mathcal{S}$ whose surface area is being calculated; (and for 2D surfaces in 3D space we'll usually prefer the letter $\mathcal{S}$ so that it doesn't look like a region).

A cartoon of the situation is shown in Figure 64. This picture is really important to understand, so take a while to let it sink in.

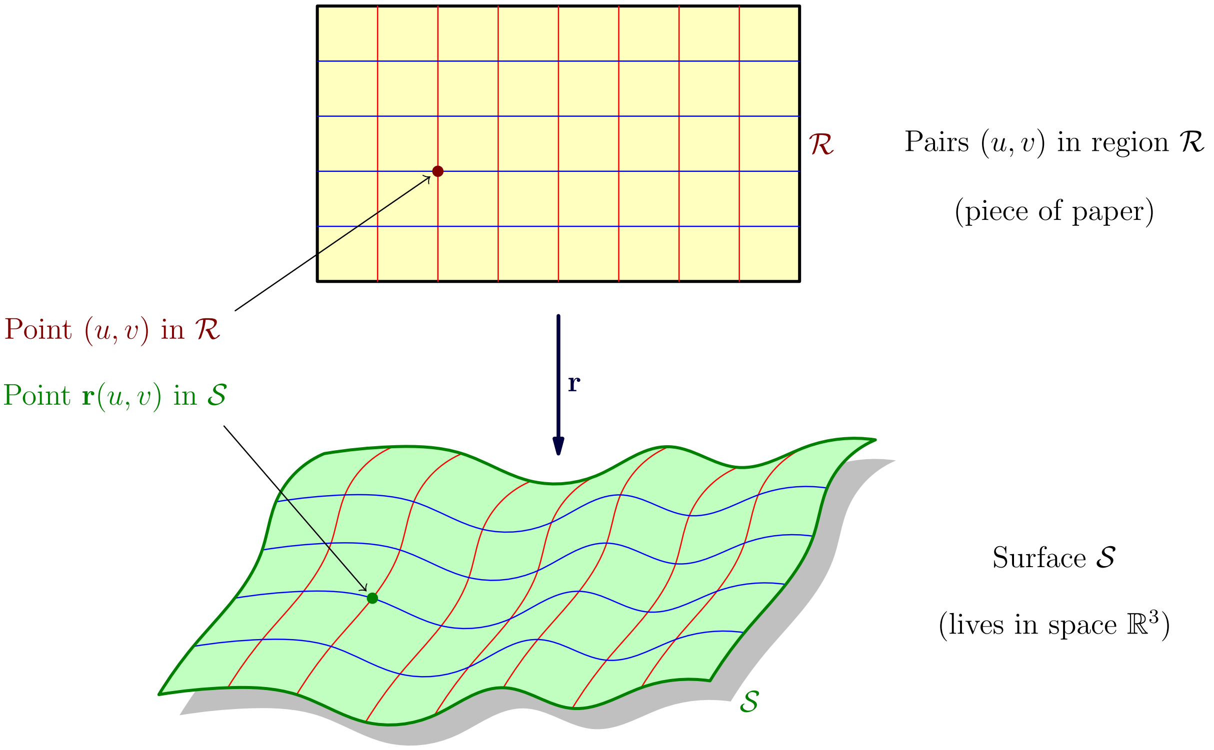

Figure 64: Illustration of how to think of a parametrization conceptually, using cartography. Imagine a piece of paper $\mathcal{R}$ showing the surface $\mathcal{S}$ as it lives in space. (In this cartoon, $\mathcal{S}$ might be described as a mountain range.) A pair (u, v) on the paper could be thought of like longitude and latitude; it should mark some point $\mathbf{r}(u, v)$ on the surface $\mathcal{S}$. Hence we write parametrizations as $\mathbf{r} : \mathcal{R} \to \mathbb{R}^3$ and identify $\mathcal{S}$ with $\mathbf{r}$.

§28.2 [TEXT] Examples of parametrized surfaces

In fact, the Earth is another good example because spherical coordinates gives you a parametrization that uses a rectangular sheet of paper.

Example of a parametrization: the spherical Earth

Consider the surface of the unit sphere, say $x^2 + y^2 + z^2 = 1$. One parametrization $\mathbf{r}$ is given from the spherical coordinate system by

$$\mathbf{r}(\varphi, \theta) = (\sin\varphi \cos\theta, \varphi \sin\theta, \cos\varphi)$$

across the range $0 \le \varphi \le \pi$ and $0 \le \theta < 2\pi$. That is, as θ and φ vary across these ranges, we get every point on the sphere exactly once. See Figure 65.

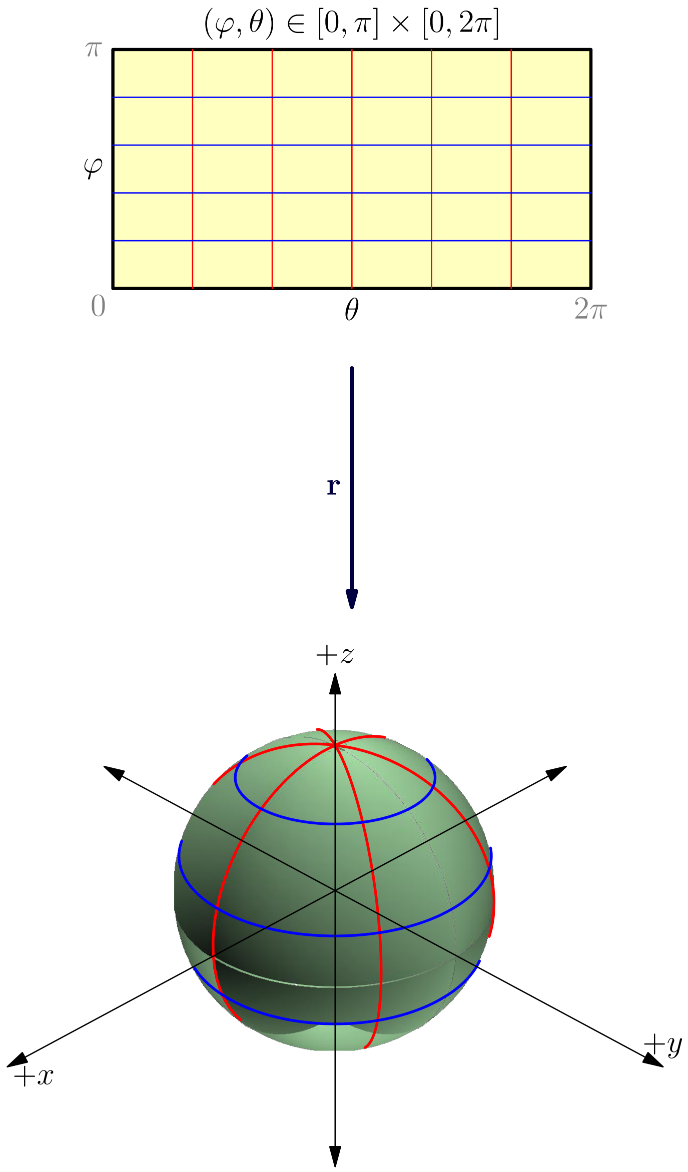

Figure 65: Consider a unit sphere; then the parametrization $\mathbf{r}(\varphi, \theta) = (\sin \varphi \cos \theta, \varphi \sin \theta, \cos \varphi)$ corresponds to longitude and latitude. In this cartoon, one should imagine the yellow sheet of paper being a map of the Earth, drawn in green. The blue and red gridlines on the sheet of paper trace out longitude and latitude lines on the Earth. (The piece of paper is rotated to have θ on the bottom and φ on the left, to make it look a bit more natural.)

If this feels familiar, it's because we used more or less the same analogy for change of variables — cartography. The Earth is round, but you can still draw a rectangular world map. So what we call $\mathbf{r}(u, v)$ here is playing the same role that our transition map $\mathbf{T}$ did back when we did change-of-variables. The only difference is that in change of variables, we had $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$ in the 2D case and $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^3$ in the 3D case. But for parametrizing a surface in $\mathbb{R}^3$, we have $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ instead. (That is, in change-of-variables we make a n-dimensional map of an n-dimensional region, but here we make a 2 -dimensional map of a surface living in $\mathbb{R}^3$.)

Example: A hemisphere whose map is printed on circular paper

In both Figure 64 and Figure 65 we used rectangular paper. But we could easily use non-rectangular paper as well. For example, suppose $\mathcal{R}$ is the region $x^2 + y^2 \le 1$ and we consider the surface

$$\mathbf{r}(x, y) = \left(x, y, \sqrt{1 - (x^2 + y^2)}\right).$$

Then this would give us a parametrization of a *hemisphere*: the part of the sphere $x^2 + y^2 + z^2 = 1$ with $z \ge 0$. Pictorially, this corresponds to drawing a circular map of the Northern hemisphere by taking a birds-eye view from the North pole. See Figure 66.

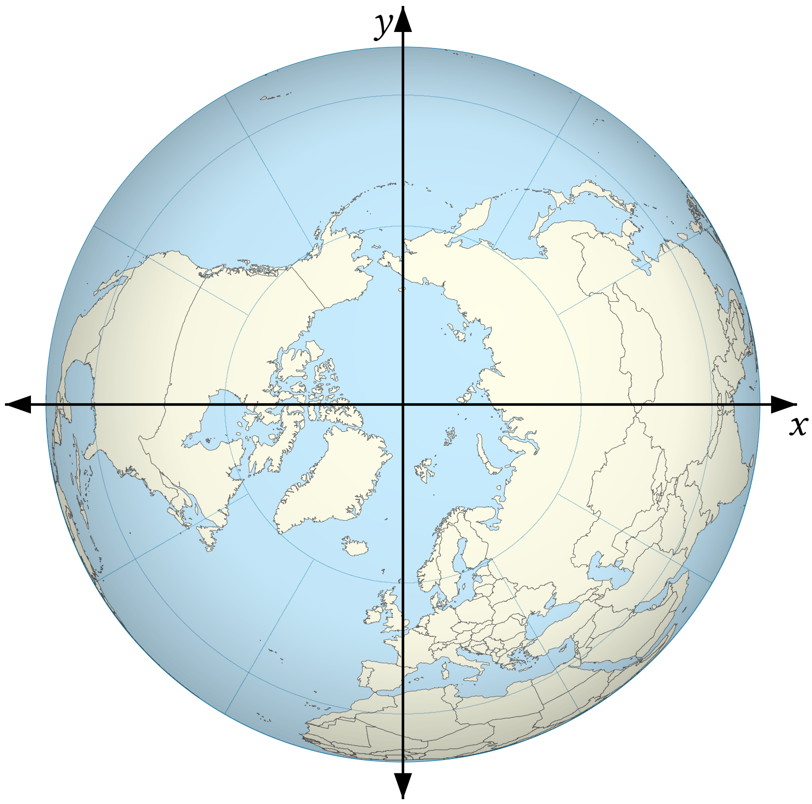

Figure 66: The northern hemisphere of the Earth, drawn on a 2D piece of paper. Image adapted from the public domain.

Remark: Graphs of functions are a common kind of surface

Note in the example we just did in Figure 66, we chose the variable names x and y rather than u and v since they match the x-component and y-component of $\mathbf{r}(x, y)$, giving us fewer different letters to juggle. And we'll do this in general: if our parametrization would *a priori* be written as

$$\mathbf{r}(u, v) = \langle u, v, f(u, v) \rangle$$

for (u, v) in some region $\mathcal{R}$, then we'll usually prefer to use the variable names

$$\mathbf{r}(x, y) = \langle x, y, f(x, y) \rangle$$

instead.

This happens quite often. Such surfaces are sometimes called **graphs** of the function f, because you think of them as the portions of the plots of $z = f(x, y)$ for some function $f : \mathcal{R} \to \mathbb{R}$. We'll give two more examples in a moment.

Another example of a graph: the plane $x + 2y + 3z = 6$ with $x, y, z \geq 0$

As another example, let's consider the part of the plane $x + 2y + 3z = 6$ that lies in the first octant $x, y, z \geq 0$. It's drawn in the right half of Figure 67.

To parametrize it, we'll again use the "graph" idea again: we imagine projecting our surface directly down to the xy-plane (where $z = 0$) and then our piece of paper is whatever points are in the shadow. In this case, our region $\mathcal{R}$ is the part of the xy-plane cut out by $x, y \geq 0$ and $x + 2y \leq 6$, shown in the left half of Figure 67. The equation parametrization is then given exactly by

$$\mathbf{r}(x, y) = \left\langle x, y, \frac{6 - x - 2y}{3} \right\rangle.$$

Optionally, one can draw the region $\mathcal{R}$ into the 3D sketch too; this is the shaded bottom triangle in the right half of Figure 67.

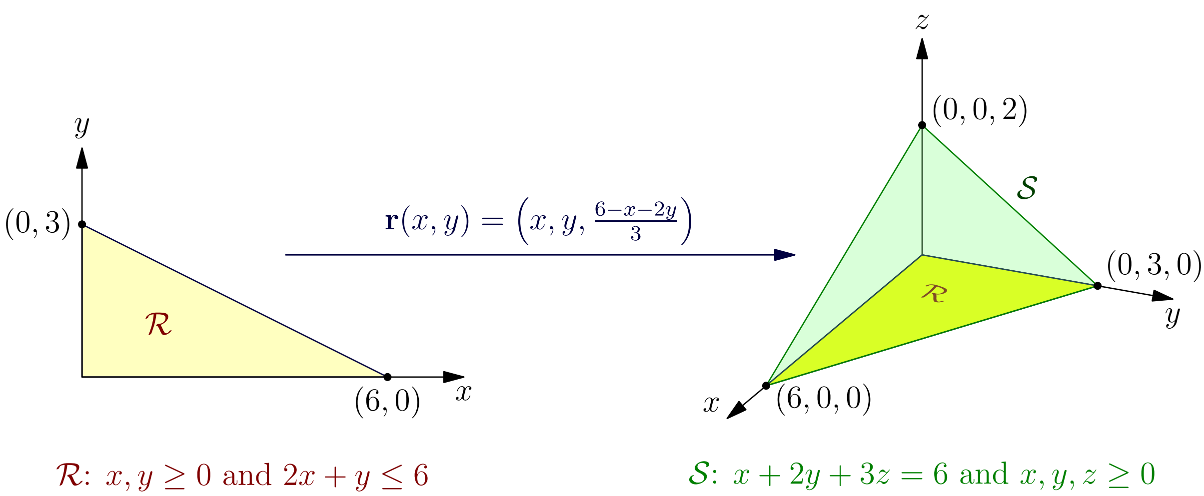

Figure 67: The part of the plane $x + 2y + 3z = 6$ being parametrized by its projection to the xy-plane. Because we're viewing the plane as a graph, we opt to use the letters $\mathbf{r}(x, y)$ rather than $\mathbf{r}(u, v)$.

Example: yz-plane

Consider the entire yz plane in $\mathbb{R}^3$ (that is, the points with $x = 0$). Then one can parametrize it by $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ (so our paper $\mathcal{R} = \mathbb{R}^2$ is infinite!) defined by the equation

$$\mathbf{r}(u, v) = (0, u, v).$$

(I suppose it would've made sense to rename the variables to $\mathbf{r}(y, z) = (0, y, z)$, but it doesn't matter.)

§28.3 [SIDENOTE] Parametrizations are still flexible

Like in Chapter 12, the parametrization of a surface is not unique, and you get flexibility in how you parametrize it. For example, for the simple yz-plane we just did, we give an example of an overly complicated parametrization.

Our piece of paper will be $\mathcal{R} = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ (that is, a square of side length π) and our parametrization $\mathbf{r} : \mathcal{R} \to \mathbb{R}^3$ will be defined by

$$\mathbf{r}(u, v) = (0, (\tan u)^3, \log(e^u + 5) + \tan v).$$

This is really a valid parametrization: you can verify every point in the yz-plane appears exactly once on our map. It even has a region $\mathcal{R}$ with finite area. But it's so ugly you would never want to use it.

§28.4 [EXER] Exercises

Exercise 28.1. Consider a surface $\mathcal{S}$ given by the parametrization $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\mathbf{r}(u, v) = \langle u - v, u + v, uv \rangle$$

for all (u, v) in $\mathbb{R}^2$. Compute the real number k for which the point $P = (2, 8, k)$ lies on $\mathcal{S}$.

Chapter 29. Scalar-field line and surface integrals

Think back to Figure 41. So far we've talked about everything except the three entries labeled "line integral" and "surface integral". This chapter will talk about them. For clarity, I will actually call these **scalar-field line integral** and **scalar-field surface integral**.

The reason for this naming is that later we'll meet vector-field variants of the line and surface integral that play a much bigger role in 18.02. Indeed we mentioned in Section 21.2 that these scalar-field integrals are only used for a few specific cases.

§29.1 [TEXT] Arc length, and its generalization to the scalar-field line integral

We've actually met arc length already back in Part Delta! I'll restate it again here for convenience, but this is a repeat:

Definition: Arc length

If the parametrization $\mathbf{r}(t) : \mathbb{R} \to \mathbb{R}^n$ traces out a path in $\mathbb{R}^n$, the **arc length** is defined as

$$L = \int_{t=\text{start time}}^{\text{stop time}} |\mathbf{r}'(t)| \, \mathrm{d}t.$$

Warning: There are no red arrows for these integrals

We don't like these scalar-field line and surface integrals as much; they just don't behave that well, in part because of the awkward absolute value. For example, Stokes' theorem — the biggest theorem in the 18.02 course — doesn't work for arc length (or anything else in this chapter).

In particular: **you don't get Fundamental Theorem of Calculus for arc length**. To make that warning explicit, note two common "wrong guesses":

$$\int_{t=a}^{b} |\mathbf{r}'(t)| \, \mathrm{d}t \neq |\mathbf{r}(b)| - |\mathbf{r}(a)|$$
$$\int_{t=a}^{b} |\mathbf{r}'(t)| \, \mathrm{d}t \neq |\mathbf{r}(b) - \mathbf{r}(a)|.$$

This is a tempting mistake to make and I've seen it happen; you might hope the fundamental theorem of calculus works somehow for $|\mathbf{r}'(t)|$ in analogy to how $\int_{x=a}^{b} f'(x) \, \mathrm{d}x = f(b) - f(a)$ for differentiable functions $f : \mathbb{R} \to \mathbb{R}$. But that's simply not the case. There's just no analog of FTC for arc length. (Indeed, the arc length on the left-hand side depends on *how* you travel from $\mathbf{r}(a)$ to $\mathbf{r}(b)$ — a straight line will be shortest mileage, a windy meander with detours will be much longer mileage. So you can't possibly know just from the starting point and the destination point how long of a route you took.)

More generally, if the parametrization $\mathbf{r}(t) : \mathbb{R} \to \mathbb{R}^n$ traces out a path in $\mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ is a function, then the **scalar-field line integral** of f is defined by

$$\int_{t=\text{start time}}^{\text{stop time}} f(\mathbf{r}(t)) |\mathbf{r}'(t)| \, \mathrm{d}t.$$

However, we won't use this definition in this class, except for the special case $f = 1$ for arc length.

</> **Type signature**

The surface area of a surface $\mathcal{S}$ is a scalar (and doesn't depend on how the surface is parametrized). The scalar-field line integral is also a scalar.

As I mentioned in the shorthand table (Table 13), many other sources abbreviate

$$\mathrm{d}s := |\mathbf{r}'(t)|\,\mathrm{d}t.$$

Whenever this shorthand is being used, one frequently cuts out the start and stop time too. The way this is done is, you let $\mathcal{C}$ denote the curve that $\mathbf{r}(t)$ traces out. Then we can abbreviate

$$\int_{t=\text{start time}}^{\text{stop time}} f(\mathbf{r}(t))|\mathbf{r}'(t)|\,\mathrm{d}t = \int_{\mathcal{C}} f\,\mathrm{d}s.$$

In particular, taking $f = 1$, the arc length formula gets abridged to $L = \int_{\mathcal{C}} \mathrm{d}s$.

§29.2 [TEXT] Surface area

Okay, so in analogy are surface area and the scalar-field surface integral. We use what we learned about parametrization from Chapter 28.

Definition: Surface area

If the parametrization $\mathbf{r}(u,v) : \mathcal{R} \to \mathbb{R}^3$ cuts out a surface $\mathcal{S}$ in $\mathbb{R}^3$, the **surface area** is given by

$$\operatorname{SurfArea}(\mathcal{S}) := \iint_{\mathcal{R}} \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\,\mathrm{d}v.$$

Yes, there's a cross product. Yes, it sucks (see Section 6.4). This is one case where you probably would prefer to use the shorthand

$$\mathrm{d}S := \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\,\mathrm{d}v$$

so that one can swallow surface area into just

$$\operatorname{SurfArea}(\mathcal{S}) := \iint_{\mathcal{S}} \mathrm{d}S$$

where we also cut out the region $\mathcal{R}$ on our cartographer's map from the notation; instead we write $\mathcal{S}$ directly.

Where does the $\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|$ come from? The way to picture this is via Figure 68 (actually analogous to the picture we drew in Section 24.6 when justifying the Jacobian). If you imagine our region $\mathcal{R}$ as a piece of paper having red and blue gridlines, then $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ correspond to little arrows on the surface along the gridlines on $\mathcal{S}$. But way back when we introduced the cross product, it had a geometric definition that stated:

- The magnitude of the cross product corresponds to the area of the little "cell" on the surface in the gridlines, shaded in Figure 68. So when we add all of them, we should get the surface area!

- The direction of the cross product is perpendicular to both the horizontal and vertical gridlines, so in fact the cross product should be thought of as *normal* to the surface. Right now we don't care about this yet, but it'll matter later on in Chapter 38.

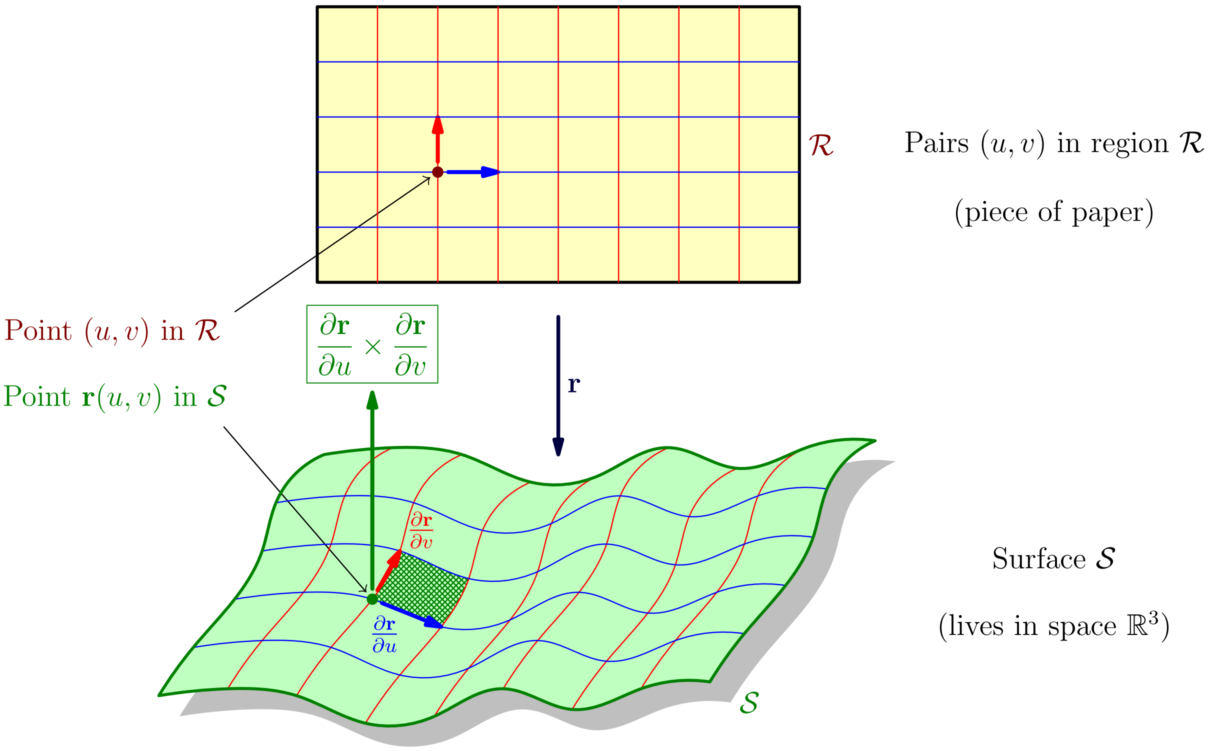

Figure 68: Consider surface $\mathcal{S}$ parametrized by $\mathbf{r} : \mathcal{R} \to \mathbb{R}^3$. The cross product of the two partial derivatives is drawn in green. The magnitude of the cross product corresponds to the small shaded area.

More generally if we have a function $f : \mathbb{R}^3 \to \mathbb{R}$ we could define the **scalar-field surface integral** of f over $\mathcal{S}$ as $\iint_{\mathcal{R}} f(\mathbf{r}(u, v)) \left| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right| \, \mathrm{d}u \, \mathrm{d}v$; however this definition will not be used in this class except for the special case $f = 1$ for surface area. But if we did use it, we could have an abbreviation $\iint_{\mathcal{S}} f \, \mathrm{d}S$.

> </> **Type signature**
>
> The scalar-field surface integral (and hence surface area as well) outputs a scalar.

§29.3 [RECIPE] Surface area (done directly)

Here's surface area in recipe format.

Recipe for surface area, manually

1. Parametrize the surface by some $\mathbf{r}(u, v) : \mathcal{R} \to \mathbb{R}^3$ for some 2D region $\mathcal{R}$ (ideally something simple like a circle or rectangle).
2. Compute the partial derivatives $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ (both are three-dimensional vectors at each point).
3. Compute the cross product $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$.
4. Compute the magnitude $\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|$ of this cross product.
5. Integrate the entire thing over $\mathcal{R}$ using any of the methods for double integrals (such as horizontal/vertical slicing, polar coordinates, change of variables, etc.).

Tip: We'll make a table of common cross products next chapter

For this chapter we'll compute the cross product by hand in the recipe above. However, this will get tedious really quickly. So in the next chapter, Chapter 30, we're actually just going to calculate all the cross products for most "common" cases all in one place, and refer to it later.

Here is a really ugly example to start, to give you some practice with spherical coordinates.

Example: Surface area of a sphere

Compute the surface area of the unit sphere $x^2 + y^2 + z^2 = 1$.

Solution. We will bludgeon our way through this task with sheer brute force using the formula above via spherical coordinates. (We'll show a more elegant solution later in Section 30.3.)

The parametrization $\mathbf{r}$ is given from the spherical coordinate system by

$$\mathbf{r}(\varphi, \theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi),$$

across the range

$$0 \le \theta \le 2\pi \quad \text{and} \quad 0 \le \varphi \le \pi$$

for our region $\mathcal{R}$. The partial derivatives are thus

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial \varphi} &= \langle \cos\varphi\cos\theta, \cos\varphi\sin\theta, -\sin\varphi \rangle \\
\frac{\partial \mathbf{r}}{\partial \theta} &= \langle -\sin\varphi\sin\theta, \sin\varphi\cos\theta, 0 \rangle.
\end{aligned}$$

We brute force our way through the entire cross product. We have

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \cos\varphi\cos\theta & \cos\varphi\sin\theta & -\sin\varphi \\ -\sin\varphi\sin\theta & \sin\varphi\cos\theta & 0 \end{vmatrix} \\
&= (0\cdot\cos\varphi\sin\theta - \sin\varphi\cos\theta\cdot(-\sin\varphi))\mathbf{e_1} \\
&\quad -(0\cdot\cos\varphi\cos\theta - (-\sin\varphi\sin\theta)\cdot(-\sin\varphi))\mathbf{e}_2 \\
&\quad +(\sin\varphi\cos\theta\cdot\cos\varphi\cos\theta + \sin\varphi\sin\theta\cdot\cos\varphi\sin\theta)\mathbf{e}_3 \\
&= (\sin^2\varphi\cos\theta)\mathbf{e}_1 + (\sin^2\varphi\sin\theta)\mathbf{e}_2 + (\sin\varphi\cos\varphi\sin^2\theta + \sin\varphi\cos\varphi\cos^2\theta)\mathbf{e}_3 \\
&= (\sin^2\varphi\cos\theta)\mathbf{e}_1 + (\sin^2\varphi\sin\theta)\mathbf{e}_2 + (\sin\varphi\cos\varphi)\mathbf{e}_3
\end{aligned}$$

since $\cos^2\theta + \sin^2\theta = 1$. If we take the magnitude ,we get

$$\begin{aligned}
\left|\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta}\right| &= \sqrt{(\sin^2\varphi\cos\theta)^2 + (\sin^2\varphi\sin\theta)^2 + (\sin\varphi\cos\varphi)^2} \\
&= \sqrt{\sin^4\varphi\cos^2\theta + \sin^4\varphi\sin^2\theta + \sin^2\varphi\cos^2\varphi} \\
&= \sqrt{\sin^4\varphi(\cos^2\theta + \sin^2\theta) + \sin^2\varphi\cos^2\varphi} \\
&= \sqrt{\sin^4\varphi\cdot 1 + \sin^2\varphi\cos^2\varphi} \\
&= \sqrt{\sin^2\varphi(\sin^2\varphi + \cos^2\varphi)} \\
&= \sqrt{\sin^2\varphi\cdot 1} = |\sin\varphi|.
\end{aligned}$$

Thank the lord it's a simple answer. Great, now we can calculate the surface area of the sphere:

$$\begin{aligned}
\operatorname{SurfArea}(\text{sphere}) &= \int_{\theta=0}^{2\pi}\int_{\varphi=0}^{\pi} \left|\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi}\right| \, \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi}\int_{\varphi=0}^{\pi} |\sin\varphi| \, \mathrm{d}\varphi \, \mathrm{d}\theta \\
&= \left(\int_{\varphi=0}^{\pi} |\sin\varphi| \, \mathrm{d}\varphi\right)\left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta\right) \\
&= \left(\int_{\varphi=0}^{\pi} \sin\varphi \, \mathrm{d}\varphi\right)\left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta\right) \\
&= [-\cos\varphi]_{\varphi=0}^{\pi}\cdot 2\pi \\
&= \boxed{4\pi}.
\end{aligned}$$

$\square$

Digression on the direction of the cross product

As we said earlier when drawing Figure 68, in general if you parametrize a surface $\mathcal{S}$ by $\mathbf{r}(u, v)$, then $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ is a vector which is normal to both $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$. That is, the direction of this cross product is the normal vector to the tangent plane of the surface $\mathcal{S}$ at $\mathbf{r}(u, v)$.

Of course, since we took an absolute value, the direction gets discarded for surface area. But if you are really observant you might have noticed that the computed cross product is

$$(\sin^2 \varphi \cos \theta)\mathbf{e}_1 + (\sin^2 \varphi \sin \theta)\mathbf{e}_2 + (\sin \varphi \cos \varphi)\mathbf{e}_3 = \sin \varphi \cdot \mathbf{r}(\theta, \varphi)$$

which happens to be a multiple of the corresponding point on the sphere; and this is why, because for a sphere, $\mathbf{r}(\theta, \varphi)$ happens to be perpendicular to the tangent plane.

And here is an example that is a little less computationally intensive.

Sample Question

Compute the surface area of the cone defined by $z = \sqrt{x^2 + y^2} \le 1$.

Solution. The given cone can be parametrized using Cartesian coordinates as:

$$\mathbf{r}(x, y) = \left(x, y, \sqrt{x^2 + y^2}\right)$$

where (x, y) lies within the disk $x^2 + y^2 \le 1$.

Compute the partial derivatives of $\mathbf{r}$ with respect to x and y:

$$\frac{\partial \mathbf{r}}{\partial x} = \left\langle \frac{\partial x}{\partial x}, \frac{\partial y}{\partial x}, \frac{\partial z}{\partial x} \right\rangle = \left\langle 1, 0, \frac{x}{\sqrt{x^2 + y^2}} \right\rangle$$
$$\frac{\partial \mathbf{r}}{\partial y} = \left\langle \frac{\partial x}{\partial y}, \frac{\partial y}{\partial y}, \frac{\partial z}{\partial y} \right\rangle = \left\langle 0, 1, \frac{y}{\sqrt{x^2 + y^2}} \right\rangle.$$

Hence the cross product is

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & \frac{x}{\sqrt{x^2+y^2}} \\ 0 & 1 & \frac{y}{\sqrt{x^2+y^2}} \end{vmatrix} \\
&= \left(0 \cdot \frac{y}{\sqrt{x^2 + y^2}} - 1 \cdot \frac{x}{\sqrt{x^2 + y^2}}\right)\mathbf{e}_1 - \left(1 \cdot \frac{y}{\sqrt{x^2 + y^2}} - 0 \cdot \frac{x}{\sqrt{x^2 + y^2}}\right)\mathbf{e}_2 \\
&\qquad + (1 \cdot 1 - 0 \cdot 0)\mathbf{e}_3 \\
&= \left\langle -\frac{x}{\sqrt{x^2 + y^2}}, -\frac{y}{\sqrt{x^2 + y^2}}, 1 \right\rangle
\end{aligned}$$

Now, compute the magnitude of this cross product:

$$\begin{aligned}\left|\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y}\right| &= \sqrt{\left(-\frac{x}{\sqrt{x^2+y^2}}\right)^2 + \left(-\frac{y}{\sqrt{x^2+y^2}}\right)^2 + 1^2} \\ &= \sqrt{\frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2} + 1} = \sqrt{2}.\end{aligned}$$

That's really convenient: we got a constant! Hence

$$\operatorname{SurfArea}(\text{cone}) = \iint_{x^2+y^2\le 1} \sqrt{2}\,\mathrm{d}A = \sqrt{2}\operatorname{Area}(x^2+y^2 \le 1) = \boxed{\sqrt{2}\pi}. \qquad \square$$

§29.4 [EXER] Exercises

Exercise 29.1. Consider a surface $\mathcal{S}$ given by the parametrization $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\mathbf{r}(u,v) = \langle u - v, u + v, uv \rangle$$

for all (u,v) in $\mathbb{R}^2$. Compute the tangent plane to $\mathcal{S}$ at the point $(3, 7, 10)$.

Chapter 30. Pre-computed cross products for common surfaces

TL;DR: cross products are too annoying, so we pre-compute them all.

§30.1 [TEXT] Pre-computed formulas for the cross product

As the examples last chapter show, it's actually really annoying to compute the cross product by hand. Consequently, we can make our lives a lot easier if we pre-compute what the cross product works out to for some common situations, so we don't have to redo it by hand every time we need it.

In these notes we will pre-compute five different cross products:

- Any *graph*, i.e. a surface of the form $z = f(x, y)$ (the cone we discussed is a good example).
- Any level surface $g(x, y, z) = c$, over some xy-region.
- A flat surface in the xy-plane (which could also be yz or zx parallel).
- The curved part of a cylinder of radius R centered along the z-axis, where the parameters are θ and z.
- The surface of a sphere of radius R centered at the origin, where the parameters are φ and θ.

As it turns out, in 18.02 it's likely these are the *only* five situations you will see.

The table showing the results is Table 14. Note that for surface area, you only need the *absolute value* of the cross product (fourth column). But I'm going to include the entire vector too, because we'll later need to reuse this table in Chapter 38 (where one reformats it as Table 23). At that point, we will actually need to know the direction the vector points in too, not just the absolute value.

Surface	**Param's**	$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$	$\left\lvert \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\rvert \, \mathrm{d}u \, \mathrm{d}v$
$z = f(x, y)$	(x, y)	$\left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle$	$\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \, \mathrm{d}x \, \mathrm{d}y$
Level surface $g(x, y, z) = c$ over an xy-region	(x, y)	$\frac{\nabla g}{\partial g / \partial z}$	$\frac{\lvert \nabla g \rvert}{\lvert \partial g / \partial z \rvert} \, \mathrm{d}x \, \mathrm{d}y$
Flat surface $z = c$	(x, y)	$\langle 0, 0, 1 \rangle$	$\mathrm{d}x \, \mathrm{d}y$
Cylindrical coords with fixed R $\mathbf{r}(\theta, z) = (R \cos\theta, R \sin\theta, z)$	(θ, z)	$\langle R \cos\theta, R \sin\theta, 0 \rangle$	$R \, \mathrm{d}\theta \, \mathrm{d}z$
Spherical coords with fixed R $\mathbf{r}(\varphi, \theta) = (R \sin\varphi \cos\theta,$ $R \sin\varphi \sin\theta, R \cos\varphi)$	(φ, θ)	$R \sin\varphi \cdot \mathbf{r}(\varphi, \theta)$	$R^2 \sin\varphi \, \mathrm{d}\varphi \, \mathrm{d}\theta$ (if $0 \le \varphi \le \pi$)

Table 14: Pre-computed formulas for the cross product in five most common situations, which are likely to be all you need.

Recall the following geometric idea from the earlier Figure 64, when we described where the cross product was coming from:

> **Idea**
>
> The vector $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ is normal to the tangent plane to the surface at each point.

This can help a lot with remembering the third column of Table 14. For example:

- For the level surface $g(x, y, z) = c$, you should remember from Chapter 15 that ∇g is normal to the tangent plane of the level surface, hence the cross product is a multiple of ∇g as needed.
- The normal vector to (the curved part of) a cylinder points straight away from the z-axis away from the origin, which $\langle R\cos\theta, R\sin\theta, 0\rangle$ indeed does.
- For the sphere, the normal vector should point straight away from the center of the sphere, and indeed $\sin(\varphi) \cdot \mathbf{r}(\varphi, \theta)$ is a multiple of the direction.

Again, to re-iterate: for surface area you actually only need the fourth column $\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|$, but

- I think $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ is actually easier to remember than $\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|$, because of the geometric interpretation above;
- starting from Chapter 38 you will need the full data of $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ in the third column.

The first and second rows of Table 14 above are quite versatile, so in these notes I'll call them "magic" formulas because they save us so much work. In contrast, the other rows are for more specialized situations.

§30.2 [TEXT] Table 14 row 1: For a graph (surface of the form $z = f(x, y)$)

So imagine your surface is given by $z = f(x, y)$ for some f over some region $\mathcal{R}$ in the xy plane (e.g. the cone had $\mathcal{R} = \{x^2 + y^2 \le 1\}$). (e.g. the cone we just did was $f(x, y) = \sqrt{x^2 + y^2}$) What we're going to do is try to capture the boilerplate work of the cross product into a single formula that we can just remember, so we don't have to redo the cross product again.

The parametrization we expect to use is

$$\mathbf{r}(x, y) = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}.$$

The partial derivatives are

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial x} &= \begin{pmatrix} \frac{\partial x}{\partial x} \\ \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial z}{\partial x} \end{pmatrix} \\ \frac{\partial \mathbf{r}}{\partial y} &= \begin{pmatrix} \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial y} \\ \frac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial z}{\partial y} \end{pmatrix}. \end{aligned}$$

Hence, in this case we arrive at

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & \frac{\partial z}{\partial x} \\ 0 & 1 & \frac{\partial z}{\partial y} \end{vmatrix} \\ &= -\frac{\partial z}{\partial x}\mathbf{e}_1 - \frac{\partial z}{\partial y}\mathbf{e}_2 + \mathbf{e}_3 \\ &= -\frac{\partial f}{\partial x}\mathbf{e}_1 - \frac{\partial f}{\partial y}\mathbf{e}_2 + \mathbf{e}_3. \end{aligned}$$

Let's write this down now.

Memorize: Magic cross product formula for graphs $z = f(x, y)$

Consider a surface given by $z = f(x, y)$ with f differentiable. Then for the obvious parametrization $\mathbf{r}(x, y) = (x, y, f(x, y))$ we have

$$\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = \left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle.$$

In particular, the surface area becomes

$$\operatorname{SurfArea}(\mathcal{S}) = \iint_{\mathcal{R}} \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dx\, dy.$$

You'll find this formula written in a lot of other textbooks and it's worth knowing (I would say you should memorize the full magic cross product formula, since it's trivial to get the magnitude from it.) Let's see how it can captures the boilerplate in the cone example.

Sample Question

Compute the surface area of the cone defined by $z = \sqrt{x^2 + y^2} \le 1$.

Solution. Letting $f(x, y) = \sqrt{x^2 + y^2}$, this time we skip straight to

$$\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}$$
$$\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}.$$

Hence we got a shortcut to the vector $\langle -\frac{x}{\sqrt{x^2+y^2}}, -\frac{y}{\sqrt{x^2+y^2}}, 1 \rangle$ we found before. We find its magnitude in the same way:

$$\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} = \sqrt{1 + \frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2}} = \sqrt{2}.$$

Now

$$\operatorname{SurfArea}(\text{cone}) = \iint_{x^2+y^2 \le 1} \sqrt{2}\, dA = \sqrt{2}\operatorname{Area}(x^2 + y^2 \le 1) = \boxed{\sqrt{2}\pi}.$$

□

§30.3 [TEXT] Table 14 row 2: For a level surface $g(x, y, z) = c$

However, we can get an even better formula in a lot of cases using implicit differentiation. The basic idea is that we would prefer to think of the cone as $x^2 + y^2 - z^2 = 0$, so that we don't need to think about square roots. And that's exactly a level surface.

So imagining a *level surface* $g(x, y, z) = c$ instead, where each (x, y) in our region $\mathcal{R}$ has exactly one $z = z(x, y)$ value. On paper, you imagine solving for z in terms of x and y, and then using

$$\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = -\frac{\partial z}{\partial x}\mathbf{e}_1 - \frac{\partial z}{\partial y}\mathbf{e}_2 + \mathbf{e}_3$$

but we'd like to not have to solve for z in such a brute way.

The trick is to consider the gradient of g and use the chain rule. You might remember that

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

is pretty easy to calculate, usually. However, if we take the partial derivative of

$$g(x, y, z) = c$$

with respect to x and y, the derivative of c vanishes while the chain rule gives

$$0 = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial x} \implies \frac{\partial g}{\partial x} = -\frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial x}.$$

Similarly $\frac{\partial g}{\partial y} = -\frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial y}$. Hence

$$\nabla g = \left\langle -\frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial x}, -\frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial y}, \frac{\partial g}{\partial z} \right\rangle = \frac{\partial g}{\partial z} \cdot \left\langle -\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y}, 1 \right\rangle.$$

Digression on the chain rule

You might be spooked by the minus sign here, as I was, since if you just look at the fractions the expression seems wrong. This is why I don't like to remember the chain rule as just "cancel the fractions", because in some contexts you'll get equations like this that don't seem correct.

The context to remember here is that $z = z(x, y)$ is itself a function of x and y that holds on to the requirement $g(x, y, z(x, y)) = c$; that is, if x changes a little, $z = z(x, y)$ should change in an "opposite" way to ensure $g = c$ is still true.

How much should the change be? It might be easiest to reason through two applications of linear approximation. If ε is some small displacement, then linear approximation is saying that

$$\begin{aligned} g(x+\varepsilon, y, z(x+\varepsilon, y)) &\approx g\left(x+\varepsilon, y, z(x, y) + \frac{\partial z}{\partial x} \cdot \varepsilon\right) \\ &\approx g(x, y, z(x, y)) + \nabla g \cdot \begin{pmatrix} \varepsilon \\ 0 \\ \frac{\partial z}{\partial x} \cdot \varepsilon \end{pmatrix} \\ &= g(x, y, z(x, y)) + \left[\frac{\partial g}{\partial x} + \frac{\partial g}{\partial z} \cdot \frac{\partial z}{\partial x}\right]\varepsilon. \end{aligned}$$

Hence we want the bracketed coefficient of ε to be zero, which is the equation we got before.

Something really good is happening here, because the cross product we wanted just sits on the right-hand side! Because of this, we have managed to derive the following miraculous identity.

Memorize: Magic cross product formula for a level surface

Let g be differentiable and consider the level surface $g(x,y,z)=c$. Let $\mathcal{S}$ be a part of this level surface described implicitly by some function $z=f(x,y)$, and suppose also that $\frac{\partial g}{\partial z}\neq 0$ over $\mathcal{R}$. Then for the obvious parametrization $\mathbf{r}(x,y)=(x,y,f(x,y))$ we have

$$\frac{\partial \mathbf{r}}{\partial x}\times\frac{\partial \mathbf{r}}{\partial y}=\frac{\nabla g}{\partial g/\partial z}.$$

The reason this magic identity is even better is that there is no need to differentiate f or even to determine it. Let's see it in action by redoing our example with a cone.

Sample Question

Compute the surface area of the cone defined by $z=\sqrt{x^2+y^2}\leq 1$.

Solution. The cone is the part of the level surface of $g(x,y,z)=x^2+y^2-z^2$ with $z\geq 0$. (We know in fact $f(x,y)=\sqrt{x^2+y^2}$, but we won't use this.) Now we can jump straight to

$$\frac{\nabla g}{\frac{\partial g}{\partial z}}=\frac{\langle 2x,2y,-2z\rangle}{-2z}=\left\langle -\frac{x}{z},-\frac{y}{z},1\right\rangle.$$

The magnitude of this vector is

$$\sqrt{\left(-\frac{x}{z}\right)^2+\left(-\frac{y}{z}\right)^2+1}=\sqrt{\frac{x^2+y^2}{z^2}+1}=\sqrt{2}$$

so we get

$$\operatorname{SurfArea}(\text{cone})=\iint_{x^2+y^2\leq 1}\sqrt{2}\,\mathrm{d}A=\sqrt{2}\operatorname{Area}(\{x^2+y^2\leq 1\})=\boxed{\sqrt{2}\pi}.$$ □

If you compare this carefully with $z=\sqrt{x^2+y^2}$, you'll see this is *still* the same solution as the first magic formula, which is in turn *still* the same solution as when we really used bare hands. But the shortcuts are nice because it means you don't have to think about the cross product at all.

Now as we promised, let's show how to find surface area for a sphere without having to slog through the pain of spherical coordinates.

Sample Question

Compute the surface area of the sphere $x^2+y^2+z^2=1$.

Solution. We'll find the surface area for the hemisphere with $z\geq 0$ and then double it. We could view the hemisphere as $z=f(x,y)=\sqrt{1-(x^2+y^2)}$, but to avoid square roots we're much happier by letting

$$g(x,y,z)=x^2+y^2+z^2$$

and considering the hemisphere as the chunk of the level surface with $z\geq 0$ and $x^2+y^2\leq 1$. In that case,

$$\frac{\nabla g}{\partial g/\partial z} = \frac{\langle 2x, 2y, 2z \rangle}{2z} = \left\langle \frac{x}{z}, \frac{y}{z}, 1 \right\rangle.$$

This time the magnitude of the vector is

$$\sqrt{\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 + 1} = \sqrt{\frac{x^2+y^2+z^2}{z^2}} = \frac{1}{z} = \frac{1}{\sqrt{1-(x^2+y^2)}}.$$

Hence, we need to integrate

$$\operatorname{SurfArea}(\text{hemisphere}) = \iint_{x^2+y^2\le 1} \frac{1}{\sqrt{1-(x^2+y^2)}} \, dx \, dy.$$

To nobody's surprise, we use polar coordinates to change this to

$$\begin{aligned} \operatorname{SurfArea}(\text{hemisphere}) &= \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \frac{1}{\sqrt{1-r^2}} (r \, dr \, d\theta) \\ &= \left(\int_{\theta=0}^{2\pi} d\theta \right) \left(\int_{r=0}^{1} \frac{r}{\sqrt{1-r^2}} \, dr \right). \end{aligned}$$

The left integral is 2π. For the inner integral, use the u-substitution $u = 1 - r^2 \implies \frac{du}{dr} = -2r$ to get

$$\int_{r=0}^{1} \frac{r}{\sqrt{1-r^2}} \, dr = \int_{u=1}^{0} -\frac{1}{2} u^{-\frac{1}{2}} \, du = \int_{u=0}^{1} \frac{1}{2} u^{-\frac{1}{2}} \, du = \left[u^{\frac{1}{2}} \right]_{u=0}^{1} = 1.$$

Hence

$$\operatorname{SurfArea}(\text{hemisphere}) = 2\pi \cdot 1 = 2\pi$$

and the surface area of the sphere is thus $2\pi \cdot 2 = \boxed{4\pi}$. □

§30.4 [TEXT] Table 14 row 3: For a flat surface

This is the really easy special case of $z = f(x, y)$ when $f(x, y) = c$ is constant. Your parametrization is just

$$\mathbf{r}(x, y) = \langle x, y, c \rangle.$$

I hesitated to include this row because it's so easy and is a special case of the first row, but it's common enough I decided I might as well toss it in. However, you should have no problem deriving this yourself even in your sleep; it's literally

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial x} &= \langle 1, 0, 0 \rangle \\ \frac{\partial \mathbf{r}}{\partial y} &= \langle 0, 1, 0 \rangle \end{aligned}$$

and the cross product of these is $\langle 0, 0, 1 \rangle$, so there you go.

Note that you might encounter flat surfaces parallel to the xz or yz planes instead, in which case you should just swap the roles of the variables.

§30.5 [TEXT] Table 14 row 4: For the curved part of the cylinder in cylindrical coordinates

If you have a cylinder aligned with the z-axis, then you don't want to be using xy-plane as parameters, because most pairs (x, y) do not get used at all. Thus, we'll instead use cylindrical coordinates as

$$\mathbf{r}(\theta, z) = (R\cos\theta, R\sin\theta, z).$$

Compute the partial derivatives:

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \theta} &= \langle -R\sin\theta, R\cos\theta, 0\rangle \\ \frac{\partial \mathbf{r}}{\partial z} &= \langle 0, 0, 1\rangle.\end{aligned}$$

The cross product is pretty easy to evaluate in this case:

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial z} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -R\sin\theta & R\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= R\cos\theta \mathbf{e}_1 + R\sin\theta \mathbf{e}_2 \\ &= \langle R\cos\theta, R\sin\theta, 0\rangle.\end{aligned}$$

That's not too bad! We can take the absolute value of this:

$$\mathrm{d}S = R\,\mathrm{d}\theta\,\mathrm{d}z.$$

Tip: $\mathrm{d}S$ for the cylinder can be remembered geometrically

The way to remember this is that "$\mathrm{d}S\,\mathrm{d}r \approx \mathrm{d}V$": if you multiply a bit of surface by a bit of the radial component, you get a chunk of volume of the sphere. And since we saw in Chapter 26 that $\mathrm{d}V = r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z$, the formula for $\mathrm{d}S$ is what you get when you divide out $\mathrm{d}r$ and set $r = R$.

Alternatively, each patch on the cylinder can be thought of as a little rectangle of height $\mathrm{d}z$ and width $R\,\mathrm{d}\theta$.

§30.6 [TEXT] Table 14 row 5: For the curved part of the sphere, in spherical coordinates

We already saw the sphere is actually handled by our magic formula for level surfaces, so if you're fine using xy-coordinates you are good to go. Nonetheless, in the event you need spherical coordinates, here is the result.

We actually computed this already while working out the sphere's surface area by brute force: if we take the parametrization

$$\mathbf{r}(\varphi, \theta) = (R\sin\varphi\cos\theta, R\sin\varphi\sin\theta, R\cos\varphi),$$

then if we repeat the brutal calculation from Section 29.3 with an extra R tacked on, we get

$$\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} = (R^2\sin^2\varphi\cos\theta)\mathbf{e}_1 + (R^2\sin^2\varphi\sin\theta)\mathbf{e}_2 + (R^2\sin\varphi\cos\varphi)\mathbf{e}_3.$$

This formula might look ugly until you realize that it's actually just

$$\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} = R \sin \varphi \cdot \mathbf{r}(\varphi, \theta).$$

Since $|\mathbf{r}(\varphi, \theta)| = R$, we get

$$\mathrm{d}S := \left| \frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} \right| d\varphi d\theta = R^2 \sin \varphi \, \mathrm{d}\varphi \, \mathrm{d}\theta.$$

Here I'm dropping the absolute value bars around $\sin \varphi$ because our spherical coordinate convention requires $0 \le \varphi \le \pi$.

Tip: $\mathrm{d}S$ for the sphere can be remembered geometrically

The way to remember this is that "$\mathrm{d}S \, \mathrm{d}\rho \approx \mathrm{d}V$": if you multiply a bit of surface by a bit of the radial component, you get a chunk of volume of the sphere. And since we saw in Chapter 27 that $\mathrm{d}V = \rho^2 \sin \varphi \, \mathrm{d}\rho \, \mathrm{d}\varphi \, \mathrm{d}\theta$, the formula for $\mathrm{d}S$ is what you get when you divide out $\mathrm{d}\rho$ and set $\rho = R$.

§30.7 [RECIPE] Recap of surface area

Let's write a new recipe for surface area now that we have Table 14.

Recipe for surface area upgraded with Table 14

To compute the surface area of a surface $\mathcal{S}$:

1. Get the cross product $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ for a parametrization $\mathbf{r}$ using the following checklist.
   - If you are using (x, y)-coordinates to parametrize (meaning $\mathcal{S}$ is a graph $z = f(x, y)$ or a level surface), use the magic formulas in rows 1 or 2 of Table 14.
   - For a flat surface, it's easy (row 3 of Table 14).
   - If $\mathcal{S}$ is specifically given by cylindrical/spherical coordinates with fixed radius, use rows 4 or 5 of Table 14.
   - Otherwise, evaluate the cross product manually:
     - Pick a parametrization $\mathbf{r}(u, v) : \mathcal{R} \to \mathbb{R}^3$ of the surface $\mathcal{S}$. Sort of like in Section 12.7, you have some freedom in how you set the parametrization.
     - Compute $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ (both are three-dimensional vectors at each point).
     - Compute the cross product $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ as in Chapter 6.
2. Take the magnitude of the cross product to get a number for each point on the surface.
3. Integrate it over $\mathcal{R}$ using any of the methods for double integrals (such as horizontal/vertical slicing, polar coordinates, change of variables, etc.).

§30.8 [EXER] Exercises

Exercise 30.1. Compute the surface area of the surface defined by $z = x^2 + y^2 \le 1$.

Exercise 30.2 (Archimedes hat-box theorem)**.** Let $-1 < a < b < 1$ be real numbers. Consider the unit sphere $x^2 + y^2 + z^2 = 1$ and the cylinder $x^2 + y^2 = 1$. Show that the portions of their (lateral) surface areas which lie between $z = a$ and $z = b$ have equal area. See Figure 69.

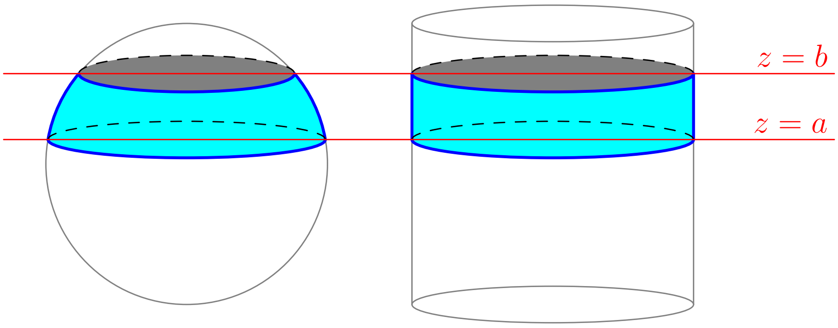

Figure 69: Figure for Exercise 30.2. Show that the two blue lateral surface areas are equal.

Part India: Line integrals of vector fields over a curve

For comparison, Part India corresponds to §14, §15, §20, §21 of Poonen's notes.

Chapter 31. Vector fields

§31.1 [TEXT] Vector fields

In Part Golf, we only considered integrals of scalar-valued functions. However, in Part India and Juliett we will meet a **vector field**, which is another name for a function that inputs points and outputs vectors.

Definition

A **vector field** for $\mathbb{R}^n$ is a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ that assigns to each point $P \in \mathbb{R}^n$ a vector $\mathbf{F}(P) \in \mathbb{R}^n$.

In contrast, we might use the word **scalar field** for the old kind of functions $f : \mathbb{R}^n \to \mathbb{R}$ that gave a number at each input point rather than a vector.

You actually have met a lot of vector fields before:

Example

Every gradient is an example of a vector field! That is, if $f : \mathbb{R}^n \to \mathbb{R}$, then ∇f is a vector field for $\mathbb{R}^n$.

This case is so important that there's a word for it:

Definition

A vector field for $\mathbb{R}^n$ is called **conservative** if it happens to equal ∇f for some function $f : \mathbb{R}^n \to \mathbb{R}$.

In Part India and Juliett we'll meet vector fields that aren't conservative too.

Type signature

For standalone vector fields, we'll always use capital bold letters like $\mathbf{F}$ to denote them. That said, remember ∇f is *also* a vector field. So that's why the operator ∇ itself is typeset a little bit bold.

Like the gradient, you should draw inputs to $\mathbf{F}$ as *points* (dots) but the outputs as *vectors* (arrows). Don't mix them.

For 2D and 3D vector fields, we'll often write

$$\mathbf{F} = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix} \quad \mathbf{F} = \begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}.$$

I think other sources often use P, Q, R instead, but right now I'm using those for points too, so I'll use lowercase letters.

§31.2 [TEXT] How do we picture a vector field?

There's a lot of ways to picture a vector field, especially in physics. For consistency, I'm going to pick *one* such framework and write all my examples in terms of it. So **in my book, all examples will be aquatic** in nature; but if you can't swim[24], you should feel free to substitute your own. Imagine an electric field. Or a black hole in outer space. Or air currents in the atmosphere. Whatever works for you!

Anyway, for my book, we'll use the following picture:

Idea

Imagine a flowing body of water (ocean, river, whirlpool, fountain, etc.) in $\mathbb{R}^2$ or $\mathbb{R}^3$. Then at any point P, we draw a tiny arrow $\mathbf{F}(P)$ indicating the direction and speed of the water at the point P. You could imagine if you put a little ball at the point P, the current would move the ball along that arrow.

Sounds a lot like the gradient, right? Indeed, conservative vector fields are a big family of vector fields, and so we should expect they fit this picture pretty neatly. But the thing about conservative vector fields is this: ∇f, as a vector field, is always rushing *towards* whatever makes the value of f bigger. Whereas generic vector fields might, for example, go in loops. Let's put these examples into aquatic terms.

Example of a conservative vector field: going downstream a river

Let's imagine we have a river with a strong current. We'll make the important assumption that the river only goes one way: that is, if you go along the current, you never end up back where you started. In real life, this often occurs if the river goes down a mountain, so as you go down the river you're losing elevation.

If you do this, you can define a "downstream function" $f : \mathbb{R}^3 \to \mathbb{R}$ as follows: for every point P in the river, $f(P)$ measures how far downstream you are. For example, if the river had a head, maybe we could assign f the value zero there, and then f would increase as you get farther from the bank, reaching the largest value at the mouth. (For mountainous rivers, f might instead be thought of as decreasing in elevation.)

Then the vector field corresponding to the river is the gradient ∇f. Remember, the gradient ∇f tells you what direction to move in to increase f. And if you throw a ball into a river, its motion could be described simply as: the ball moves downstream.

[24] Doesn't MIT make you pass a swim test, though?

Figure 70: A river flowing from the base of a waterfall. Black arrows point along the direction of the gradient.

Example of a non-conservative vector field: a whirlpool

Now imagine instead you have a whirlpool. If you throw a ball in it, it goes in circles around vertex of the whirlpool. This doesn't look anything like the river! If you have a river, you never expect a ball to come back to the same point after a while, because it's trying to go downstream. But with a whirlpool, you keep going in circles over and over.

If you draw the vector field corresponding to a whirlpool, it looks like lots of concentric rings made by tiny arrows. That's an example of a non-conservative vector field.

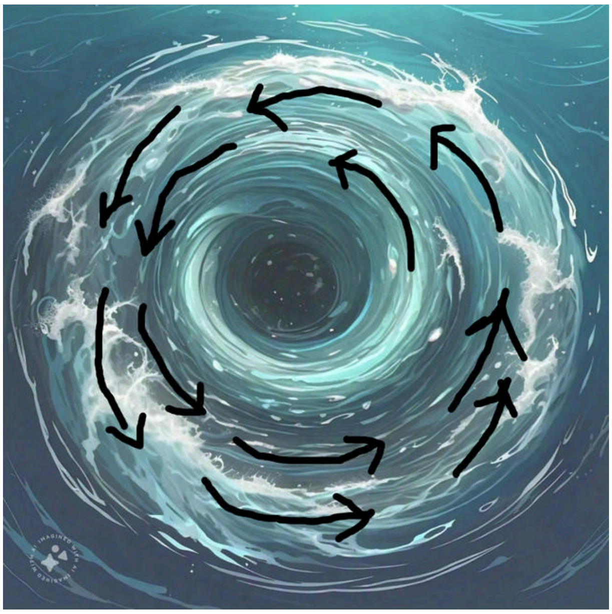

Figure 71: A whirlpool. Round and round we go. Not a conservative vector field.

§31.3 [TEXT] How to draw a vector fields

If you're actually given a formula for $\mathbf{F}(x, y)$, you can sketch the corresponding vector field by the following procedure.

Recipe for sketching a vector field

To draw a cartoon of a vector field:

1. Pick a bunch of points P that you want to draw the arrows at.
2. For each P, draw a little arrow starting at P in the same direction as $\mathbf{F}(P)$. (To make the cartoon readable, you usually scale down the magnitude of the arrow.)

In practice, to make the pictures not look absurd, people will typically draw the arrows a lot smaller than they really are. For example, if you get that $\mathbf{F}(10, 10) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, then strictly speaking the arrow starting at $(10, 10)$ should reach all the way to $(12, 13)$. But if you do this, your cartoon image looks absurd. So people will typically adjust the arrow sizes in the cartoon to be a *tiny* arrow, still pointing the right way, but with much smaller magnitude. For the cartoon it's usually more important the *relative* size of the arrows is correct; drawing the absolute values to scale is unnecessary.

The classic easy-to-draw example is $\mathbf{F} = \begin{pmatrix} x \\ y \end{pmatrix}$, in which for every point P, one just points straight away from the origin. See Figure 72 for that. For the record, in this figure (and the other figures in this section), the length of all the little arrows is scaled exactly 30% compared to the true length.

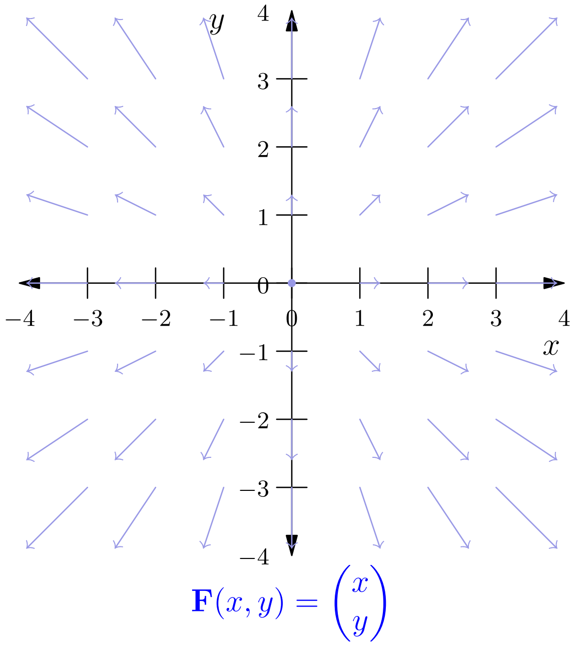

Figure 72: A plot of the first vector field that's drawn in every class, good old $\mathbf{F}(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$.

Okay, here are eight more pictures to train your instincts. For each one, try to pick a few points like $(2, 0)$ or $(3, -1)$ and so on and verify that the arrow starting at that point points in the way you expect.

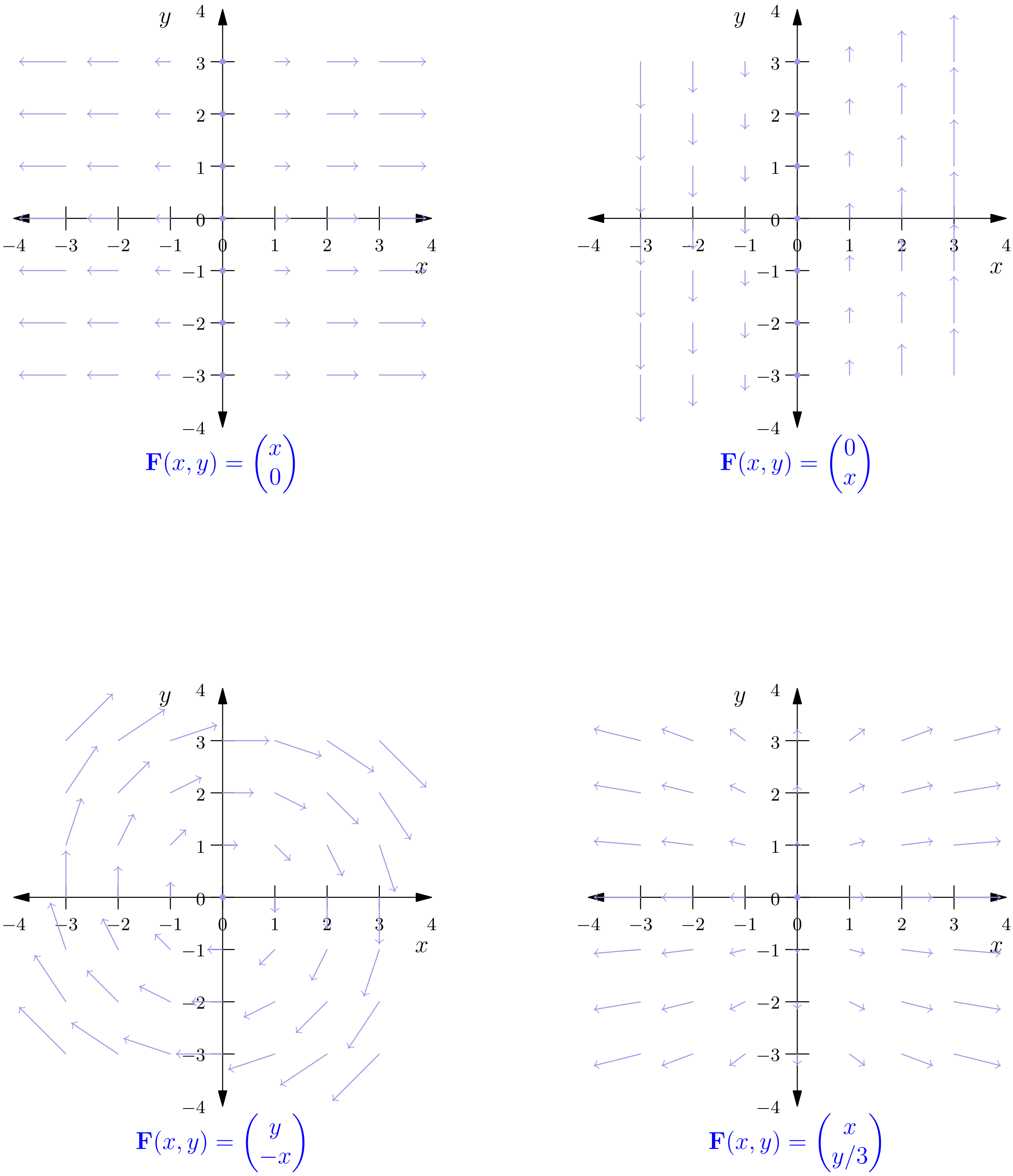

Figure 73: Four more examples of vector fields in $\mathbb{R}^2$. The example $\mathbf{F}(x,y) = \left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right)$ has just horizontal arrows that get longer as $|x|$ grows. The example $\mathbf{F}(x,y) = \left(\begin{smallmatrix} 0 \\ x \end{smallmatrix}\right)$ has just vertical arrows, similar story. The example $\mathbf{F}(x,y) = \left(\begin{smallmatrix} y \\ -x \end{smallmatrix}\right)$ is swirly. The example $\mathbf{F}(x,y) = \left(\begin{smallmatrix} x \\ \frac{y}{3} \end{smallmatrix}\right)$ is slanted.

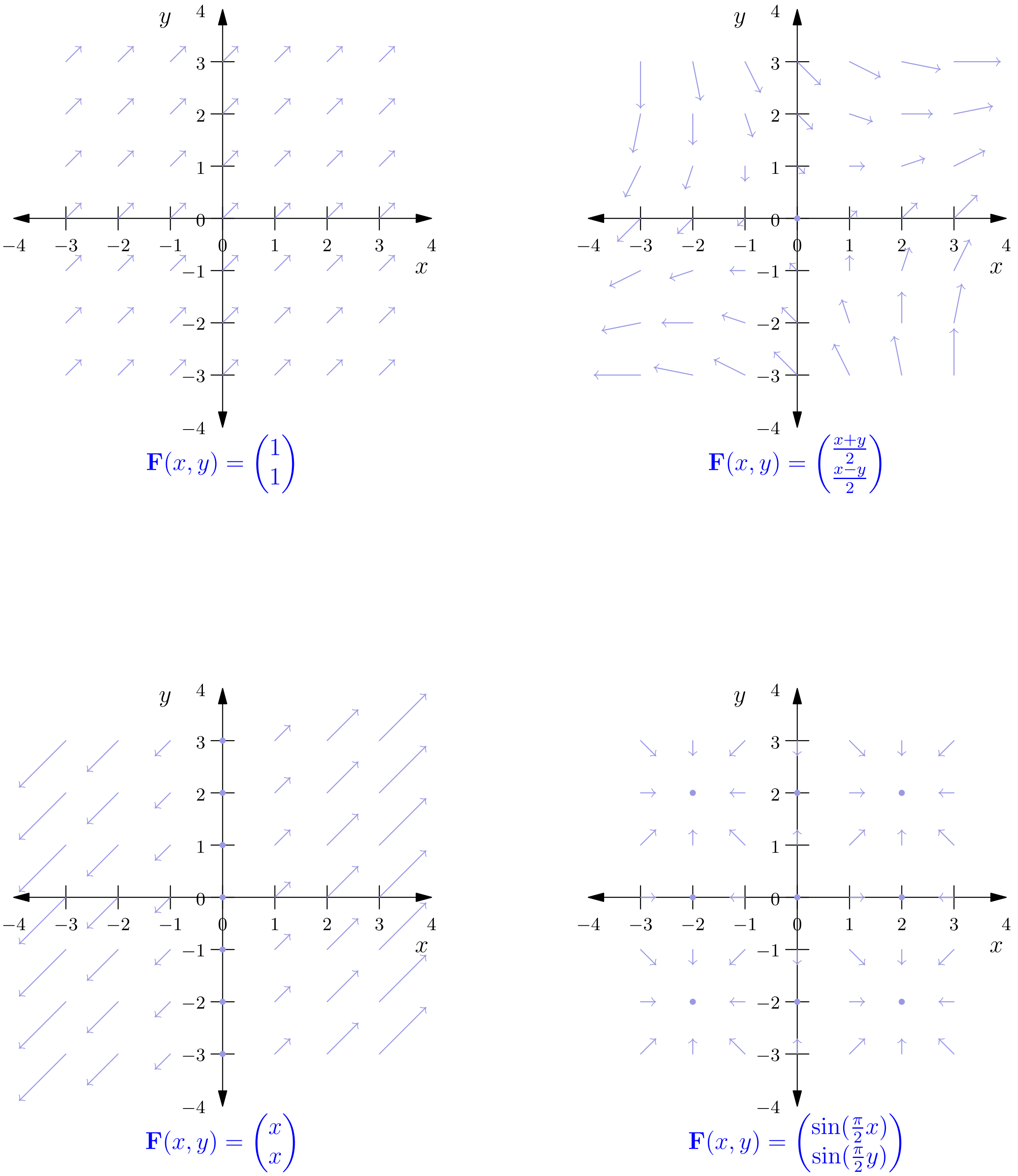

Figure 74: Another four examples of vector fields in $\mathbb{R}^2$. The vector field $\mathbf{F}(x, y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is just constant. The vector field $\mathbf{F}(x, y) = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix}$ has some tilted arrows. The vector field $\mathbf{F}(x, y) = \begin{pmatrix} x \\ x \end{pmatrix}$ has all arrows at $45°$ angles, growing in length with larger $|x|$. And the goofy $\mathbf{F}(x, y) = \begin{pmatrix} \sin(\frac{\pi}{2}x) \\ \sin(\frac{\pi}{2}y) \end{pmatrix}$ has an oscillating behavior.

§31.4 [TEXT] Preview of integration over vector fields

As we mentioned in Section 21.2, the line integral and surface integral we encountered in Part Golf (which had a scalar-valued function) are actually the ugly ducklings that don't get used. For most cases, if you are doing a line integral or surface integral, you actually want **vector-valued line and surface integrals**, where one takes a line or surface integral over the entire field.

That's when the type signatures go crazy.

In order for this to be even remotely memorable, what I'm going to do is augment the previous Figure 41 with pictures corresponding to the situations in which we might integrate a vector field. The new chart can also be downloaded as a large PDF version.

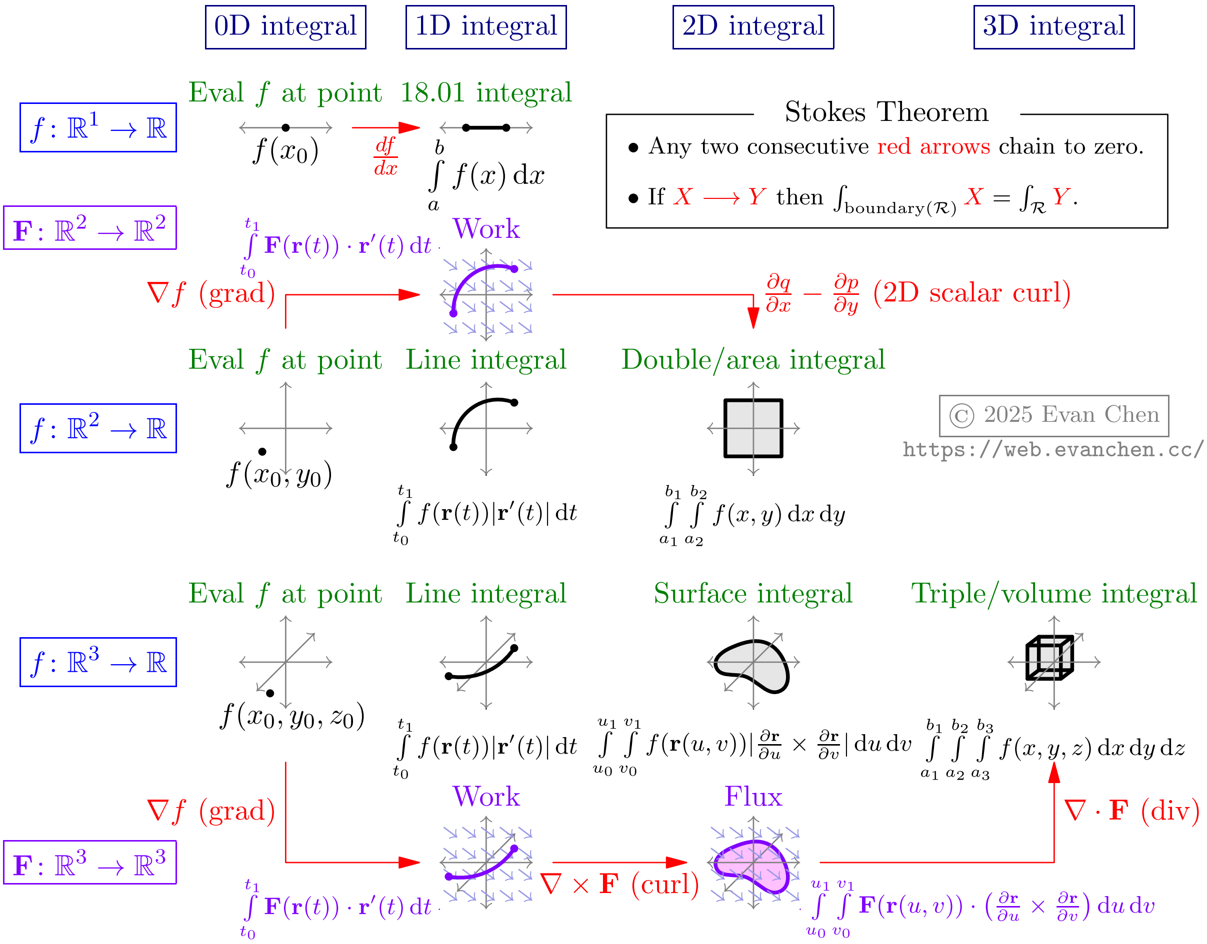

Figure 75: Upgraded Figure 41 with vector fields. Download at `https://web.evanchen.cc/textbooks/poster-stokes.pdf`.

There are two new features of Figure 75 compared to the old version: the three purple pictures and the six red arrows. We'll define them all over the next few sections, so just a few words now.

§31.4.1 The three purple pictures

There are **three new pictures in purple**: they are **work** (covered in Part India) and **flux** (covered in Part Juliett). Basically, these are the only two situations in which we'll be integrating over a vector field:

- We have a path along a vector field and want to measure the *work* of the vector field *along* that path (in the physics sense); this is the focus of Part India.
- We have a surface in a 3D vector field and want to measure the *flux* of the vector field *through* the surface; this is the focus of Part Juliett.

These terms will be defined next chapter.

> **</> Type signature**
>
> The new purple things are *still* all scalar quantities, i.e. work and flux are both numbers, not vectors.

As we mentioned in Section 21.2, the purple pictures will basically replace the corresponding green ones. Conversely, vector fields will usually only be integrated in the situations described in the purple picture. This is summarized in Table 15.

	For scalar fields $f : \mathbb{R}^n \to \mathbb{R}$	For vector fields $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$
18.01 integral	✅ Used all the time	❌ Never used in 18.02
Line integral	🔵 Only for arc length	✅ Used all the time (work)
Double/Area integral	✅ Used all the time	❌ Never used in 18.02
Surface integral	🔵 Only for surface area	✅ Used all the time (flux)
Triple/Volume integral	✅ Used all the time	❌ Never used (except in Equation 12)

Table 15: What the various kinds of integrals are used for. The integrals with ❌ markers never appear in 18.02, so they don't appear in the chart Figure 75 either. (In 18.02 it could happen that line and surface integrals are used for scalar fields if you need mass of a wire or metal plate, but that's quite rare I think.)

This bears repeating:

> **💡 Idea**
>
> We'll pretty much not use the scalar field integrals besides for arc length and surface area. In other words, you can mostly ignore the green pictures in the poster Figure 75 that got replaced by purple ones.

§31.4.2 The six red arrows

There are also **six new red arrows**. They indicate transformations on functions: a way to take one type of function and use it to build another function.

For example, the gradient ∇ is the one we've discussed: if you start with a scalar-valued function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient creates into a vector field $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$. (The $\frac{\mathrm{d}f}{\mathrm{d}x}$ in the $f : \mathbb{R}^1 \to \mathbb{R}$ case is also just the gradient, though a bit more degenerate.)

We'll soon meet three more transformations:

- **2D curl**, which converts a vector field on $\mathbb{R}^2$ *back* into a scalar-valued function;
- **3D curl**, which converts a vector field on $\mathbb{R}^3$ into *another* vector field;
- **divergence**, which converts a vector field on $\mathbb{R}^3$ *back* into a scalar-valued function.

§31.5 [TEXT] Foreshadowing of Stokes' theorem

All the red arrows have two important properties I will tell you *right now*. Because I haven't talked much about any of the red arrows yet, **you won't be able to understand what this means yet**. That's okay. We'll go over what this means for each individual red arrow means when we get to it. However, I want tell you *in advance* that every time we meet a red arrow, there will be a case of Stokes' theorem that applies to it.

Memorize: Two red arrows gives zero

In Figure 75, if you follow two red arrows consecutively, you get zero.

Memorize: Generalized Stokes' Theorem, for 18.02

In Figure 75, take any of the six red arrows $X \to Y$. Let $\mathcal{R}$ be a compact region. Then the integral of X over the **boundary** of $\mathcal{R}$ equals the integral of Y over $\mathcal{R}$:

$$\int_{\text{boundary}(\mathcal{R})} X = \int_{\mathcal{R}} Y.$$

In fact let me tell you what generalized Stokes' theorem says, in vague non-precise terms (we'll make precise later), for each of these six red arrows:

Evaluation → 18.01 integral For the $\frac{\mathrm{d}f}{\mathrm{d}x}$ arrow joining evaluation to the 18.01 integral, it's the fundamental theorem of calculus. The region $\mathcal{R}$ is the line segment $[a, b]$, and the boundary is the two endpoints a and b. Then we have the *fundamental theorem of calculus* from 18.01:

$$f(b) - f(a) = \int_a^b \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x.$$

Evaluation → line integral (× 2) There are two such red arrows, but the statement is the same for both. Suppose $\mathbf{r}(t)$ parametrizes a path joining point P to Q (say, the line segment PQ, or some more curvy path). The region $\mathcal{R}$ is this path, and the endpoints are P and Q: Then we get the *fundamental theorem of calculus for line integrals*:

$$f(Q) - f(P) = \int_{t=\text{start time}}^{\text{stop time}} \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t.$$

The right-hand side is the work done by ∇f on the path $\mathbf{r}$.

If you use shorthand where $\mathcal{C}$ is the curve formed by $\mathbf{r}$, this could be rewritten as

$$f(Q) - f(P) = \int_{\mathcal{C}} \nabla f \cdot \mathrm{d}\mathbf{r}.$$

Line integral → double/area integral Suppose now $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ is a vector field. Write $\mathbf{F} = (p(x, y), q(x, y))$. Let $\mathcal{R}$ be some two dimensional region, like a disk. Suppose further that the *boundary* of $\mathcal{R}$ is parametrized by a curve $\mathbf{r}(t)$ (e.g. the circumference of the disk). Then *Green's theorem* says that

$$\int_{t=\text{start time}}^{\text{stop time}} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y.$$

The weird expression $\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}$ in the right-hand side is called the 2D scalar curl, but we haven't defined this term yet.

If you use shorthand as in Table 13, this can be simplified. Let $\mathcal{C}$ be the curve formed by $\mathbf{r}$, this could be rewritten as

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A.$$

There's a second form of Green's theorem I'll show you when I get to it.

Line integral → surface integral Suppose now $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field. Let $\mathbf{r}_2(u, v)$ parametrizes some two dimensional surface $\mathcal{S}$ (like a metal sheet), and suppose further that the *boundary* of this surface is parametrized by a curve $\mathbf{r}_1(t)$ (e.g. the edges of the sheet). Then the classical version of *Stokes' theorem* says that

$$\int_{t=\text{start time}}^{\text{stop time}} \mathbf{F}(\mathbf{r}_1(t)) \cdot \mathbf{r}_1'(t) \, \mathrm{d}t = \iint_{\mathcal{S}} (\nabla \times \mathbf{F})(\mathbf{r}_2(u, v)) \cdot \left(\frac{\partial \mathbf{r}_2}{\partial u} \times \frac{\partial \mathbf{r}_2}{\partial v} \right) \mathrm{d}u \, \mathrm{d}v.$$

The nonsense expression $\nabla \times \mathbf{F}$ is called the curl, defined next chapter in Chapter 32.

The shorthand version following Table 13 is much easier to read, because the shorthand $\mathbf{n} \, \mathrm{d}S$ stands for the entire hunk $\left(\frac{\partial \mathbf{r}_2}{\partial u} \times \frac{\partial \mathbf{r}_2}{\partial v} \right) \mathrm{d}u \, \mathrm{d}v$. Suppose the curve for $\mathbf{r}_1$ is denoted $\mathcal{C}$. Then the above equation compresses all the way down to

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}_1 = \iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n} \, \mathrm{d}S.$$

so yes, that does save a lot of characters.

Surface integral → triple/volume integral Suppose now $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field. Let $\mathcal{T}$ be some three-dimensional solid (e.g. metal ball). Suppose further the boundary of $\mathcal{R}$ is parametrized by some two-dimensional surface $\mathbf{r}(u, v)$ (e.g. metal sphere), which we call $\mathcal{S}$. Then the *divergence theorem* says that

$$\iint_{\mathcal{S}} \mathbf{F}(\mathbf{r}(u, v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \mathrm{d}u \, \mathrm{d}v = \iiint_{\mathcal{T}} (\nabla \cdot \mathbf{F})(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z.$$

If we adopt shorthand again, this reads just

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S = \iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V.$$

The nonsense expression $\nabla \cdot \mathbf{F}$ is called the divergence, defined next chapter in Chapter 32.

Again, these bullets will not make sense to you yet (except the first one, which is the 18.01 fundamental theorem of calculus), because there are several undefined terms. Instead, treat this as a template for all the theorem statements you are going to learn soon:

> 💡 **Idea**
>
> Every red arrow in the poster Figure 75 has an associated Stokes result.

Know this going in to each of the concepts to follow.

Figure 76: It's all Stokes' theorem.

§31.6 [TEXT] Stay determined

This is probably super overwhelming right now, and Figure 75 might be frightening to look at because there's so much information in it. Don't worry, we'll take Figure 75 apart one piece at a time over the rest of the semester.

- In Chapter 32, I'll tell you how to compute each of grad, curl, div. This chapter has no integration in it, so if you hate integration, you'll like this chapter.
- I'll start talking about work in Chapter 33. I'll define it, and then I'll show you how it ties in to the fundamental theorem of calculus (which are some of the cases of generalized Stokes' theorem).
- Then Chapter 38 will define flux. This will let us talk about the rest of the cases of generalized Stokes' theorem.

§31.7 [EXER] Exercises

Exercise 31.1. Take a few deep breaths, touch some grass, and have a nice drink of water, so that you can look at Figure 75 without feeling fear.

Exercise 31.2. Print out a copy of the high-resolution version of Figure 75 (which can be downloaded at `https://web.evanchen.cc/textbooks/poster-stokes.pdf`) and hang it in your room.

Chapter 32. Grad, curl, and div

The goal of this chapter is to *define* each individual red arrow in the poster. For each red arrow, we'll show you

- How to compute it, and
- How to visualize it in an aquatic setting.

There is no integration in this chapter, and so it's actually pretty straightforward.

Red arrow	Symbol	Input type	Output at each point	Example input	Example output
Gradient	∇f	Scalar field	Vector	Measure of distance from top of waterfall	Waterfall current pointing to lower elevation
Curl	$\nabla \times f$	3D vec. field	3D vector	Whirlpool current	Arrow aligned with rotation axis, magnitude is rotation speed
Divergence	$\nabla \cdot f$	3D vec. field	Scalar	Water flow	Pump/drain speed
2D curl	$\frac{\partial g}{\partial y} - \frac{\partial f}{\partial x}$ where $\mathbf{F} = (f, g)$	2D vec. field	Scalar	Whirlpool current	Angular velocity
2D div	$\nabla \cdot f$	2D vec. field	Scalar	Water flow	Pump/drain speed

Table 16: The red arrows, plus an extra 2D div that's a modified version of 2D curl.

§32.1 [SIDENOTE] Aquatics are unlikely to improve your exam score

By the way, a quick word about aquatics. For each of these, I'm trying to tell you how to think of the quantity in terms of real life. This may help you internalize and remember the results. However, on the actual 18.02 exam, you will find that most of the functions you are taking curl's or divergence's of are rather artificial functions. So your aquatic intuition is more or less useless for actually doing calculation.

It's kind of like how, you were told in 18.01 that derivatives measured rate of change. But then on the calculus final exam you were asked things like "differentiate $f(x) = \sin(e^x)^2$". It probably wasn't much help to know that $f'(x)$ was the rate of change of f, because the function f is completely artificial and would never appear in real life. The question was really testing whether you can apply a recipe with the chain rule to get $f'(x) = 2e^x \sin(e^x) \cos(e^x)$.

The same is true in this chapter. Exam questions about grad, curl, div tend to use artificial functions. So the aquatic intuition is not going to be directly helpful and you just need to be good at following the recipe.

For this reason, in these notes, I'm not even going to bother trying to explain where the curl and div formulas come from. Many have tried and many have failed. If you want to see the grown-ups discuss this, see `https://mathoverflow.net/q/21881/70654`, where the top comment is "My advice: at this level, stick *strictly* to the textbook".

§32.2 [TEXT] Gradient

You already know how to do this from Chapter 15. The function f assigns some number to every point in $\mathbb{R}^n$, and then ∇f points in the direction that f increases most rapidly. In our aquatic examples, you could imagine you have a waterfall, f measures the distance from the top of the waterfall, and ∇f just points straight down.

§32.3 [TEXT] Curl

Here's the definition of curl in 3D space.

Definition of curl

Suppose

$$\mathbf{F}(x, y, z) = \begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}$$

is a 3D vector field. Then the **curl** of $\mathbf{F}$ is the vector field defined by

$$\operatorname{curl} \mathbf{F} := \nabla \times \mathbf{F} := \begin{pmatrix} \frac{\partial r}{\partial y} - \frac{\partial q}{\partial z} \\ \frac{\partial p}{\partial z} - \frac{\partial r}{\partial x} \\ \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \end{pmatrix}.$$

Type signature

The curl takes in only a 3D vector field. The curl at each point is a 3D vector (i.e. the curl of a 3D vector field is itself a 3D vector field).

Tip: How to memorize curl

In practice, everyone remembers this formula using the following mnemonic:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ p & q & r \end{vmatrix}.$$

This equation does not pass type-safety, because it's a "matrix" whose entries are some combination of functions, vectors, and partial derivative operators, so it absolutely does not make sense. Nonetheless, if you ignore all the type safety warnings and try to "expand" this expression, you will find that it basically gives you the formula for curl above. (Try it.)

This is why $\nabla \times \mathbf{F}$ is the notation chosen. You could almost imagine

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

in which case the determinant above is the old mnemonic for the cross product. Again, this makes absolutely zero sense math-wise. It's only a convenient way to remember the formula, but it works really well because you only have to remember "$\nabla \times \mathbf{F}$".

Sample Question

Compute the curl of the vector field

$$\mathbf{F}(x,y,z) = \begin{pmatrix} xy \\ yz \\ zx \end{pmatrix}.$$

Solution. Let $p(x,y,z) = xy$, $q(x,y,z) = yz$, $r(x,y,z) = zx$. We can compute the first component of the curl by calculating

$$\begin{aligned} \frac{\partial r}{\partial y} &= \frac{\partial}{\partial y}(zx) = 0 \\ \frac{\partial q}{\partial z} &= \frac{\partial}{\partial z}(yz) = y. \end{aligned}$$

Hence:

$$\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z} = (0 - y) = -y$$

is the first component of the curl.

The second and third components are done in the same way. The second component is

$$\left(\frac{\partial p}{\partial z} - \frac{\partial r}{\partial x}\right) = (0 - z) = -z$$

and the third component is

$$\left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) = (0 - x) = -x.$$

Hence

$$\nabla \times \mathbf{F} = \boxed{\begin{pmatrix} -y \\ -z \\ -x \end{pmatrix}}.$$ □

Now let's talk about aquatic intuition. Suppose as we do for most of our examples that our vector field $\mathbf{F}$ represents the flow of water in an ocean or other body of water. We think of the curl as capturing the local rotation or swirling motion of the water at each point.

Here's how you can imagine curl:

1. **Place a small paddle wheel** (or imagine a small object that can rotate, like a stick with flags at each end) in the water at a point where you want to measure the curl.
2. **Observe how the wheel rotates** due to the flow of water. If the water is flowing uniformly in a straight line, the wheel will not rotate. In this case, the curl at that point is zero because there's no local rotation in the flow.
3. **If the wheel spins**, this indicates that there is local rotational motion in the water. The curl at the point is a 3D vector, so there are two pieces of information:

- The **direction** in which the wheel spins corresponds as follows to the direction of the curl. The **axis of rotation** of the wheel will point in the direction of the curl vector at that point. For example, if the water causes the wheel to rotate counterclockwise when viewed from above, the curl vector points upward. If the wheel rotates clockwise, the curl vector points downward.
- The **magnitude** of the curl is related to how fast the wheel spins. A faster rotation means a stronger curl, indicating more intense local rotational motion in the flow of water.

Example of high and low curl:

- In a region where water is circulating in a whirlpool-like pattern, the curl is high because the water is rapidly rotating around a central point.
- In a calm, straight-moving current, the curl is low or zero because the water doesn't exhibit any significant rotation.

In short, the curl of a vector field $\mathbf{F}$ in our context of water flow measures the tendency of the water to rotate around each point, rather than simply move in a straight line. Visualizing it with a paddle wheel helps convey the idea of local rotational motion, with the curl vector indicating the direction and strength of that rotation.

Now, I promised you earlier that any two red arrows put together give 0. So now we prove the following.

Memorize: Curl of conservative field is zero

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a function (aka scalar field), and let ∇f be the corresponding conservative vector field. Then (assuming ∇f is continuously differentiable), the curl of ∇f is $\mathbf{0}$ at every point i.e.

$$\operatorname{curl}(\nabla f) = \nabla \times (\nabla f) = \mathbf{0}.$$

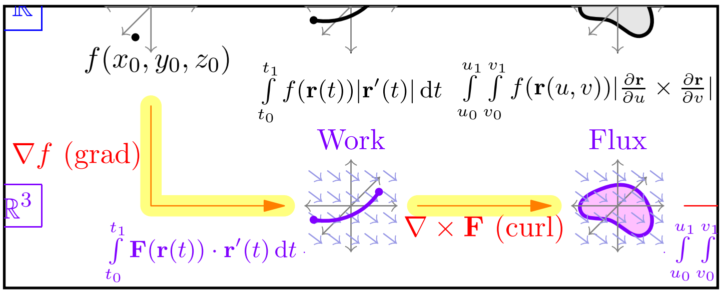

Figure 77: Cut-out of two red arrows from the poster Figure 75 that chain to give zero.

You can actually verify this theorem pretty easily by definition:

Proof. Since

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}$$

we get

$$\operatorname{curl}(\nabla f) = \begin{pmatrix} f_{zy} - f_{yz} \\ f_{xz} - f_{zx} \\ f_{yx} - f_{xy} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

because we saw in Section 16.6 that $f_{zy} = f_{yz} = 0$, etc. □

However, it's more important to have a visual understanding of why this is true. Remember, in the context of water flow, the fact that the curl of the gradient is zero means that if the flow is purely driven by a gradient (such as water moving due to pressure differences or height differences), there will be no rotational movement in the water. For example, if you have a waterfall, the water will flow directly downhill or uphill, without any swirling or spinning motion.

Here's an example showing this:

Sample Question

Compute the curl of the vector field

$$\mathbf{F}(x, y, z) = \begin{pmatrix} y^2 - \sin(x) \\ 2xy + 4yz \\ e^z + 2y^2 \end{pmatrix}.$$

Secretly, we happen to know the right-hand side is the gradient of the function $f(x, y, z) = y^2 x + \cos x + 2y^2 z + e^z + C$; this was the last example in Section 16.3. So with this insider information we expect the answer should come out to $\mathbf{0}$. Indeed, it does:

Solution. Let $p(x, y, z) = y^2 - \sin(x)$, $q(x, y, z) = 2xy + 4yz$, $r(x, y, z) = e^z + 2y^2$. First compute $\frac{\partial r}{\partial y}$ and $\frac{\partial q}{\partial z}$:

$$\frac{\partial r}{\partial y} = \frac{\partial}{\partial y}(e^z + 2y^2) = 4y$$

$$\frac{\partial q}{\partial z} = \frac{\partial}{\partial z}(2xy + 4yz) = 4y.$$

Compute the first component:

$$\left(\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z}\right) = 4y - 4y = 0.$$

For the second component, compute $\frac{\partial p}{\partial z}$ and $\frac{\partial r}{\partial x}$:

$$\frac{\partial p}{\partial z} = \frac{\partial}{\partial z}(y^2 - \sin x) = 0$$
$$\frac{\partial r}{\partial x} = \frac{\partial}{\partial x}(e^z + 2y^2) = 0.$$

Hence

$$\left(\frac{\partial p}{\partial z} - \frac{\partial r}{\partial x}\right) = 0 - 0 = 0.$$

Finally, compute $\frac{\partial q}{\partial x}$ and $\frac{\partial p}{\partial y}$:

$$\begin{aligned}\frac{\partial q}{\partial x} &= \frac{\partial}{\partial x}(2xy + 4yz) = 2y \\ \frac{\partial p}{\partial y} &= \frac{\partial}{\partial y}(y^2 - \sin x) = 2y.\end{aligned}$$

Compute the third component:

$$\left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) = 2y - 2y = 0.$$

So the curl of the vector field $\mathbf{F}(x, y, z)$ is $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \boxed{\mathbf{0}}$. □

Hence this is an example of a *conservative* vector field, which we'll talk more about later.

§32.4 [TEXT] Divergence

Definition of divergence

Suppose

$$\mathbf{F}(x, y, z) = \begin{pmatrix} p(x, y, z) \\ q(x, y, z) \\ r(x, y, z) \end{pmatrix}$$

is a 3D vector field. Then the **divergence** of $\mathbf{F}$ is the scalar field defined by

$$\operatorname{div}(\mathbf{F}) := \nabla \cdot \mathbf{F} := \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z}.$$

Type signature

Divergence takes vector fields as input. The divergence at each point is a *number* (i.e. is a scalar field).

Tip: How to memorize divergence

The notation $\nabla \cdot \mathbf{F}$ is supposed to also be a mnemonic. If you continue the analogy where

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

then $\nabla \cdot \mathbf{F}$ looks like a dot product that does the right thing. Again, only for memory; this is totally nonsense math-wise.

Sample Question

Compute the divergence of the vector field

$$\mathbf{F}(x,y,z) = \begin{pmatrix} xy \\ yz \\ zx \end{pmatrix}.$$

Solution. Let $p(x,y,z) = xy$, $q(x,y,z) = yz$, $r(x,y,z) = zx$. Then

$$\begin{aligned} \frac{\partial p}{\partial x} &= \frac{\partial}{\partial x}(xy) = y \\ \frac{\partial q}{\partial y} &= \frac{\partial}{\partial y}(yz) = z \\ \frac{\partial r}{\partial z} &= \frac{\partial}{\partial z}(zx) = x. \end{aligned}$$

Sum the partials to get the divergence:

$$\begin{aligned} \nabla \cdot \mathbf{F} &= \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \\ &= y + z + x = \boxed{x+y+z}. \end{aligned}$$

□

Now let's talk about aquatics. Again suppose $\mathbf{F}$ represents the flow of water in an ocean or another body of water. To visualize the divergence of a vector field $\mathbf{F}$, think of the divergence as measuring how much the water is **spreading out** or **converging** at each point.

Here's how you can imagine it:

1. **Place a small marker** (such as a floating object) at a point in the water flow. The goal is to observe how the flow of water behaves around that point.
2. **If the water appears to be flowing outward**, as if water is being emitted from that point, the divergence is **positive** at that location. This suggests that more water is moving away from the point than toward it, indicating a local source of water.
3. **If the water appears to be flowing inward**, as if the water is being sucked into that point, the divergence is **negative**. This suggests that water is converging at the point, indicating a local sink or depletion of water.
4. **If there is no noticeable net flow inward or outward** (the water moves but does not spread out or converge), the divergence is zero. This indicates that there is no net change in how much water is entering or leaving that point.

Example of high and low divergence:

- In a region where water is being pumped outward from a source, the divergence is high and positive, indicating that the water is spreading out from that point.
- In a region where water is being drawn into a drain, the divergence is negative, indicating that the water is converging toward that point.
- In a region where the water is flowing uniformly with no sources or sinks, the divergence is zero because there is no net flow into or out of any point.

In summary, the divergence of a vector field $\mathbf{F}$ in the context of water flow measures the rate at which water is spreading out (positive divergence) or converging (negative divergence) at each point. If there is neither spreading nor converging, the divergence is zero.

Digression on divergence of curl

I told you any two red arrows give you zero in the poster Figure 75. So there is technically a theorem that says the divergence of a curl is 0: that is,

$$\operatorname{div}(\operatorname{curl}(\mathbf{F})) = \nabla \cdot (\nabla \times \mathbf{F}) = 0$$

assuming $\mathbf{F}$ has continuous second partial derivatives. However, I don't think we ever use this in 18.02. The problem is that our description of divergence assumes that our vector field is thought of like a water current, but $\operatorname{curl}(\mathbf{F})$ is a vector field that describes how fast something rotates, and those arrows are emphatically *not* a water current (or anything resembling one).

§32.5 [TEXT] 2D scalar curl

The 2D scalar curl is a little more unnatural. The physical interpretation is the same, but if you have a 2D body of water, there's only two ways to rotate: either clockwise or counterclockwise. (In contrast, if you put a paddle wheel into the ocean, there are *lots* of ways it can rotate.)

So the 2D scalar curl, true to its name, only **outputs a number at each point**, which you think of as an angular velocity of the spinning paddle wheel. Unlike with the usual 3D curl it's no longer needed to specify an entire vector so that you can talk about the direction of rotation. Instead we take the convention that

- positive numbers mean counterclockwise spin,
- negative numbers mean clockwise spin.

Definition of 2D scalar curl

Suppose

$$\mathbf{F}(x, y) = \begin{pmatrix} p(x, y) \\ q(x, y) \end{pmatrix}$$

is a 2D vector field. Then the **2D scalar curl** of $\mathbf{F}$ is the scalar field defined by

$$\operatorname{curl} \mathbf{F} := \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}.$$

Tip: 2D scalar curl is a special case of 3D scalar curl

The mnemonic $\nabla \times \mathbf{F}$ actually still works if you just pretend $\mathbf{F}$ is a 3D vector field where the z-coordinate is always zero. That is, given $\mathbf{F} = \left(\begin{smallmatrix} p(x,y) \\ q(x,y) \end{smallmatrix}\right)$, consider the mnemonic

$$\nabla \times \begin{pmatrix} p(x,y) \\ q(x,y) \\ 0 \end{pmatrix}.$$

If you follow through, you will find you get

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ p(x,y) & q(x,y) & 0 \end{vmatrix}.$$

All the terms involving $\frac{\partial}{\partial z}$ disappear, because there's no z anywhere. So only the terms in front of $\mathbf{e}_3$ survive, and you get

$$\begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ p(x,y) & q(x,y) \end{vmatrix} \mathbf{e}_3 = \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathbf{e}_3.$$

And there's the 2D scalar curl, the coefficient of $\mathbf{e}_3$.

Sample Question

Compute the 2D scalar curl of

$$\mathbf{F}(x,y) = \begin{pmatrix} x \cos y \\ e^x + \sin y \end{pmatrix}.$$

Solution. The 2D scalar curl is given by

$$\operatorname{curl} F = \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}.$$

Given $q(x,y) = e^x + \sin y$, we have

$$\frac{\partial q}{\partial x} = \frac{\partial}{\partial x}(e^x + \sin y) = e^x + 0 = e^x.$$

Given $p(x,y) = x \cos y$ we have

$$\frac{\partial p}{\partial y} = \frac{\partial}{\partial y}(x \cos y) = x(-\sin y) = -x \sin y.$$

Hence

$$\operatorname{curl} \mathbf{F}(x,y) = \boxed{e^x + x \sin y}.$$ □

It's still true (and indeed follows from the 3D version) that:

Memorize

The 2D scalar curl of a conservative 2D vector field is zero at every point.

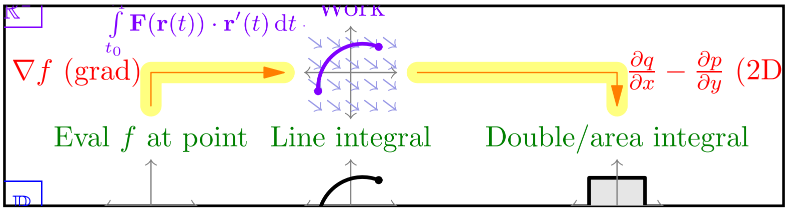

Figure 78: These two red arrows from the poster Figure 75 also chain to give zero.

Let's see an example of that.

Sample Question

Compute the 2D scalar curl of

$$\mathbf{F}(x,y) = \begin{pmatrix} 3x^2 + 4xy + y^2 \\ 2x^2 + 2xy - 3y^2 \end{pmatrix}.$$

Secretly, we happen to know the right-hand side is the gradient of the function $f(x,y) = x^3 + 2x^2y + xy^2 - y^3$, because we did this example Section 16.3. So the 2D scalar curl should be 0, and indeed it is.

Solution. Given $q(x,y) = 2x^2 + 2xy - 3y^2$:

$$\begin{aligned} \frac{\partial q}{\partial x} &= \frac{\partial}{\partial x}(2x^2 + 2xy - 3y^2) \\ &= 4x + 2y \end{aligned}$$

Given $p(x,y) = 3x^2 + 4xy + y^2$:

$$\begin{aligned} \frac{\partial p}{\partial y} &= \frac{\partial}{\partial y}(3x^2 + 4xy + y^2) \\ &= 4x + 2y. \end{aligned}$$

Hence

$$\begin{aligned} \operatorname{curl} \mathbf{F}(x,y) &= \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \\ &= (4x + 2y) - (4x + 2y) = \boxed{0}. \end{aligned}$$

□

§32.6 [TEXT] 2D divergence

This is *not* a red arrow in the picture. But it comes up in one version of Green's theorem, and it's actually exactly the same as 3D. So I'll just mention it briefly.

Definition of divergence

Suppose

$$\mathbf{F}(x,y) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$$

is a 2D vector field. Then the **divergence** of $\mathbf{F}$ is the scalar field defined by

$$\operatorname{div} \mathbf{F} := \nabla \cdot \mathbf{F} := \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}.$$

The aquatic interpretation is the same too, just in 2D bodies of water.

Sample Question

Compute the divergence of

$$\mathbf{F}(x,y) = \begin{pmatrix} x \cos y \\ e^x + \sin y \end{pmatrix}.$$

Solution. In two dimensions, the divergence is given by

$$\nabla \cdot \mathbf{F} = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}.$$

Given $p(x,y) = x \cos y$:

$$\frac{\partial p}{\partial x} = \frac{\partial}{\partial x}(x \cos y) = \cos y.$$

Given $q(x,y) = e^x + \sin y$:

$$\frac{\partial q}{\partial y} = \frac{\partial}{\partial y}(e^x + \sin y) = \cos y.$$

Hence

$$\begin{aligned} \nabla \cdot \mathbf{F}(x,y) &= \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \\ &= \cos y + \cos y = \boxed{2 \cos y}. \end{aligned}$$

□

Sample Question

Compute the divergence of

$$\mathbf{F}(x,y) = \begin{pmatrix} 3x^2 + 4xy + y^2 \\ 2x^2 + 2xy - 3y^2 \end{pmatrix}.$$

Solution. Given $p(x,y) = 3x^2 + 4xy + y^2$:

$$\frac{\partial p}{\partial x} = \frac{\partial}{\partial x}(3x^2 + 4xy + y^2) = 6x + 4y$$

Given $q(x, y) = 2x^2 + 2xy - 3y^2$:

$$\frac{\partial q}{\partial y} = \frac{\partial}{\partial y}(2x^2 + 2xy - 3y^2) = 2x - 6y.$$

Hence

$$\begin{aligned} \nabla \cdot \mathbf{F}(x, y) &= \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \\ &= (6x + 4y) + (2x - 6y) = \boxed{8x - 2y}. \end{aligned}$$

□

§32.7 [EXER] Exercises

Exercise 32.1. Consider the force of gravity $\mathbf{G}$ exerted by a point mass of mass m at a point O. Show that

$$\nabla \cdot \mathbf{G} = 0$$

at every point *except* O.

Chapter 33. Work (aka line integrals), and how to compute them with bare hands

This chapter defines a so-called *line integral*, the first of the two purple pictures in our poster Figure 75. For now, we'll only view this cell in isolation, so we'll give you the definition and show you how to use it with bare-hands.

However, it's worth saying now: **there will be shortcuts** to computing line integrals that bypass the work of parametrization. Those shortcuts are given by the *red arrows* in the Figure 75. In fact, the entire next chapter, Chapter 34, is dedicated to these shortcuts.

§33.1 [TEXT] Work

We now define the leftmost purple pictures in our poster Figure 75. When we have a vector field $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ as a path $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ through it, we can define the **work** on it.

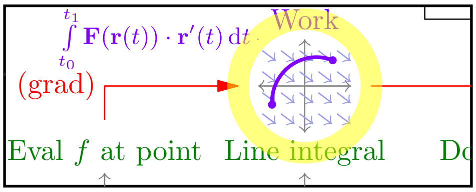

Figure 79: The work integral circled from the giant poster in Figure 75.

Definition of work

The **work** of $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ done on a path $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ is defined as

$$\int_{t=\text{start time}}^{\text{stop time}} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t.$$

Type signature

Keep in mind the types of the inputs; see Table 17. The work is a scalar quantity (there is a dot product inside the integrand, so it outputs a number).

Symbol	Name	Input type	Output type
$\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$	Parametrization of a path	Scalar t (time)	Point (dot) in $\mathbb{R}^n$
$\mathbf{r}' : \mathbb{R} \to \mathbb{R}^n$	Velocity vector for $\mathbf{r}$	Scalar t (time)	Vector (arrow) in $\mathbb{R}^n$
$\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$	Vector field	Point (dot) in $\mathbb{R}^n$	Vector (arrow) in $\mathbb{R}^n$

Table 17: The type signatures of the objects in the work integral.

This is commonly abbreviated with shorthand in two ways.

- First, we mention a new row of the shorthand in Table 13:

$$\mathrm{d}\mathbf{r} := \mathbf{r}'(t)\,\mathrm{d}t.$$

- Second, often the time parametrization is suppressed from the notation and we just write $\int_{\mathcal{C}}$ instead, where $\mathcal{C}$ denotes the curve that $\mathbf{r}(t)$ traces out. In this context we always consider the curve to be **directed**, i.e. one of the endpoints is the starting point, and the other is the ending point.

That means the above work integral can be rewritten as just

$$\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}.$$

Mercifully, the shorthand still writes a dot product in the symbols, to remind you that, yes, you should be evaluating a dot product when you compute this. I can't imagine how much confusion it would cause if the shorthand didn't have the dot product.

§33.2 [SIDENOTE] Aquatic interpretation of work

Letting $\mathbf{F}$ represent water current as always, the "work" done along a trajectory can be thought of as:

> **Idea**
>
> The work tells you how much the water current helps or hinders the movement of a swimmer through the water.

To compute this, consider a trajectory along which an object, such as a boat or swimmer, moves through the water. The current vector field $\mathbf{F}$ at any point describes the speed and direction of the water flow at that location. The work done by the current as the object follows a path $\mathcal{C}$ depends on the alignment of the current with the object's movement along that path.

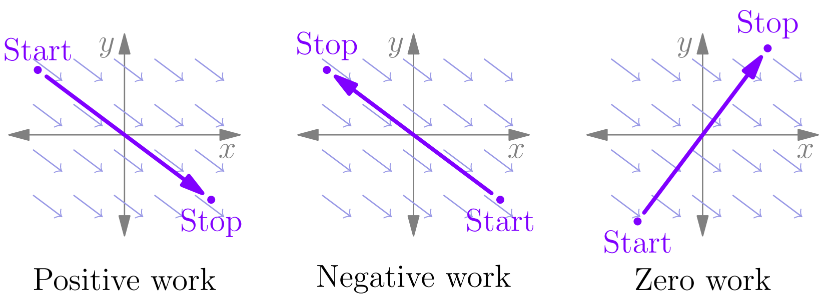

Figure 80: An illustration of the sign of work for a given vector field $\mathbf{F}$ which points roughly southeast. In the leftmost image, the path moves directly along $\mathbf{F}$ and hence the work is positive. In the middle image, $\mathbf{F}$ moves directly against $\mathbf{F}$ instead and hence the work is negative. In the rightmost image, $\mathbf{F}$ moves perpendicular to $\mathbf{F}$ and the work is zero instead.

1. *When the current aligns with the path*: If the direction of $\mathbf{F}$ aligns with the direction of the trajectory at a point, the current contributes positively to the work along that segment, effectively aiding the motion. This is experienced as a "push" in the direction of travel. The dot products are positive since at each point $P = \mathbf{r}(t)$, the vectors $\mathbf{F}(P)$ and $\mathbf{r}'(t)$ align well, and the work is a sum of a lot of positive numbers.

2. *When the current opposes the path*: If the current direction opposes the trajectory at any point, it contributes negatively to the work, effectively resisting the motion. In this case, the object has to work against the current, experiencing it as a "drag" force that slows its progress. The dot products are negative since at each point $P = \mathbf{r}(t)$, the vectors $\mathbf{F}(P)$ and $\mathbf{r}'(t)$ point against each other, and the work is a sum of a lot of negative numbers.

3. *When the current flows perpendicularly to the path*: If $\mathbf{F}$ is perpendicular to the trajectory at a point, it does no work in the direction of travel, as the current neither aids nor resists the movement along the path. The effect of the current in this case would primarily cause a lateral drift rather

than a forward or backward push along the trajectory. The dot products are zero in this case: at each point $P = \mathbf{r}(t)$, the vectors $\mathbf{F}(P)$ and $\mathbf{r}'(t)$ are perpendicular.

An illustration of all three situations is shown in Figure 80.

§33.3 [TEXT] Visualizing line integrals via dot products

If you want to visualize the integral, you can imagine walking along the path cut out by $\mathbf{r}$. At each point, you draw the tangent vector $\mathbf{r}'(t)$ to the path, and also look at the arrow for the vector field $\mathbf{F}(\mathbf{r}(\mathbf{t}))$ at that point. There's a dot product of these two vectors, which is a number. The line integral adds up all these numbers.

The light blue and purple in Figure 81 are totally separate

When drawing a cartoon like in Figure 81, it might be useful to keep in mind that there are two parts to the picture:

- the curve $\mathcal{C}$ and its parametrization $\mathbf{r}(t)$ (purple in Figure 81)
- the vector field $\mathbf{F}(x, y)$ (light blue arrows in Figure 81)

Remember, **these two parts have nothing to do with each other**. That is:

- When you're sketching the light blue arrows for $\mathbf{F}(x, y)$, you should only look at $\mathbf{F}$ and completely ignore $\mathcal{C}$ and $\mathbf{r}$.
- Similarly, when sketching the purple path $\mathcal{C}$, ignore $\mathbf{F}$ completely.

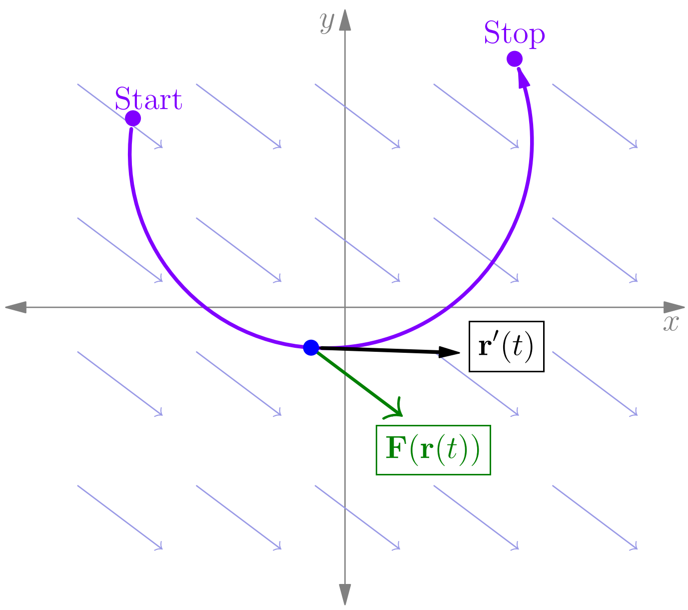

Figure 81: Cartoon of the dot products being added up by the work integral. Imagine adding up all the dot products $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$.

In practice, if you actually want to do the integral calculation, you end up having to do a parametrization, so the shorthand hides how much effort will be needed. For example, in the following exercise, $\mathcal{C}$ is the upper half of the circle $x^2 + y^2 = 1$, directed from the point $(1, 0)$ to $(-1, 0)$. (Again, for work integrals, we always require a specification of which way the integral moves along $\mathcal{C}$, if we choose to hide the parametrization $\mathbf{r}(t)$ from the notation.)

Tip: You get flexibility in parametrizations, as in Section 12.7

The work integral depends on which direction you walk along the path (it negates if you flip the start and stop point), but it doesn't depend on exactly how exactly you parametrize the path.

Thus, the comments from Section 12.7 apply here: if you're saying, parametrizing the semicircle $(1, 0)$ to $(-1, 0)$ the blue arc in later Figure 82), you should probably use $\mathbf{r}(t) = (\cos(t), \sin(t))$ for $0 \le t \le \pi$.

You could also use $\mathbf{r}(t) = (\cos(\pi t), \sin(\pi t))$ for $0 \le t \le 1$. Or if you wanted to annoy the grader, you could even use $\mathbf{r}(t) = (\cos(\pi t^2), \sin(\pi t^2))$ for $0 \le t \le 1$, which traces out the same arc at an irregular rate. Since these all give the same answer, you should pick the parametrization that makes the calculation easiest for you.

Tip: Splicing is OK

There's no issue with cutting up the path into multiple parts. For example, if $\mathcal{C}$ is a closed loop consisting of walking along the perimeter of the square, just cut it into the four line segments.

§33.4 [RECIPE] Computing line integrals by bare-hands via parametrization

Going back to our definition, here it is in recipe form.

Recipe for computing line integrals with bare-hands parametrization

To compute the line integral of $\mathbf{F}$ over the curve $\mathcal{C}$:

1. Pick **any** parametrization $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ of the curve $\mathcal{C}$, including specifying the start and stop times.
    - As described in Section 12.7, you have some freedom in how you set the parametrization: it only matters you start and end at the right place, and trace out the exact curve $\mathcal{C}$. So you should pick a parametrization that makes your calculation easier.
2. Calculate the derivative $\mathbf{r}'(t)$.
3. Calculate the dot product $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$.
4. Integrate this from the start time to the stop time.

Let's show some examples of how to calculate this in practice.

Sample Question

Compute the line integral of the vector field $\mathbf{F}(x, y) = \begin{pmatrix} 2y \\ 3x \end{pmatrix}$ along the following two curves:

- the upper half of the circle $x^2 + y^2 = 1$, oriented counterclockwise (blue in Figure 82).
- the line segment from $(1, 0)$ to $(-1, 0)$ (brown in Figure 82).

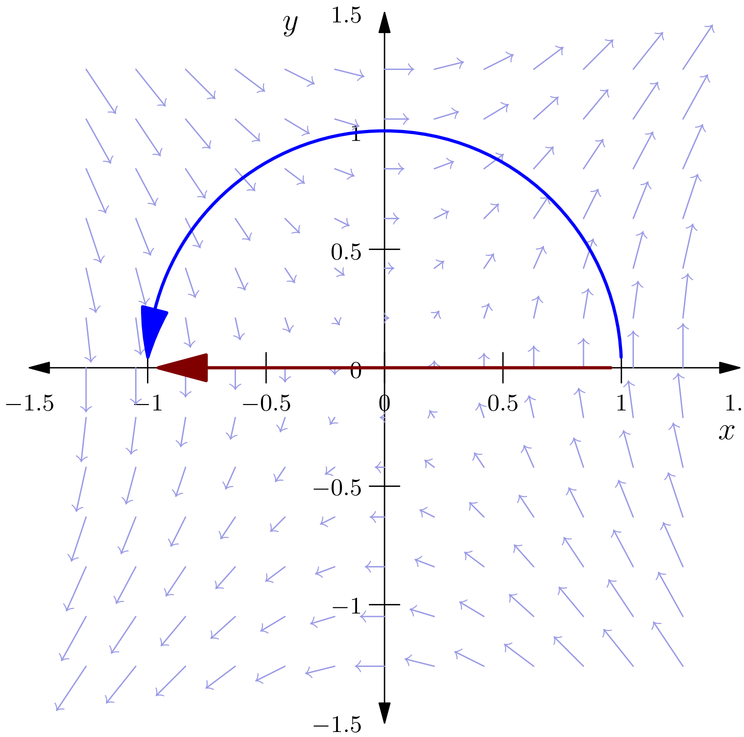

Figure 82: Two example of a work integral in the vector field $\mathbf{F}(x, y) = \binom{2y}{3x}$. The blue path is the upper semicircle of $x^2 + y^2 = 1$; the brown one is a straight line.

Solution. We do both parts; to prevent drowning in subscripts, we'll use $\mathcal{C}$ and $\mathbf{r}$ for the curve and parametrization for each part.

Before jumping into the calculation, look at Figure 82 to get a sense of what's going on. The blue arc has mixed signs: near the start and end of the arc, the dot products we're adding are positive as the small arrows line up well with the blue path. But we're moving against the current near the top. Since the arrows near the start are longer, you might guess the work integral is a small positive number, and you'd be right.

Meanwhile, along the brown arrow, all the arrows are perpendicular to our trajectory. We should expect the total work to thus be 0, and indeed it is.

- Let's first do the problem when $\mathcal{C}$ is the arc. The upper half of the circle $x^2 + y^2 = 1$ can be parametrized by:

$$\mathbf{r}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad \text{where } t \text{ ranges from } 0 \text{ to } \pi.$$

Substitute the parameterization into the vector field:

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(\cos t, \sin t) = \begin{pmatrix} 2\sin t \\ 3\cos t \end{pmatrix}.$$

Differentiate $\mathbf{r}(t)$ with respect to t:

$$\mathbf{r}'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}.$$

The line integral of $\mathbf{F}$ along $\mathcal{C}$ is given by:

$$\begin{aligned}
\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{t=0}^{\pi} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t \\
&= \int_{t=0}^{\pi} \begin{pmatrix} 2\sin t \\ 3\cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \mathrm{d}t \\
&= \int_{t=0}^{\pi} \left[-2\sin^2 t + 3\cos^2 t\right] \mathrm{d}t.
\end{aligned}$$

To simplify these trig expressions, we use the fact that

$$\sin^2 t = \frac{1-\cos 2t}{2}, \quad \cos^2 t = \frac{1+\cos 2t}{2}.$$

Substitute these into the integral:

$$\begin{aligned}
\int_{t=0}^{\pi} \left[-2\cdot\frac{1-\cos 2t}{2} + 3\cdot\frac{1+\cos 2t}{2}\right] \mathrm{d}t &= \int_{t=0}^{\pi} \left[-(1-\cos 2t) + \frac{3}{2}(1+\cos 2t)\right] \mathrm{d}t \\
&= \int_{t=0}^{\pi} \left[\frac{1}{2} + \frac{5}{2}\cos 2t\right] \mathrm{d}t.
\end{aligned}$$

The term $\int_{t=0}^{\pi} \cos(2t)\,\mathrm{d}t$ is zero by symmetry, so the final integral is $\boxed{\pi/2}$.

- Now let's suppose $\mathcal{C}$ is the brown line segment shown. Parametrize the curve $\mathcal{C}$ as $\mathbf{r}(t) = (1-2t, 0)$, where $0 \le t \le 1$. (You could also use $\mathbf{r}(t) = (1-t, 0)$ for $0 \le t \le 2$ if you prefer, or any other parametrization starting from $(1,0)$ and ending at $(-1,0)$; you'll get the same answer.)

  Differentiate $\mathbf{r}(t)$ with respect to t:

  $$\mathbf{r}'(t) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$

  Meanwhile, the parameterization into the vector field is:

  $$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(0, 1-2t) = \begin{pmatrix} 0 \\ 3-6t \end{pmatrix}.$$

  The dot product is identically equal to zero:

  $$\begin{pmatrix} -2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 3-6t \end{pmatrix} = 0.$$

  So the line integral is $\boxed{0}$ as well. □

In particular the work integral in general depends on which path you take: we got different answers for the blue and brown path above. It's only for the so-called **conservative** vector fields, which we'll talk about more in a moment, for which work integrals are path-independent.

§33.5 [TEXT] Even more shorthand: $p\,\mathrm{d}x + q\,\mathrm{d}y$

Then notation $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ can *still* be contracted further: there is another shorthand that hides both $\mathbf{F}$ and $\mathbf{r}$ altogether. Here it is:

Definition: More shorthand for line integrals

In $\mathbb{R}^2$, suppose $\mathbf{F}(x,y) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$. Then the work integral can further be abbreviated as

$$\int_{\mathcal{C}} (p\,\mathrm{d}x + q\,\mathrm{d}y).$$

Analogously, suppose we have a vector field $\mathbf{F}(x,y,z) = \begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}$ for $\mathbb{R}^3$. Then the work integral can further be abbreviated as

$$\int_{\mathcal{C}} (p\,\mathrm{d}x + q\,\mathrm{d}y + r\,\mathrm{d}z).$$

Remark

Here's the reason the shorthand is written like so. For simplicity, let's say we're in the 2D case and $\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$. Then $\begin{pmatrix} p \\ q \end{pmatrix} \cdot \mathbf{r}' = \begin{pmatrix} p \\ q \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} = p \cdot x' + q \cdot y' = p \cdot \frac{\mathrm{d}x}{\mathrm{d}t} + q \cdot \frac{\mathrm{d}y}{\mathrm{d}t}$. Hence, if we are integrating $\int_{t=\text{start}}^{\text{stop}} \begin{pmatrix} p \\ q \end{pmatrix} \cdot \mathbf{r}'\,\mathrm{d}t$, we could imagine "cancelling" the $\mathrm{d}t$ out, the expression we'd get looks like $p\,\mathrm{d}x + q\,\mathrm{d}y$.

For 18.02 purposes, all of this is only for mnemonic purposes; we don't actually define what any of the d symbols mean, so we can't make a more precise statement than that.

If any of p, q, r are zero, that term can also be omitted entirely. So for example, in 2D, if you see

$$\int_{\mathcal{C}} y\,\mathrm{d}x$$

you should take this shorthand to mean

$$\int_{\mathcal{C}} y\,\mathrm{d}x := \int_{\mathcal{C}} (y\,\mathrm{d}x + 0\,\mathrm{d}y) = \int_{\mathcal{C}} \begin{pmatrix} y \\ 0 \end{pmatrix} \cdot \mathrm{d}\mathbf{r}.$$

Let's do an example to practice the weird $\mathrm{d}x$ and $\mathrm{d}y$ shorthand, along a different path.

Sample Question

Let $\mathcal{C}$ denote the arc of the parabola $y = x^2$ starting from $(-1,1)$ and moving right to $(1,1)$. Compute the line integral

$$\int_{\mathcal{C}} y^{2/3}\,\mathrm{d}x.$$

Solution. First we need to expand the shorthand with $\mathrm{d}x$ and $\mathrm{d}y$. Recall that $p\,\mathrm{d}x + q\,\mathrm{d}y$ is shorthand for the vector field being $\begin{pmatrix} p \\ q \end{pmatrix}$. So where $y^{2/3}\,\mathrm{d}x = y^{2/3}\,\mathrm{d}x + 0\,\mathrm{d}y$, we expand the shorthand as

$$\int_{\mathcal{C}} y^{2/3}\,\mathrm{d}x = \int_{\mathcal{C}} \begin{pmatrix} y^{2/3} \\ 0 \end{pmatrix} \cdot \mathrm{d}\mathbf{r} = \int_{t=\text{start time}}^{\text{stop time}} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t$$

where $\mathbf{F}(x,y) := \begin{pmatrix} y^{2/3} \\ 0 \end{pmatrix}$ refers to the vector field encoded by the $y^{2/3}\,\mathrm{d}x$ shorthand.

Again, if you look at the sketch in Figure 83, we're expecting a positive work: all the arrows are pointing right, and the path $\mathcal{C}$ in red is moving right as well, so all the dot products are positive. (Again, if you imagine the blue arrows as a river current, it's definitely helping you swim, even if it's not directly aligned since you're not swimming straight east.)

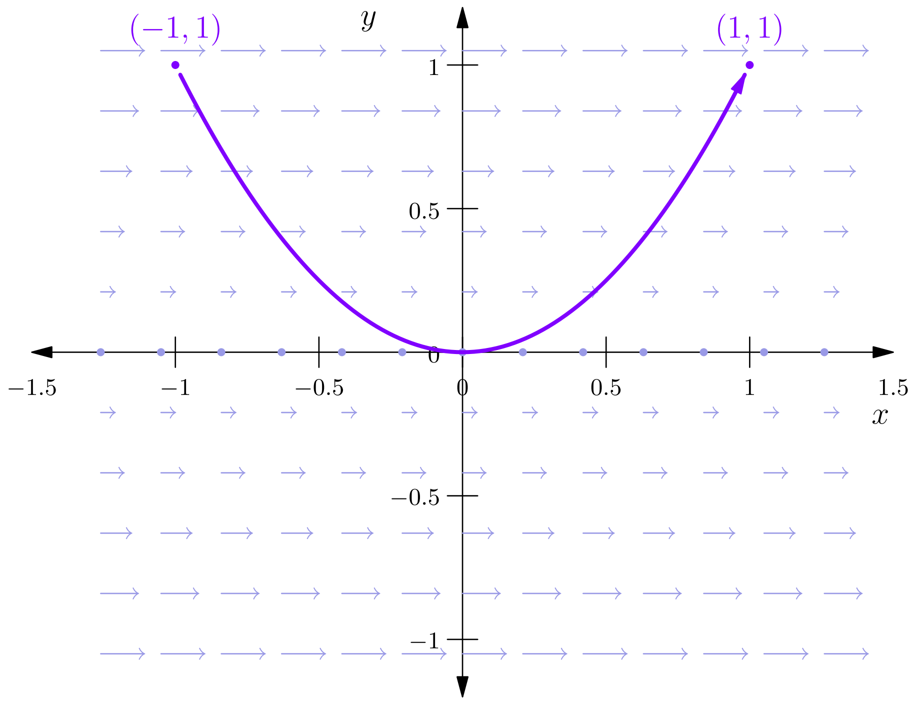

Figure 83: The vector field $\mathbf{F}(x,y) = \begin{pmatrix} y^{2/3} \\ 0 \end{pmatrix}$ (little light blue horizontal arrows) and an arc $\mathcal{C}$ from the parabola $y = x^2$ in it (purple).

The arc of the parabola we're trying to traverse needs to start at $(-1,1)$ and end at $(1,1)$. Just to make things concrete, examples of points we expect to pass through in our path are

$$(-1,1) \longrightarrow \left(-\frac{1}{2}, \frac{1}{4}\right) \longrightarrow \left(-\frac{1}{3}, \frac{1}{9}\right) \longrightarrow (0,0) \longrightarrow \left(\frac{1}{3}, \frac{1}{9}\right) \longrightarrow \left(\frac{1}{2}, \frac{1}{4}\right) \longrightarrow (1,1).$$

Anyway, we choose to parametrize the time as varying in $-1 \le t \le 1$ with

$$\mathbf{r}(t) = (t, t^2).$$

Now if we throw everything in, we have

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(t, t^2) = \begin{pmatrix} t^{\frac{4}{3}} \\ 0 \end{pmatrix}$$

and

$$\mathbf{r}'(t) = \begin{pmatrix} 1 \\ 2t \end{pmatrix}.$$

So the overall line integral becomes

$$\int_{t=-1}^{1} \underbrace{\begin{pmatrix} t^{\frac{4}{3}} \\ 0 \end{pmatrix}}_{=\mathbf{F}(\mathbf{r}(t))} \cdot \underbrace{\begin{pmatrix} 1 \\ 2t \end{pmatrix}}_{=\mathbf{r}'(t)} \, \mathrm{d}t = \int_{t=-1}^{1} t^{\frac{4}{3}} \, \mathrm{d}t = \left[\frac{3}{7} t^{\frac{7}{3}}\right]_{t=-1}^{1} = \frac{6}{7}.$$ □

§33.6 [EXER] Exercises

Exercise 33.1 (Suggested by Ting-Wei Chao)**.** Let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Let

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ using direct parametrization.

Exercise 33.2. Let $\mathcal{C}$ be a curve in $\mathbb{R}^2$ from $(0, 0)$ to $(2, 3)$ whose arc length is 7. Let $\mathbf{F}$ be a vector field with the property that for any point P on the curve,

- $\mathbf{F}(P)$ has magnitude 5;
- $\mathbf{F}(P)$ makes a 45° angle with the tangent vector to $\mathcal{C}$ at P (the tangent vector points along the direction of $\mathcal{C}$).

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$.

Chapter 34. Shortcuts for work: conservative vector fields and Green's theorem

In the last chapter we showed the definition of work and how to compute $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$ with bare-hands by using parametrization. However, parametrization can be a lot of work. So the purpose of this chapter is to show you under what conditions you can get away with *not* having to do a parametrization. There will be two such categories:

- If $\mathbf{F}$ is a *conservative* vector field, then the *fundamental theorem of calculus* is the way to go.
- If you're working in 2D and $\mathcal{C}$ is a closed loop, then Green's theorem is the way to go.

We'll show you both of these now.

§34.1 [TEXT] The fundamental theorem of calculus for line integrals

We now show the first Stokes result. It corresponds to a statement for the red arrow shown below.

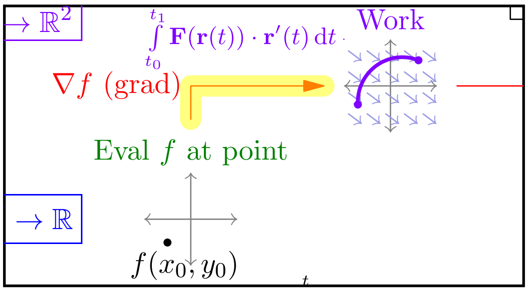

Figure 84: The FTC for line integrals is the Stokes statement for the "grad" red arrows in the poster Figure 75.

Memorize: FTC for line integrals

Suppose $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is a conservative vector field, given by $\mathbf{F} = \nabla f$ for some potential function $f : \mathbb{R}^n \to \mathbb{R}$. Then for any curve $\mathcal{C}$ from a point P to a point Q we have

$$\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} = f(Q) - f(P).$$

This has several important consequences. If you know $\mathbf{F}$ is conservative, then

- For *any* loop (i.e. a curve from a point to itself), the work integral is zero.
- More generally, all the work integrals are *path independent* if $\mathcal{C}$ and $\mathcal{C}'$ are two different paths from P to Q but the endpoints are the same, the work integrals will both be equal.

Indeed in the first case the work integral is $f(P) - f(P) = 0$ and the second case it equals $f(Q) - f(P)$ (for both $\mathcal{C}$ and $\mathcal{C}$'). But those two bullets are nice philosophically because they don't even require you to know anything at all about the function f.

In other words:

Idea: Practical consequences of FTC for line integrals

If you *already know* $\mathbf{F}$ is conservative, then

- If you also know the potential function f, then work integrals are extremely easy to calculate: just compute $f(\text{ending point}) - f(\text{starting point})$.
- If you don't know the potential function f, use the methods in Chapter 16 to find it.
- If the starting point and ending point are the same you don't even need to find f. The work integral is always 0.

Let's see this concretely with a conservative vector field. We'll use

$$\mathbf{F}(x, y) = \begin{pmatrix} 2x + 1 \\ 3y \end{pmatrix}$$

which, if we follow the recipe from Chapter 16, we can recover the f such that $\mathbf{F} = \nabla f$:

$$f(x, y) = x^2 + x + \frac{3}{2}y^2.$$

Thus, now that f is known, line integrals are trivial to compute:

Sample Question

Compute the line integral of the vector field $\mathbf{F}(x, y) = \begin{pmatrix} 2x+1 \\ 3y \end{pmatrix}$ along the following two curves:

- the upper half of the circle $x^2 + y^2 = 1$, oriented counterclockwise (blue in Figure 85).
- the line segment from $(1, 0)$ to $(-1, 0)$ (brown in Figure 85).

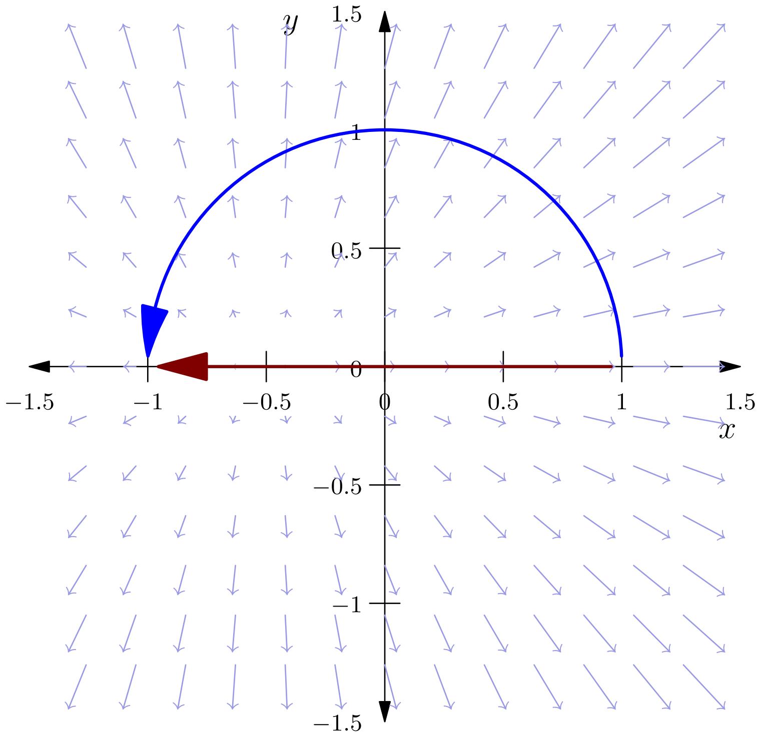

Figure 85: The same brown and blue path from Figure 82, but this time with a different vector field, now conservative.

Solution. Because we know the potential function

$$f(x,y) = x^2 + x + \frac{3}{2}y^2$$

the answer to both parts is the same:

$$f(-1,0) - f(1,0) = 0 - 2 = \boxed{-2}.$$

□

For comparison, we show how we could have computed the line integrals "by hand" for each of the bullets above, if we were not clever enough to notice that $\mathbf{F}$ was conservative. Of course, we do this knowing that the two answers better be equal (to -2).

- Work on the blue path, which is again $\mathbf{r}(t) = (\cos t, \sin t)$ for $0 \le t \le \pi$ with $\mathbf{r}'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$. The values of the new vector field $\mathbf{F}$ along the curve are

  $$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(\cos t, \sin t) = (2\cos t + 1, 3\sin t)$$

  Hence, the dot product being integrated is

  $$\begin{aligned} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) &= (2\cos t + 1)(-\sin t) + (3\sin t)(\cos t) \\ &= -2\cos t \sin t - \sin t + 3\sin t \cos t = \cos t \sin t - \sin t \end{aligned}$$

  Integrate with respect to t from 0 to $\frac{\pi}{2}$:

  $$\int_{t=0}^{\pi} (\cos t \sin t - \sin t)\,\mathrm{d}t = \int_{t=0}^{\pi} \left(\frac{\sin(2t)}{2} - \sin t \right) \mathrm{d}t = [-\cos(2t) - \cos(t)]_{t=0}^{\pi} = -2.$$

- Work on the brown line segment, parametrized again as $\mathbf{r}(t) = (1-2t, 0)$, where $t \in [0,1]$, and

  $$\mathbf{r}'(t) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$

  Putting parameterization into the new vector field gives:

  $$\mathbf{F}(\mathbf{r}(t)) = \begin{pmatrix} 2(1-2t)+1 \\ 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 3-4t \\ 0 \end{pmatrix}.$$

  The line integral of $\mathbf{F}$ along $\mathcal{C}$ is given by:

  $$\begin{aligned} \int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{t=0}^{1} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,\mathrm{d}t = \int_{t=0}^{1} -2(3-4t)\,\mathrm{d}t \\ &= \int_{t=0}^{1} (8t-6)\,\mathrm{d}t = \left(4t^2 - 6t\right)_{t=0}^{1} = -2. \end{aligned}$$

Sample Question

Suppose $\mathcal{C}$ is any path from $(1, 100)$ to $(42, 1337)$. Compute

$$\int_{\mathcal{C}} 5\,\mathrm{d}x.$$

Solution. Expanding the shorthand $5\,\mathrm{d}x = 5\,\mathrm{d}x + 0\,\mathrm{d}y$, the vector field we're integrating over is the constant vector field $\mathbf{F}(x,y) = \binom{5}{0}$. (In the cartoon, every blue arrow points directly east and has the same length 5.) This is certainly conservative: the potential function

$$f(x,y) = 5x$$

can be found just by guessing or via the method in Chapter 16. Indeed, $\nabla f = \binom{5}{0}$ as we needed.

So now that we know $\mathbf{F}$ is conservative and have found a potential function f, we can forget about parametrizing $\mathcal{C}$ and just write directly

$$\int_{\mathcal{C}} 5\,\mathrm{d}x = f(42,1337) - f(1,100) = 5\cdot 42 - 5\cdot 1 = 5(42-1) = \boxed{204}.$$ □

Remark

In general, the vector field encoded by $c\,\mathrm{d}x$ for any constant c is conservative with potential function $f(x,y) = cx$. Hence, $\int_{\mathcal{C}} c\,\mathrm{d}x = c\int_{\mathcal{C}} \mathrm{d}x$ will always just equal to c times the total change in x.

§34.2 [TEXT] Okay, but how do you tell whether $\mathbf{F}$ is conservative?

We saw that when $\mathbf{F}$ is conservative, the curl $\nabla \times \mathbf{F}$ is zero. It turns out that if $\mathbf{F}$ is defined everywhere, then the reverse is true too: that is, we can use $\nabla \times \mathbf{F}$ as a criteria for checking conservative fields.

Memorize: Conservative $\iff \nabla \times \mathbf{F} = \mathbf{0}$

Assume here the vector field is continuously differentiable and defined everywhere on $\mathbb{R}^2$ or $\mathbb{R}^3$.

- A vector field $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\mathbf{F}(x,y) = \binom{p(x,y)}{q(x,y)}$ is conservative if and only if the 2D scalar curl is zero everywhere:
$$\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = 0.$$
- A vector field $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is conservative if and only if the curl $\nabla \times \mathbf{F}$ is zero everywhere. In symbols, if $\mathbf{F}(x,y) = \begin{pmatrix} p(x,y,z) \\ q(x,y,z) \\ r(x,y,z) \end{pmatrix}$ then we need all three components of the curl to equal 0:
$$\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z} = \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} - \frac{\partial r}{\partial x} = 0.$$

This should look familiar: it's the same thing I told you in Section 16.6. The only thing that's change is that I now have an aquatic interpretation of all the equations: that we require the 2D or 3D curl to be zero. But the equations are the same.

Remark

This theorem also fails if the vector field $\mathbf{F}$ is only defined on part of $\mathbb{R}^n$ and the region is not *simply connected*. In that case it only works one direction — that is, if $\nabla \times \mathbf{F} \neq \mathbf{0}$ then $\mathbf{F}$ is definitely not conservative, but some non-conservative fields also satisfy $\nabla \times \mathbf{F} = \mathbf{0}$.

Sample Question

For which real number c is the vector field

$$\mathbf{F} = \begin{pmatrix} e^{\cos x} + xy^5 \\ cx^2y^4 + \log(y^2+1) \end{pmatrix}$$

a conservative vector field?

Solution. We need the number c such that

$$\begin{aligned} \frac{\partial}{\partial x}(cx^2y^4 + \log(y^2+1)) &= \frac{\partial}{\partial y}(e^{\cos x} + xy^5) \\ \iff c \cdot 2xy^4 &= 5xy^4 \end{aligned}$$

holds for all real numbers x and y. This occurs only when $\boxed{c = \frac{5}{2}.}$. □

Sample Question

For which real numbers a, b is the vector field

$$\mathbf{F} = \begin{pmatrix} y^2 + ax^2z + e^x \\ bxy + z\cos(yz) \\ x^3 + y\cos(yz) \end{pmatrix}.$$

a conservative vector field?

Solution. Let

$$\begin{aligned} p &= y^2 + ax^2z + e^x \\ q &= bxy + z\cos(yz) \\ r &= x^3 + y\cos(yz). \end{aligned}$$

We need to seek (a, b) such that the curl $\nabla \cdot \mathbf{F}$ is zero, that is

$$\frac{\partial r}{\partial y} - \frac{\partial q}{\partial z} = \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = \frac{\partial p}{\partial z} - \frac{\partial r}{\partial x} = 0.$$

We compute all six partial derivatives in turn.

- For the first component of the curl to be zero, we need the following to partials to be equal:

  $$\begin{aligned} \frac{\partial r}{\partial y} &= \frac{\partial}{\partial y}(x^3 + y\cos(yz)) = \cos(yz) + y(-\sin(yz))z = \cos(yz) - yz\sin(yz) \\ \frac{\partial q}{\partial z} &= \frac{\partial}{\partial z}(bxy + z\cos(yz)) = \cos(yz) + z(-\sin(yz))y = \cos(yz) - yz\sin(yz). \end{aligned}$$

  But this is always true, regardless of a and b.
- For the second component of the curl to be zero, we need the following to partials to be equal:

$$\begin{aligned}\frac{\partial p}{\partial z} &= \frac{\partial}{\partial z}(y^2 + ax^2 z + e^x) = ax^2 \\ \frac{\partial r}{\partial x} &= \frac{\partial}{\partial x}(x^3 + y\cos(yz)) = 3x^2.\end{aligned}$$

This occurs if and only if $a = 3$.

- For the third component of the curl to be zero, we need the following to partials to be equal:

$$\begin{aligned}\frac{\partial q}{\partial x} &= \frac{\partial}{\partial x}(bxy + z\cos(yz)) = by \\ \frac{\partial p}{\partial y} &= \frac{\partial}{\partial y}(y^2 + ax^2 z + e^x) = 2y.\end{aligned}$$

This occurs if and only if $b = 2$.

Hence $\boxed{(a, b) = (3, 2)}$ is the only answer. □

Digression

In particular, there should be a potential function for each of the two examples above.

In the first example, it's not easy to *write down* a potential function, because $e^{\cos x}$ has no easily-expressed anti-derivative. (Though $\log(y^2+1)$ does; it turns out to be $y(\log(y^2+1)-2) + 2\arctan(y)$.) So we are content that some potential function *does exist* even if it cannot be written down using familiar functions.

On the other hand, the second example can be integrated readily enough, by following the procedure in Chapter 16: one should get

$$f(x, y, z) = e^x + \cos(yz) + x^3 z + xy^2 + C.$$

§34.3 [TEXT] Green's theorem (2D only)

We expect there should be a Stokes result as well for the red arrow joining the 2D work integral to an area integral.

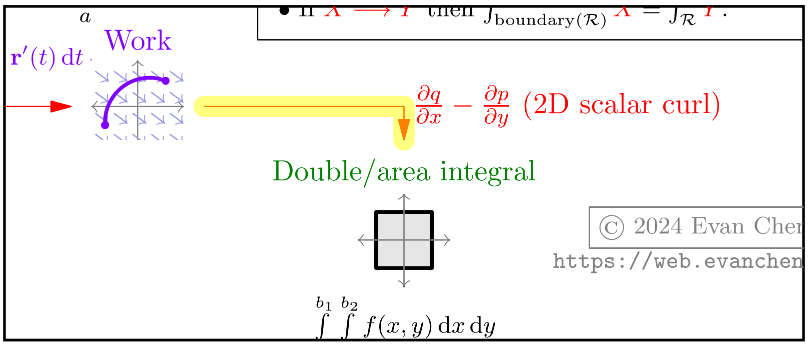

Figure 86: Green's theorem is the Stokes result for the above red arrow from the poster Figure 75.

Here is what it says.

Memorize: Green's theorem for converting work to curl

Suppose $\mathcal{C}$ is a closed loop parametrized by $\mathbf{r}(t)$ that encloses a region $\mathcal{R}$ counterclockwise. Then for any vector field $\mathbf{F} = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$, conservative or not, we have

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \oint_{\mathcal{C}} (p \, \mathrm{d}x + q \, \mathrm{d}y) = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A.$$

There's a new symbol $\oint_{\mathcal{C}}$ on the left, but it has the same meaning as $\int_{\mathcal{C}}$. The circle is there to emphasize that $\mathcal{C}$ is a closed loop, i.e. it's required to have the same start and ending point (unlike the other curves we saw in earlier examples). In other words:

Definition of $\oint$

$\oint_{\mathcal{C}}$ means "$\int_{\mathcal{C}}$ but with an extra optional reminder that $\mathcal{C}$ is a loop". (The reminder is optional, i.e. you are not obligated to add the circle even when $\mathcal{C}$ is a loop.)

Note this doesn't require $\mathbf{F}$ to be conservative! (All the past discussion about $\mathbf{F}$ being conservative was because we were using the red "grad" arrow in Figure 84. But we're now moving on to a new red arrow in our poster, and that assumption about a gradient isn't needed anymore.) In fact, in the event that $\mathbf{F} = \nabla f$ *is* conservative, we know that $\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = 0$: the 2D scalar curl of a conservative vector field is 0. So Green's theorem is then just saying that $\oint_{\mathcal{C}} \nabla f \, \mathrm{d}\mathbf{r} = 0$ which we already knew.

Tip: Always use counterclockwise orientation

Whenever $\mathcal{C}$ is a closed loop in $\mathbb{R}^2$, we'll basically always assume that the direction we walk around it is counterclockwise. It's considered bad manners to break this convention and have a loop oriented clockwise.

Green's theorem gives us a way to short-circuit a bunch of calculations that we were doing by hand earlier in the case where our loop is closed. Here are a few.

Sample Question

Calculate the line integral

$$\oint_{\mathcal{C}} (x^3 - y) \, \mathrm{d}x + (x + y^3) \, \mathrm{d}y,$$

where $\mathcal{C}$ is the circle $x^2 + y^2 = 4$ oriented counterclockwise.

Solution. Let $\mathcal{R}$ denote the region enclosed by $\mathcal{C}$. We use Green's theorem with the vector field

$$p(x,y) = x^3 - y, \quad q(x,y) = x + y^3.$$

Calculate the partial derivatives of q with respect to x and p with respect to y:

$$\frac{\partial q}{\partial x} = \frac{\partial}{\partial x}(x + y^3) = 1, \frac{\partial p}{\partial y} = \frac{\partial}{\partial y}(x^3 - y) = -1.$$

Substitute the partial derivatives into Green's theorem:

$$\oint_{\mathcal{C}} p\,\mathrm{d}x + q\,\mathrm{d}y = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) \mathrm{d}A = \iint_{\mathcal{R}} (1-(-1))\,\mathrm{d}A = \iint_{\mathcal{R}} 2\,\mathrm{d}A.$$

The region $\mathcal{R}$ is the disk defined by $x^2 + y^2 \le 4$, which is a circle of radius 2, hence with area

$$\operatorname{Area}(\mathcal{R}) = \pi r^2 = \pi(2)^2 = 4\pi.$$

So the answer is $2 \cdot 4\pi = \boxed{8\pi}$. □

Sample Question

Evaluate the line integral

$$\oint_{\mathcal{C}} (y\,\mathrm{d}x - x\,\mathrm{d}y)$$

where $\mathcal{C}$ is the triangle with vertices at $(0,0)$, $(1,0)$, and $(0,1)$, oriented counterclockwise.

Solution. Let $\mathcal{R}$ denote the interior of the triangle. By Green's theorem:

$$\oint_{\mathcal{C}} p\,\mathrm{d}x + q\,\mathrm{d}y = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) \mathrm{d}A,$$

where $P(x,y) = y$ and $Q(x,y) = -x$. Calculate $\frac{\partial q}{\partial x}$ and $\frac{\partial p}{\partial y}$:

$$\frac{\partial q}{\partial x} = -1, \quad \frac{\partial p}{\partial y} = 1.$$

Hence

$$\iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) \mathrm{d}A = \iint_{\mathcal{R}} (-1-1)\,\mathrm{d}A = \iint_{\mathcal{R}} -2\,\mathrm{d}A.$$

The area of the triangle $\mathcal{R}$ is:

$$\operatorname{Area}(\mathcal{R}) = \frac{1}{2} \cdot \text{base} \cdot \text{height} = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}.$$

Thus:

$$\iint_{\mathcal{R}} -2\,\mathrm{d}A = -2 \cdot \frac{1}{2} = \boxed{-1}.$$

□

Sample Question

Calculate the line integral

$$\oint_{\mathcal{C}} (x^2\,\mathrm{d}y - y^2\,\mathrm{d}x)$$

where C is the boundary of the square with vertices at $(1,1)$, $(0,1)$, $(0,0)$, and $(1,0)$, oriented counterclockwise.

Solution. Let $\mathcal{R}$ denote the interior of the square. By Green's theorem:

$$\oint_{\mathcal{C}} p \,\mathrm{d}x + q \,\mathrm{d}y = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A,$$

where $p(x,y) = -y^2$ and $q(x,y) = x^2$. Calculate $\frac{\partial q}{\partial x}$ and $\frac{\partial p}{\partial y}$:

$$\frac{\partial q}{\partial x} = 2x, \quad \frac{\partial p}{\partial y} = -2y.$$

Substitute these into Green's theorem:

$$\iint_{\mathcal{R}} (2x - (-2y)) \,\mathrm{d}A = \iint_{\mathcal{R}} (2x + 2y) \,\mathrm{d}A.$$

Since $\mathcal{R}$ is a square with side length 1 centered at the origin, integrate over x and y from 0 to 1:

$$\iint_{\mathcal{R}} (2x + 2y) \,\mathrm{d}A = 2 \int_{y=0}^{1} \int_{x=0}^{1} (x + y) \,\mathrm{d}x \,\mathrm{d}y.$$

Evaluate the inner integral with respect to x:

$$\left[\frac{x^2}{2} + yx \right]_{x=0}^{1} = y + \frac{1}{2}.$$

Evaluate the integral with respect to y:

$$2 \int_{y=0}^{1} \left(y + \frac{1}{2} \right) \mathrm{d}y = \boxed{2}.$$

□

§34.4 [SIDENOTE] A picture explaining why Green's Theorem for work should be true (not a formal proof)

Here is an extremely informal explanation of what Green's Theorem is trying to say pictorially. We won't make it precise or go into the details.

Remember that the *2D scalar curl* of a 2D vector field $\mathbf{F}$ at a point P is a number that describes the counterclockwise swirl of the field $\mathbf{F}$ near P. So to draw a picture (Figure 87):

- Let's $\mathcal{C}$ be a counterclockwise loop oriented counterclockwise. For the picture, we'll draw $\mathcal{C}$ as a purple square, which encloses a region $\mathcal{R}$.
- Then we imagine breaking $\mathcal{R}$ into a bunch of tiny little squares. At each little square, we draw a little green swirl inside it that corresponds roughly to the 2D scalar curl of $\mathbf{F}$ at the center of the tiny square.

Then the integral

$$\iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A$$

can be thought of as the "sum of the green swirls" (whatever that means).

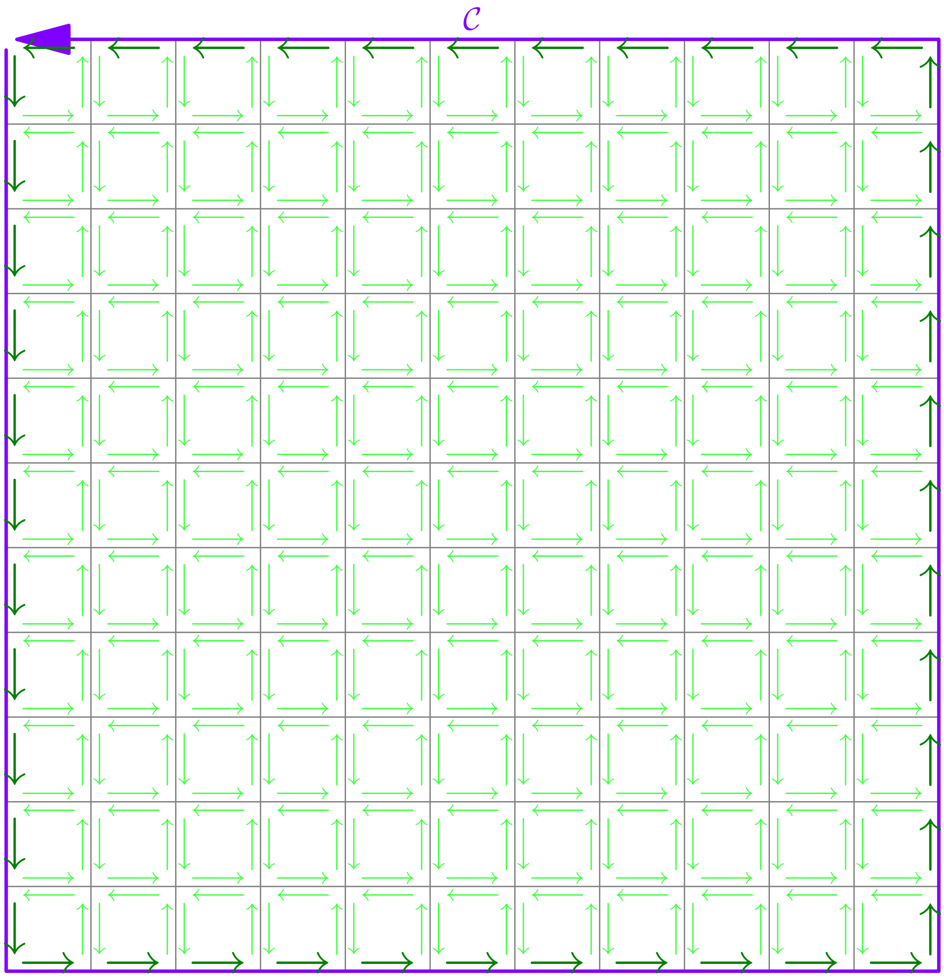

Figure 87: A rough picture of what Green's Theorem for work is trying to say: "if you add up the green swirls, then only the work along the purple boundary is left".

However, in Figure 87 you should realize: **all the swirls on the inside cancel**. Imagine one of the vertical grey walls between two grey cells: you can imagine the direction and magnitude of $\mathbf{F}$ along the wall contributes to the "swirliness" of the two adjacent cells, but whatever it contributes positively to one cell, it contributes negatively to the other one. (Again, this is all an informal picture, so I won't make this precise.)

So if we add all the green stuff, the only thing that's left is the green stuff that's just along the purple curve (drawn darker above). For example, the dark green arrows on the left correspond to how much $\mathbf{F}$ points downwards against the nearby grey walls: which exactly matches the description of the work integral of $\mathbf{F}$ along that western wall. And when you sum all four dark green currents, you just get the total work done by $\mathbf{F}$ along the purple curve $\mathcal{C}$, as desired.

§34.5 [RECIPE] Evaluating line integrals, all together now

While we gave a definition of line integrals with parametrization, we then saw right away there are a couple shortcuts, namely FTC and Green's theorem (in 2D) in certain cases. So with this, we can present a recipe that condenses these together.

Recipe for computing line integrals with possible shortcuts

Suppose we want to evaluate $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$.

1. First, check if the vector field is conservative by seeing if the curl is zero.
   - If so, don't bother parametrizing $\mathcal{C}$. Don't even look at $\mathcal{C}$ besides the endpoints. Find a potential function f for the vector field $\mathbf{F}$ and use the FTC as a shortcut: output
   $$f(\text{stop}) - f(\text{start}).$$
2. Second, if the line integral is in $\mathbb{R}^2$, check if $\mathcal{C}$ is a closed loop.
   - If so, see if Green's theorem gives you an easy shortcut:
   $$\oint_{\mathcal{C}} (p\,dx + q\,dy) = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dA.$$
3. If both of these fail, fall back to the parametrization recipe described in Section 33.4. To repeat it here:
   1. Pick **any** parametrization $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ of the curve $\mathcal{C}$, including specifying the start and stop times. As described in Section 12.7, you have some freedom in how you set the parametrization.
   2. Calculate the derivative $\mathbf{r}'(t)$.
   3. Calculate the dot product $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$.
   4. Integrate this from the start time to the stop time.

We'll give several more examples of this in Section 35.5, where we contrast it to another type of line integral, the "2D flux".

§34.6 [TEXT] Advanced technique: sealing regions

Green's Theorem is powerful enough that it can be handy even if the path $\mathcal{C}$ is not a closed loop: the idea is to "seal" the loop by adding some simple path, for which the line integral is easy to calculate. To show this technique, we bring back the first example from Section 33.4 all the way back when we first introduced how to compute work with bare hands.

Sample Question

Compute the line integral of the vector field $\mathbf{F}(x, y) = \begin{pmatrix} 2y \\ 3x \end{pmatrix}$ along the upper half of the circle $x^2 + y^2 = 1$, oriented counterclockwise. See Figure 88.

We already saw that we could compute this using bare-hands parametrization. Now we'll show how to use Green's Theorem as a shortcut by adding the line segment from $(-1, 0)$, to $(1, 0)$.

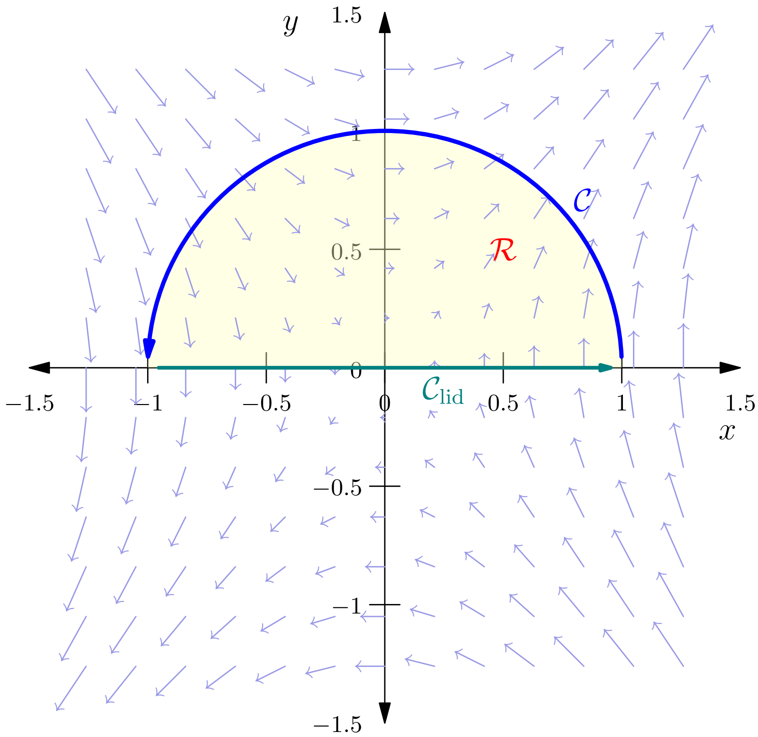

Figure 88: Evaluation of $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ by "sealing" the region, adding in a line segment joining $(-1,0)$ to $(1,0)$. The line integral across the segment is easy to compute (it equals zero, since the force is perpendicular to it.) Then Green's theorem applies to the sealed region $\mathcal{R}$.

Solution. Let $\mathcal{C}$ denote the semicircle. Because $\mathcal{C}$ is not a closed loop, Green's Theorem does not apply directly. To use it, we instead add a new line segment $\mathcal{C}_{\text{lid}}$ pointing from $(-1,0)$ to $(1,0)$. Then if we consider *both* $\mathcal{C}$ and the new lid $\mathcal{C}_{\text{lid}}$, they enclose the upper half of a disk $\mathcal{R}$ with area $\frac{\pi}{2}$, as shown in Figure 88. Hence Green's Theorem on the two-part boundary states that

$$\begin{aligned}
\underbrace{\int_{\mathcal{C}} (2y \,\mathrm{d}x + 3x \,\mathrm{d}y)}_{\text{what we want}} + \int_{\mathcal{C}_{\text{lid}}} (2y \,\mathrm{d}x + 3x \,\mathrm{d}y) &= \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A \\
&= \iint_{\mathcal{R}} \left(\frac{\partial}{\partial x}(3x) - \frac{\partial}{\partial y}(2y) \right) \mathrm{d}A \\
&= \iint_{\mathcal{R}} (3-2) \,\mathrm{d}A = \iint_{\mathcal{R}} \mathrm{d}A \\
&= \operatorname{Area}(\mathcal{R}) = \frac{\pi}{2}.
\end{aligned}$$

On the other hand, I claim that

$$\int_{\mathcal{C}_{\text{lid}}} (2y \,\mathrm{d}x + 3x \,\mathrm{d}y) = 0.$$

This is easy to compute with direct parametrization: if we parametrize the lid by $\mathbf{r}(t) = (t,0)$ for $-1 \le t \le 1$, for example, then

$$\int_{\mathcal{C}_{\text{lid}}} (2y \,\mathrm{d}x + 3x \,\mathrm{d}y) = \int_{t=-1}^{1} \begin{pmatrix} 2 \cdot 0 \\ 3 \cdot t \end{pmatrix} \cdot \mathbf{r}'(t) \,\mathrm{d}t = \int_{t=-1}^{1} \begin{pmatrix} 0 \\ 3t \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{d}t = \int_{t=-1}^{1} 0 \,\mathrm{d}t = 0.$$

Indeed one can even see it from Figure 88 directly, since the vector field is perpendicular to the x-axis along the entire lid, so the total work being 0 is not a surprise. Thus, the desired line integral is

$$\begin{aligned}\underbrace{\int_{\mathcal{C}} (2y\,\mathrm{d}x + 3x\,\mathrm{d}y)}_{\text{what we want}} &= \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) \mathrm{d}A - \int_{\mathcal{C}_{\text{lid}}} (2y\,\mathrm{d}x + 3x\,\mathrm{d}y) \\ &= \frac{\pi}{2} - 0 = \boxed{\frac{\pi}{2}}.\end{aligned}$$

□

§34.7 [EXER] Exercises

Exercise 34.1. Is the vector field

$$\mathbf{F}(x,y) = \begin{pmatrix} \sin(e^x) \\ \arctan(y^\pi + \pi^y) \end{pmatrix}$$

conservative?

Exercise 34.2. Calculate the line integral

$$\oint_{\mathcal{C}} (x^2 - y)\,\mathrm{d}x + (y^2 - x)\,\mathrm{d}y$$

where $\mathcal{C}$ is the boundary of the region enclosed by the circle $x^2 + y^2 = 4$, oriented counterclockwise.

Exercise 34.3 (Suggested by Ting-Wei Chao)**.** As in Exercise 33.1, let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Again let

$$\mathbf{F}(x,y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ this time using Green's Theorem.

Exercise 34.4 (*) (Shoelace formula)**.** Let $n \geq 3$ be an integer and suppose $\mathcal{P} = P_1 P_2 \dots P_n$ is a convex n-gon in $\mathbb{R}^2$, where the vertices $P_i = (x_i, y_i)$ are labeled counterclockwise. Use Green's theorem to prove the following formula for the area of $\mathcal{P}$:

$$\operatorname{Area}(\mathcal{P}) = \frac{1}{2} \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i).$$

Here $x_0 = x_n$ and $y_0 = y_n$ by convention, so the $i = 0$ summand is $x_n y_1 - x_1 y_n$.

Chapter 35. 2D flux

§35.1 [TEXT] Definition of 2D flux

I will grudgingly define 2D flux first, since I just went over Green's theorem. I say "grudgingly" because 2D flux is really a special case of 3D flux, but to keep things simple we've still been working in two dimensions.

The idea of flux is that you have some closed curve $\mathcal{C}$ in $\mathbb{R}^2$. When we had a work integral, we went along the curve $\mathcal{C}$ and added together the dot product of the vector field with the tangent vectors on that vector field.

With 2D flux, we instead take the dot product of the vector field with the *normal vector* rather than the tangent vector. This should be drawn as a 90° *clockwise rotation of* $\mathbf{r}'(t)$. Seriously, I can't make this up.

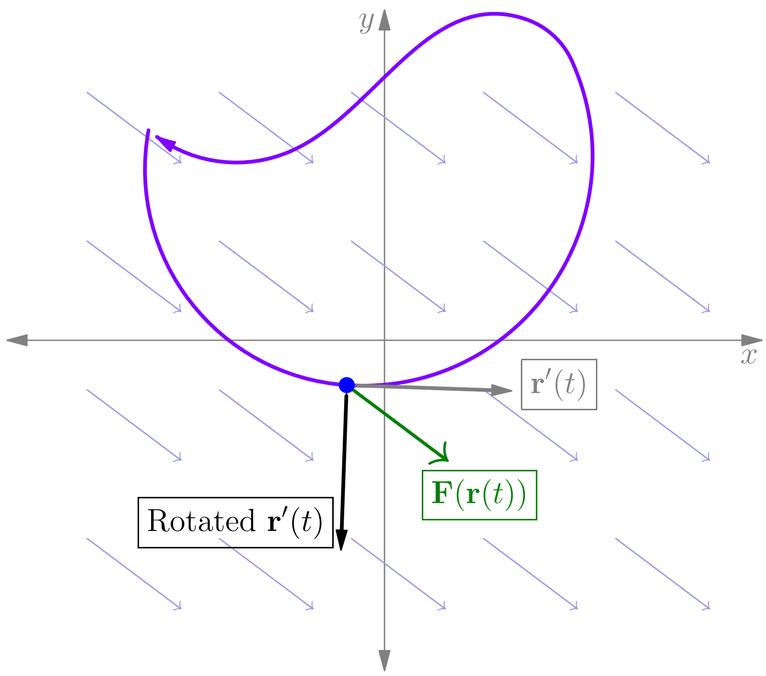

Figure 89: The 2D flux is the dot product where the tangent $\mathbf{r}'(t)$ is replaced by its rotated version.

In any case, the 2D flux is then defined as follows.

Definition of 2D flux

The 2D flux of a vector field $\mathbf{F}$ through the closed path $\mathcal{C}$ parametrized by $\mathbf{r}(t)$ is defined by

$$\int_{t=\text{start time}}^{\text{stop time}} \mathbf{F}(\mathbf{r}(t)) \cdot (90^\circ \text{ clockwise rotation of } \mathbf{r}'(t))\, dt.$$

Type signature

2D flux is a scalar quantity. It's only defined for a vector field in $\mathbb{R}^2$ piercing a closed path in $\mathbb{R}^2$.

The "90° clockwise rotation of $\mathbf{r}'(t)$" is so awkward that you can bet people immediately made up a shorthand to sweep it under the rug. I think the usual notation is

$$\mathbf{n}\,\mathrm{d}s := (90^\circ \text{ clockwise rotation of } \mathbf{r}'(t))\,\mathrm{d}t$$

so that the above thing will usually be condensed to

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}s.$$

Digression on why $\mathbf{n}\,\mathrm{d}s$ is the shorthand

I think the reason this shorthand is used is: $\mathbf{n}$ is supposed to be the "outward unit normal vector", i.e. a vector of length 1 whose direction is 90° rotated from $\mathbf{r}'(t)$. So then it needs to be scaled by the magnitude $|\mathbf{r}'(t)|$, and so we copy the old $\mathrm{d}s$ from arc length.

So this notation is consistent with the notation used for scalar-field line integrals (if you consider the scalar field $f = \mathbf{F} \cdot \mathbf{n}$). But I don't like to mention this because I want to avoid scalar-field line integrals in 18.02 for anything that isn't arc length to keep things simple.

Warning

The rotated $\mathbf{r}'(t)$ is sometimes called the "outward normal vector". However, despite the name, it only points outward if we oriented $\mathcal{C}$ counterclockwise. If $\mathcal{C}$ is clockwise it points inwards instead!

§35.2 [TEXT] Aquatic interpretation of 2D flux

Aquatically, if the curve $\mathcal{C}$ is thought of as some permeable membrane, then the 2D flux measures the rate the current passes through the membrane. Assuming $\mathcal{C}$ is oriented counterclockwise, the 2D flux is positive if water is (net) moving out of $\mathcal{C}$; it's negative if it flows in.

§35.3 [TEXT] 2D flux is a rotation of 2D work

We don't like the $\mathbf{n}\,\mathrm{d}s$ notation because we don't like scalar-field line integrals. Fortunately, there is another way to write the flux with shorthand that avoids $\mathbf{n}\,\mathrm{d}s$ notation. To see where it comes from, once again write

$$\mathbf{F}(x, y) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}.$$

Rather than rotating $\mathbf{r}'(t)$ by 90° clockwise, let's imagine we instead rotated $\mathbf{F}$ by 90° counterclockwise instead, and use:

$$(90^\circ \text{ counterclockwise rotation of } \mathbf{F}(x,y)) = \begin{pmatrix} -q(x,y) \\ p(x,y) \end{pmatrix}.$$

The idea is the following:

Idea

$$\mathbf{F} \cdot (90^\circ \text{ clockwise rotation of } \mathbf{r}') = (90^\circ \text{ counterclockwise rotation of } \mathbf{F}) \cdot \mathbf{r}'.$$

So what we've done is put the rotation thing onto the vector field instead.

Proof of the equation. To spell this out, imagine that $\mathbf{r}'(t) = \begin{pmatrix} r_1'(t) \\ r_2'(t) \end{pmatrix}$, meaning that its 90° clockwise rotation is $\begin{pmatrix} r_2'(t) \\ -r_1'(t) \end{pmatrix}$. Then the two quantities

$$\mathbf{F} \cdot (90^\circ \text{ clockwise rotation of } \mathbf{r}') = \begin{pmatrix} p \\ q \end{pmatrix} \cdot \begin{pmatrix} r_2' \\ -r_1' \end{pmatrix}$$
$$(90^\circ \text{ counterclockwise rotation of } \mathbf{F}) \cdot \mathbf{r}' = \begin{pmatrix} -q \\ p \end{pmatrix} \cdot \begin{pmatrix} r_1' \\ r_2' \end{pmatrix}$$

and equal as both are $pr_2' - qr_1'$ (strictly speaking, this quantity should be written in full as $p(\mathbf{r}(t))r_2'(t) - q(\mathbf{r}(t))r_1'(t)$, for each time t). □

The upshot of this is that we can actually change the flux into a work integral:

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = \int_{\mathcal{C}} (\mathbf{F} \text{ rotated } 90^\circ \text{ counterclockwise}) \cdot \mathrm{d}\mathbf{r}.$$

This looks a bit better but we still want to get rid of the rotation thing. But we can, because there is a shorthand for work that uses just p and q. Specifically, since $\mathbf{F}$ rotated 90° counterclockwise $= \begin{pmatrix} q \\ -p \end{pmatrix}$, we have

$$\int_{\mathcal{C}} (\mathbf{F} \text{ rotated } 90^\circ \text{ counterclockwise}) \cdot \mathrm{d}\mathbf{r} = \int_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y).$$

In summary, we get the following more readable shorthand:

Better definition of 2D flux using work shorthand

Let $\mathbf{F}(x, y) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$ be a 2D vector field and let $\mathcal{C}$ be a path in $\mathbb{R}^2$. Then the flux of $\mathbf{F}$ through $\mathcal{C}$ is defined as

$$\int_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y).$$

Tip

For this reason, we usually prefer to rotate $\mathbf{F}$ by 90° counterclockwise (rather than rotate $\mathbf{r}'$ by 90° clockwise) when doing concrete calculation, though of course they give the same result. I think it's easier to remember and more natural this way, because it makes things more consistent with the work integral. We'll use that convention in all the examples to follow.

In particular, if $\mathcal{C}$ is a loop (and that's usually the case if we're talking about flux at all) that means we can apply Green's theorem again; the resulting theorem is called *Green's theorem in flux form*. We get that

$$\oint_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y) = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A.$$

The right-hand side is 2D divergence, so it could be condensed even further to

$$\iint_{\mathcal{R}} \nabla \cdot \mathbf{F} \, \mathrm{d}A.$$

There's like four different versions of the same expression now, so let me just put everything in one place for sanity's sake:

Memorize: Green's theorem in flux form

Suppose $\mathcal{C}$ is a closed curve oriented counterclockwise enclosing a region $\mathcal{R}$. We have

$$\underbrace{\oint_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y)}_{= \oint_{\mathcal{C}} (\mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s)} = \underbrace{\iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A}_{= \iint_{\mathcal{R}} \nabla \cdot \mathbf{F} \, \mathrm{d}A}.$$

Warning: There's no FTC for flux

2D flux is conspicuously missing from our poster in Figure 75. Through this chapter, we were able to complete an analogy to get one Stokes result by translating 2D flux into 2D work and then quoting Green's theorem. However, as far as I can tell there isn't an analog of FTC that can be made this way. So actually one good thing about the notation $\mathbf{n} \, \mathrm{d}s$ is that the presence of $\mathrm{d}s$ is a good reminder that there's no FTC result.

In other words, 2D flux is conceptually missing one red Stokes arrow compared to 2D work. (I suppose if you really missed it, you could try to force it by asking whether $\begin{pmatrix} -q \\ p \end{pmatrix}$ is conservative, but I haven't seen this done. One possible reason is that 2D flux is mostly used for closed loops $\mathcal{C}$, and Green's theorem can handle that case anyway.)

§35.4 [SIDENOTE] A picture explaining why Green's Theorem for flux should be true (not a formal proof)

We can draw a figure much like the earlier Figure 87 (from Section 34.4) for Green's Theorem for flux. Remember that the quantity

$$\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = \nabla \cdot \mathbf{F}$$

is the divergence and interprets how much $\mathbf{F}$ is moving away from the point. So instead of *spirals*, we draw little green *explosions* corresponding to how fast $\mathbf{F}$ is moving out of each individual grey cell. The picture now turns into Figure 90, and

$$\iint_{\mathcal{R}} \nabla \cdot \mathbf{F} \, \mathrm{d}A$$

is drawn as the sum of the green explosions.

Like before, everything on the inside just cancels out. So what's left over is now the measure of $\mathbf{F}$ *against* the purple walls: the dark green arrows in Figure 90. And this corresponds to the 2D flux of $\mathbf{F}$ against the purple walls, as desired.

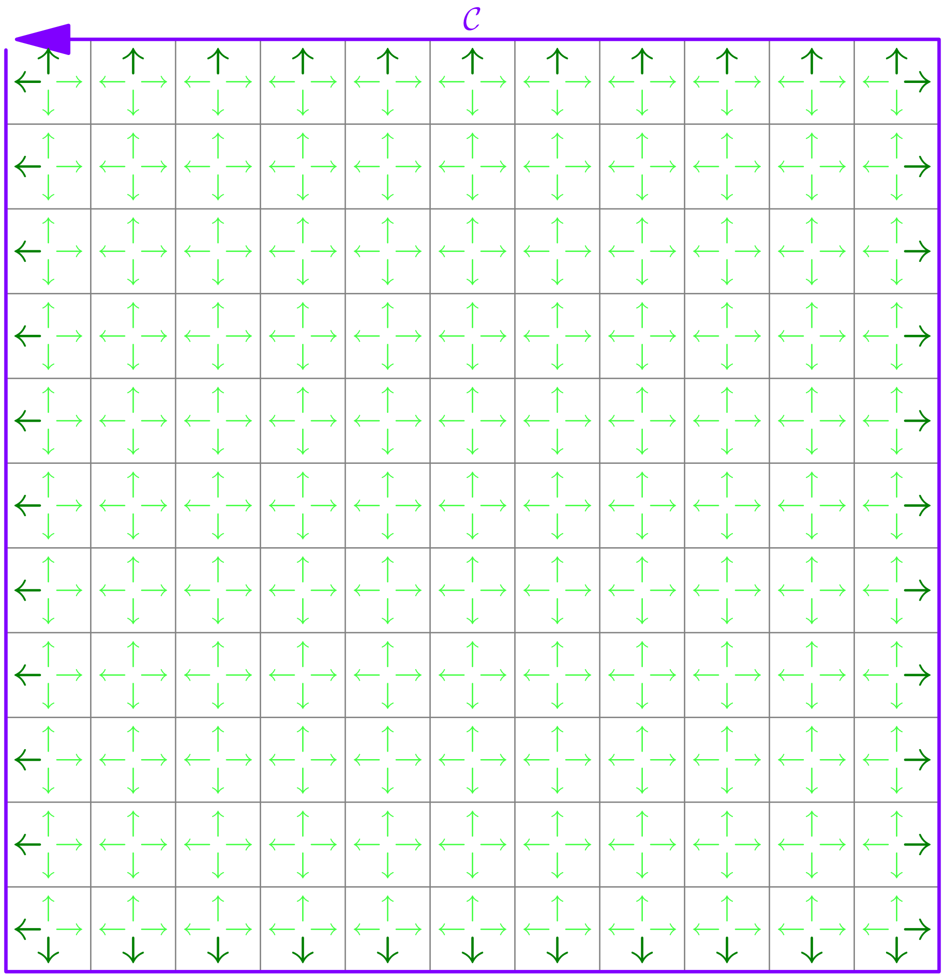

Figure 90: Roughly what Green's Theorem for flux is trying to say: summing the green explosions gives just the force of $\mathbf{F}$ against the walls.

§35.5 [RECIPE] Computing 2D flux

Recipe for computing 2D flux

1. If $\mathcal{C}$ is a closed loop, use Green's theorem as a shortcut:

$$\oint_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y) = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A.$$

2. Otherwise, do the manual recipe in Section 33.4 with $\mathbf{F} = \begin{pmatrix} p \\ q \end{pmatrix}$ replaced by its $90°$ counter-clockwise rotation $\begin{pmatrix} -q \\ p \end{pmatrix}$:
   1. Pick **any** parametrization $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ of the curve $\mathcal{C}$, including specifying the start and stop times. As described in Section 12.7, you have some freedom in how you set the parametrization.
   2. Calculate the derivative $\mathbf{r}'(t)$.
   3. Calculate the dot product $\begin{pmatrix} -q \\ p \end{pmatrix} \cdot \mathbf{r}'(t)$. (The vector field $\begin{pmatrix} -q \\ p \end{pmatrix}$ is the $90°$ counterclock-wise rotation of $\mathbf{F}$.)
   4. Integrate this from the start time to the stop time.

Here are a few examples for documentation. For each example, we actually show how to do it "manually" (by calculating a line integral) and how to do it with Green's theorem for flux.

Sample Question

Compute the flux of the vector field $\mathbf{F}(x, y) = \binom{x^2}{y^2}$ across the circle $\mathcal{C}$ defined by $x^2 + y^2 = 1$, oriented counterclockwise.

Solution. For this one, we'll actually show how to do it both using Green and manually, for comparison.

- Using Green's theorem: Green's theorem for flux states:

$$\text{Flux} = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \, \mathrm{d}A,$$

where $\mathcal{R}$ is the region enclosed by $\mathcal{C}$.

The divergence is

$$\nabla \cdot \mathbf{F} = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial y}(y^2) = 2x + 2y.$$

Therefore,

$$\text{Flux} = \iint_{\mathcal{R}} (2x + 2y) \, \mathrm{d}A.$$

Since the region $\mathcal{R}$ is the unit circle centered at the origin, and the integrand $2x + 2y$ is an odd function over this symmetric region, the integral evaluates to $\boxed{0}$. (Alternatively, integrate using polar coordinates.)

- Use the definition

$$\text{Flux} = \oint_{\mathcal{C}} (p \, \mathrm{d}y - q \, \mathrm{d}x)$$

and parametrize the curve by using

$$\mathbf{r}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \quad 0 \le t \le 2\pi$$

so

$$\mathbf{r}'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \quad 0 \le t \le 2\pi.$$

So the dot product inside the integrand is

$$\begin{aligned} (\mathbf{F} \text{ rotated } 90^\circ \text{ counterclockwise}) \cdot \mathbf{r}'(t) &= \begin{pmatrix} -q \\ p \end{pmatrix} \cdot \mathbf{r}'(t) \\ &= \begin{pmatrix} \cos(t)^2 \\ \sin(t)^2 \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \\ &= \cos^2 t \cdot \cos t - \sin^2 t \cdot (-\sin t) = \cos^3 t + \sin^3 t. \end{aligned}$$

Hence

$$\text{Flux} = \int_{t=0}^{t=2\pi} (\cos^3 t + \sin^3 t) \, \mathrm{d}t.$$

It's possible to observe from here again that the integral is symmetric; that is, for $0 \le t \le \pi$ we have $\cos^3(t) + \cos^3(t+\pi) = 0$ and $\sin^3(t) + \sin^3(t+\pi) = 0$. So again the entire contribution of the integral is $\boxed{0}$.

□

Sample Question

Compute the flux of the vector field $\mathbf{F}(x, y) = \binom{5x}{7y}$ across the square $\mathcal{C}$ with vertices at $(1, 1)$, $(-1, 1)$, $(-1, -1)$, $(1, -1)$, oriented counterclockwise.

Solution. If we were to do the line integral manually, we would have to parametrize all four sides. This would be straightforward, but it's annoying, so we'll just jump straight the shortcut with Green's theorem.

The divergence is

$$\nabla \cdot \mathbf{F} := \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = \frac{\partial}{\partial x}(5x) + \frac{\partial}{\partial y}(7y) = 12.$$

So by Green's theorem, Flux $= \iint_{\mathcal{R}} 12 \, \mathrm{d}A = 12 \operatorname{Area}(\mathcal{R}) = 12 \cdot 2^2 = \boxed{48}$ where $\mathcal{R}$ is the region enclosed by $\mathcal{C}$, a square of side length 2. □

Sample Question

Let $a, b > 0$. Compute the flux of the vector field $\mathbf{F}(x, y) = \binom{x}{y}$ across the ellipse $\mathcal{C}$ defined by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, oriented counterclockwise.

Solution. We don't really want to parametrize the ellipse[25] Again, we jump straight to Green's theorem, with

$$\nabla \cdot \mathbf{F} = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) = 1 + 1 = 2.$$

So by Green's theorem,

$$\text{Flux} = \iint_{\mathcal{R}} 2 \, \mathrm{d}A = 2 \operatorname{Area}(\mathcal{R}).$$

In a previous section (Section 23.5) we saw the area of this ellipse is was $ab\pi$; if you didn't remember this, you would go back to the change of variables and execute it. In any case, this means the flux is $2 \cdot (ab\pi) = \boxed{2ab\pi}$. □

§35.6 [RECAP] Comparison

Since the recipes for 2D flux and work look so similar, it might be helpful to compare them side by side. This comparison is shown in the table below.

[25] Although it could be done with $\mathbf{r} = (a \cos t, b \sin t)$ for $0 \le t \le 2\pi$. So it's not *that* bad.

Method	**Work** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ **(see Section 34.5)**	**2D Flux** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}s$ **(see Section 35.5)**
$\mathbf{F}$ is conservative $\implies$ FTC	If $\mathbf{F} = \nabla f$, Output $f(\text{stop}) - f(\text{start})$	*Not applicable*
$\mathcal{C}$ is a closed loop $\implies$ Green	Output $\iint_{\mathcal{R}} \underbrace{\left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right)}_{\text{2D scalar curl}} \mathrm{d}A$	Output $\iint_{\mathcal{R}} \underbrace{\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}}_{\text{Div } = \nabla \cdot \mathbf{F}} \mathrm{d}A$
Bare-hands definition Use parametrization	Output $\int (p\, \mathrm{d}x + q\, \mathrm{d}y)$ $= \int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}'(t)\, \mathrm{d}t$	Output $\int (-q\, \mathrm{d}x + p\, \mathrm{d}y)$ $= \int_{\mathcal{C}} (\mathbf{F} \text{ rot } 90^\circ \text{ ccw}) \cdot \mathbf{r}'(t)\, \mathrm{d}t$

Table 18: Comparison of the recipe for work and flux. Methods higher in the table are less work, and preferred when they apply.

If you want to see examples of this written out, see Chapter 36. There I do four examples in full, using every applicable cell of Table 18. Since it's so long, I broke it out into a separate skippable chapter.

§35.7 [EXER] Exercises

Exercise 35.1 (Suggested by Ting-Wei Chao)**.** As in Exercise 33.1 and Exercise 34.3, let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Again let

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}s$ using direct parametrization and by using Green's Theorem for flux.

Exercise 35.2. Triangle ABC has vertices $A = (-5, 0)$, $B = (9, 0)$, and C on the positive y-axis. The flux of the vector field

$$\mathbf{F}(x, y) = \begin{pmatrix} x + 7y^2 \\ x^2 + 7y \end{pmatrix}$$

across the perimeter of ABC, oriented counterclockwise, is 672. Compute the length of the perimeter of ABC.

Chapter 36. Way too many examples of work and 2D flux

This entire chapter is **review and examples only** and **can be skipped** if you know what you're doing.

The goal of this chapter is to fully write out several examples of Table 18. We'll show the entire table with four situations:

- The conservative field $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ over the unit circle oriented counterclockwise (a closed loop).
- The conservative field $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ over the line segment from $(1,4)$ to $(3,9)$.
- The non-conservative field $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ over the unit circle oriented counterclockwise (a closed loop).
- The non-conservative field $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ over the line segment from $(1,4)$ to $(3,9)$.

§36.1 Example with $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ and $\mathcal{C}$ the unit circle

Sample Question

Let $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ and let $\mathcal{C}$ be the unit circle oriented counterclockwise. Evaluate $\int \mathbf{F} \cdot d\mathbf{r}$ and $\int \mathbf{F} \cdot \mathbf{n} \, ds$.

We use the parametrization

$$\mathbf{r}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \qquad 0 \le t \le 2\pi$$

so

$$\mathbf{r}'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \qquad 0 \le t \le 2\pi.$$

In this case all five methods are applicable, see the table below.

Method	Work $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ (see Section 34.5)	2D Flux $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, ds$ (see Section 35.5)
$\mathbf{F}$ is conservative $\implies$ FTC	$f(1,0) - f(1,0) = 0$	*Not applicable*
$\mathcal{C}$ is a closed loop $\implies$ Green	$\iint_{\mathcal{R}} \underbrace{1-1}_{\text{2D scalar curl}} \, dA = 0$	Output $\iint_{\mathcal{R}} \underbrace{2+2}_{\text{Div } = \nabla \cdot \mathbf{F}} \, dA = 4\pi$
Bare-hands definition Use parametrization	$\int_{t=0}^{2\pi} \begin{pmatrix} 2\cos(t)+\sin(t) \\ \cos(t)+2\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} dt$	$\int_{t=0}^{2\pi} \begin{pmatrix} -(\cos(t)+2\sin(t)) \\ 2\cos(t)+\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} dt$

Table 19: For $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ which is conservative, with potential function $f(x,y) = x^2 + xy + y^2$, around the unit circle.

§36.1.1 Using FTC for work

The line integral is trivially zero: we don't even have to compute the potential function, because the FTC implies that we'll get $f(1,0) - f(1,0) = \boxed{0}$. In fact the potential function is $f(x,y) = x^2 + xy + y^2$ but we won't use this until the next example.

§36.1.2 Green's theorem for work

If you missed that the vector field was conservative, and you use Green's theorem, you unsurprisingly get 0 for the 2D scalar curl:

$$\iint_{\mathcal{R}} \underbrace{1-1}_{\text{2D scalar curl}} \, \mathrm{d}A = \iint_{\mathcal{R}} 0 \, \mathrm{d}A = \boxed{0}.$$

Conservative functions have vanishing curl. So actually even if you don't notice the field is conservative to start, when you try to apply Green's theorem you'll get a rather rude reminder when you realize you're just integrating the 0 function.

§36.1.3 Bare-hands for work

For the work integral, you compute it as follows:

$$\begin{aligned}
\mathbf{F}(\cos t, \sin t) \cdot \mathbf{r}'(t) &= \begin{pmatrix} 2\cos(t) + \sin(t) \\ \cos(t) + 2\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \\
&= (2\cos t + \sin t)(-\sin t) + (\cos t + 2\sin t)(\cos t) \\
&= -2\cos t \sin t - \sin^2 t + \cos^2 t + 2\cos t \sin t \\
&= \cos^2 t - \sin^2 t = \cos(2t).
\end{aligned}$$

So the integral becomes:

$$\int_0^{2\pi} \cos(2t) \, \mathrm{d}t = \boxed{0}$$

because it's an integral over two full periods of the cosine function, hence 0. (Alternatively, write $\left[\frac{\sin 2t}{2}\right]_{t=0}^{2\pi} = \frac{\sin 4\pi}{2} - \frac{\sin 0}{2} = 0 - 0 = 0$.)

§36.1.4 Green's theorem for flux

For flux, we don't get a fundamental theorem of calculus anyway, but the divergence is $2 + 2 = 4$ everywhere, which is a constant, so the flux works out to $4 \operatorname{Area}(\mathcal{R})$, which is just $\boxed{4\pi}$.

§36.1.5 Bare-hands for flux

For the flux integral, rotate the vector for the vector field (that is, look at $-q \, \mathrm{d}x + p \, \mathrm{d}y$) to get the dot product

$$\begin{aligned}
\begin{pmatrix} -(\cos(t) + 2\sin(t)) \\ 2\cos(t) + \sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} &= (2\cos t + \sin t)\cos t + (\cos t + 2\sin t)\sin t \\
&= 2\cos^2 t + \sin t \cos t + \cos t \sin t + 2\sin^2 t \\
&= 2(\cos^2 t + \sin^2 t) + 2\sin t \cos t \\
&= 2(1) + \sin 2t = 2 + \sin 2t.
\end{aligned}$$

Consequently,

$$\int_{t=0}^{2\pi} (2 + \sin 2t) \, \mathrm{d}t = 4\pi + \int_{t=0}^{2\pi} \sin 2t \, \mathrm{d}t = \boxed{4\pi}.$$

since $\int_{t=0}^{2\pi} \sin 2t$ is an integral over two full periods of the sine function.

§36.2 Example with $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ and $\mathcal{C}$ a line segment

Sample Question

Let $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ and let $\mathcal{C}$ be the path from $(1,4)$ to $(3,9)$. Evaluate $\int \mathbf{F} \cdot d\mathbf{r}$ and $\int \mathbf{F} \cdot \mathbf{n}\, ds$.

We use the same vector field, but this time we parametrize our line segment

$$\mathbf{r}(t) = \begin{pmatrix} 1+2t \\ 4+5t \end{pmatrix} \qquad 0 \le t \le 1$$

so

$$\mathbf{r}'(t) = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \qquad 0 \le t \le 2\pi.$$

This time, our table looks like this:

Method	**Work** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ **(see Section 34.5)**	**2D Flux** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, ds$ **(see Section 35.5)**
$\mathbf{F}$ is conservative $\implies$ FTC	$f(3,9) - f(1,4) = 117 - 21 = 96$	*Not applicable*
$\mathcal{C}$ is a closed loop $\implies$ Green	*Cannot use here*	*Cannot use here*
Bare-hands definition Use parametrization	$\int_{t=0}^{1} \begin{pmatrix} 2(1+2t)+(4+5t) \\ (1+2t)+2(4+5t) \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} dt$	$\int_{t=0}^{1} \begin{pmatrix} -((1+2t)+2(4+5t)) \\ 2(1+2t)+(4+5t) \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} dt$

Table 20: For $\mathbf{F} = \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix}$ which is conservative, with potential function $f(x,y) = x^2 + xy + y^2$, but this time on the line segment from $(1,4)$ to $(3,9)$.

As always, the bare-hands method is the most work, but for the flux integral we don't really have a choice because no other method is possible.

§36.2.1 Using FTC

This time, we will actually use the potential function

$$f(x,y) = x^2 + xy + y^2$$

(or really $f(x,y) = x^2 + xy + y^2 + C$ for any constant C). So we can short-circuit the entire line integral by simply evaluating

$$f(3,9) - f(1,4) = 117 - 21 = \boxed{96}.$$

§36.2.2 Bare-hands for work

For the work integral, first expand

$$\mathbf{F}(1+2t, 4+5t) = \begin{pmatrix} 2(1+2t)+(4+5t) \\ (1+2t)+2(4+5t) \end{pmatrix} = \begin{pmatrix} 2+4t+4+5t \\ 1+2t+8+10t \end{pmatrix} = \begin{pmatrix} 6+9t \\ 9+12t \end{pmatrix}.$$

Hence the dot product is

$$\mathbf{F}(1+2t, 4+5t) \cdot \mathbf{r}'(t) = \begin{pmatrix} 6+9t \\ 9+12t \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} = 2(6+9t) + 5(9+12t) = 57 + 78t.$$

Integrating this gives

$$\int_{t=0}^{1} (57+78t)\,\mathrm{d}t = \left[57t + 39t^2\right]_0^1 = 57(1) + 39(1)^2 - 0 = 57 + 39 = \boxed{96}.$$

§36.2.3 Bare-hands for flux

For the flux integral, rotate the vector for the vector field (that is, look at $-q\,\mathrm{d}x + p\,\mathrm{d}y$) to get the dot product

$$\begin{pmatrix} -(9+12t) \\ 6+9t \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} = -2(9+12t) + 5(6+9t) = 12 + 21t$$

Integrating this gives

$$\int_{t=0}^{1} (12+21t)\,\mathrm{d}t = \left[12t + \frac{21}{2}t^2\right]_{t=0}^{1} = \boxed{\frac{45}{2}}.$$

§36.3 Example with $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ and $\mathcal{C}$ the unit circle

Sample Question

Let $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ and let $\mathcal{C}$ be the unit circle oriented counterclockwise. Evaluate $\int \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ and $\int \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}s$.

Green's theorem works readily here because $\mathcal{C}$ is closed. You can also do parametrization, which is disgusting, but it works.

Method	**Work** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ **(see Section 34.5)**	**2D Flux** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}s$ **(see Section 35.5)**
$\mathbf{F}$ is conservative $\implies$ FTC	*Cannot use here*	*Not applicable*
$\mathcal{C}$ is a closed loop $\implies$ Green	$\iint_{\mathcal{R}} \underbrace{0-3}_{\text{2D scalar curl}}\,\mathrm{d}A = 0$	Output $\iint_{\mathcal{R}} \underbrace{2x+5}_{\text{Div }=\nabla\cdot\mathbf{F}}\,\mathrm{d}A = 5\pi$
Bare-hands definition Use parametrization	$\int_{t=0}^{2\pi} \begin{pmatrix} \cos(t)^2+3\sin(t) \\ 5\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}\,\mathrm{d}t$	$\int_{t=0}^{2\pi} \begin{pmatrix} -5\sin(t) \\ \cos(t)^2+3\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}\,\mathrm{d}t$

Table 21: For $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ which is not conservative, integrated over the unit circle.

§36.3.1 Green's theorem for work

For the work version, we do

$$\begin{aligned} \frac{\partial q}{\partial x} &= \frac{\partial}{\partial x}(5y) = 0 \\ \frac{\partial p}{\partial y} &= \frac{\partial}{\partial y}(x^2+3y) = 3. \end{aligned}$$

so the answer is $\iint_{\mathcal{R}} (0-3)\,\mathrm{d}A = \boxed{-3\pi}$.

§36.3.2 Bare-hands for work

We need to compute

$$\int_{t=0}^{2\pi} \begin{pmatrix} \cos(t)^2 + 3\sin(t) \\ 5\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \,\mathrm{d}t.$$

Expanding the dot product gives

$$\int_{t=0}^{2\pi} (-\cos^2 t \sin t - 3\pi \sin^2 t + 5\pi \sin t \cos t)\,\mathrm{d}t.$$

Compute each integral separately:

1. $\int_{t=0}^{2\pi} \cos^2 t \sin t \,\mathrm{d}t$: Let $u = \cos t$, then $\mathrm{d}u = -\sin t \,\mathrm{d}t$.

$$\int \cos^2 t \sin t \,\mathrm{d}t = -\int u^2 \,\mathrm{d}u = -\frac{u^3}{3} + C = -\frac{\cos^3 t}{3} + C.$$

   Evaluate from 0 to 2π:

$$\left[-\frac{\cos^3 t}{3}\right]_{t=0}^{2\pi} = -\frac{\cos^3(2\pi)}{3} + \frac{\cos^3 0}{3} = -\frac{1}{3} + \frac{1}{3} = 0.$$

2. $\int_{t=0}^{2\pi} \sin^2 t \,\mathrm{d}t$: Use the identity $\sin^2 t = \frac{1-\cos 2t}{2}$:

$$\int_{t=0}^{2\pi} \sin^2 t \,\mathrm{d}t = \frac{1}{2}\int_{t=0}^{2\pi} (1-\cos 2t)\,\mathrm{d}t = \frac{1}{2}\left[t - \frac{\sin 2t}{2}\right]_{t=0}^{2\pi} = \frac{1}{2}(2\pi - 0) = \pi.$$

3. $\int_{t=0}^{2\pi} \sin t \cos t \,\mathrm{d}t$:

Use the identity $\sin t \cos t = \frac{\sin 2t}{2}$:

$$\int_{t=0}^{2\pi} \sin t \cos t \,\mathrm{d}t = \frac{1}{2}\int_{t=0}^{2\pi} \sin 2t \,\mathrm{d}t = 0.$$

(Since the integral of sine over its full period is zero.) Combine the results to get

$$-0 - 3 \cdot \pi + 5 \cdot 0 = \boxed{-3\pi}.$$

§36.3.3 Green's theorem for flux

For the flux version, it's instead

$$\begin{aligned} \frac{\partial p}{\partial x} &= \frac{\partial}{\partial x}(x^2 + 3y) = 2x \\ \frac{\partial q}{\partial y} &= \frac{\partial}{\partial y}(5y) = 5. \end{aligned}$$

so the flux is

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \,\mathrm{d}s = \iint_{\mathcal{R}} (2x + 5)\,\mathrm{d}A.$$

By symmetry, we have

$$\iint_{\mathcal{R}} x \,\mathrm{d}A = 0$$

and we also have

$$\iint_{\mathcal{R}} 5 \, \mathrm{d}A = 5\pi$$

so we get the answer $0 + 5\pi = \boxed{5\pi}$.

> **Polar coordinates is fine too for the flux one**
>
> If you don't notice the symmetry trick, you can use polar coordinates too. Write $2x + 5 = 2r\cos\theta + 5$ and set up the flux integral as:
>
> $$\begin{aligned} \oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s &= \int_{\theta=0}^{2\pi} \int_{r=0}^{1} (2r\cos\theta + 5) r \, \mathrm{d}r \, \mathrm{d}\theta \\ &= \int_{\theta=0}^{2\pi} \int_{r=0}^{1} (2r^2\cos\theta + 5r) \, \mathrm{d}r \, \mathrm{d}\theta \\ &= 2 \int_{\theta=0}^{2\pi} \cos\theta \int_{r=0}^{1} r^2 \, \mathrm{d}r \, \mathrm{d}\theta + 5 \int_{\theta=0}^{2\pi} \int_{r=0}^{1} r \, \mathrm{d}r \, \mathrm{d}\theta. \end{aligned}$$
>
> The inner integrals are $\int_{r=0}^{1} r^2 \, \mathrm{d}r = \frac{1}{3}$ and $\int_{r=0}^{1} r \, \mathrm{d}r = \frac{1}{2}$, so we get the same answer
>
> $$2 \cdot \frac{1}{3} \int_0^{2\pi} \cos\theta \, \mathrm{d}\theta + 5 \cdot \frac{1}{2} \int_0^{2\pi} \mathrm{d}\theta = \frac{2}{3} \cdot 0 + 5 \cdot \frac{1}{2} \cdot 2\pi = 5\pi.$$

§36.3.4 Bare-hands for flux

We need to compute

$$\int_{t=0}^{2\pi} \begin{pmatrix} -5\sin(t) \\ \cos(t)^2 + 3\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \mathrm{d}t.$$

Expand the dot product:

$$\int_{t=0}^{2\pi} (\cos^3 t + 3\sin t \cos t + 5\sin^2 t) \, \mathrm{d}t.$$

Compute each integral separately:

1. $\int_{t=0}^{2\pi} \cos^3 t \, \mathrm{d}t$: Use the identity $\cos^3 t = \frac{3\cos t + \cos 3t}{4}$:

$$\int_{t=0}^{2\pi} \cos^3 t \, \mathrm{d}t = \frac{3}{4} \int_{t=0}^{2\pi} \cos t \, \mathrm{d}t + \frac{1}{4} \int_{t=0}^{2\pi} \cos 3t \, \mathrm{d}t = 0 + 0 = 0.$$

2. $\int_{t=0}^{2\pi} \sin t \cos t \, \mathrm{d}t$: Use the identity $\sin t \cos t = \frac{\sin 2t}{2}$:

$$\int_{t=0}^{2\pi} \sin t \cos t \, \mathrm{d}t = \frac{1}{2} \int_{t=0}^{2\pi} \sin 2t \, \mathrm{d}t = 0.$$

3. $\int_{t=0}^{2\pi} \sin^2 t \, \mathrm{d}t$: Use the identity $\sin^2 t = \frac{1 - \cos 2t}{2}$:

$$\int_{t=0}^{2\pi} \sin^2 t \, \mathrm{d}t = \frac{1}{2} \int_{t=0}^{2\pi} (1 - \cos 2t) \, \mathrm{d}t = \frac{1}{2}[2\pi - 0] = \pi.$$

Combine the results to get

$$0 + 3 \cdot 0 + 5 \cdot \pi = \boxed{5\pi}.$$

§36.4 Example with $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ and $\mathcal{C}$ a line segment

Sample Question

Let $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ and let $\mathcal{C}$ be the path from $(1,4)$ to $(3,9)$. Evaluate $\int \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ and $\int \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s$.

Here in both cases we have to bite the bullet — none of our shortcuts apply. As before we use the parametrization

$$\mathbf{r}(t) = \begin{pmatrix} 1+2t \\ 4+5t \end{pmatrix} \qquad 0 \le t \le 1$$

with

$$\mathbf{r}'(t) = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \qquad 0 \le t \le 2\pi.$$

Method	**Work** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ **(see Section 34.5)**	**2D Flux** $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s$ **(see Section 35.5)**
$\mathbf{F}$ is conservative $\implies$ FTC	*Cannot use here*	*Not applicable*
$\mathcal{C}$ is a closed loop $\implies$ Green	*Cannot use here*	*Cannot use here*
Bare-hands definition Use parametrization	$\int_{t=0}^{1} \begin{pmatrix} (1+2t)^2+3(4+5t) \\ 5(4+5t) \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} \mathrm{d}t$	$\int_{t=0}^{1} \begin{pmatrix} -5(4+5t) \\ (1+2t)^2+3(4+5t) \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} \mathrm{d}t$

Table 22: For $\mathbf{F} = \begin{pmatrix} x^2+3y \\ 5y \end{pmatrix}$ which is not conservative, integrated over the unit circle.

§36.4.1 Bare-hands for work

For the work integral, substitute $1+2t$ and $4+5t$ into $\mathbf{F}$:

$$\mathbf{F}(1+2t, 4+5t) = \begin{pmatrix} (1+2t)^2 + 3(4+5t) \\ 5(4+5t) \end{pmatrix} = \begin{pmatrix} 4t^2 + 19t + 13 \\ 25t + 20 \end{pmatrix}.$$

Then the dot product is Dot product:

$$\begin{pmatrix} 4t^2 + 19t + 13 \\ 25t + 20 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} = 2(4t^2 + 19t + 13) + 5(25t + 20) = 8t^2 + 163t + 126.$$

Hence

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{t=0}^{1} (8t^2 + 163t + 126) \, \mathrm{d}t.$$

Integrate term by term:

$$\begin{aligned}
\int_{t=0}^{1} 126\,\mathrm{d}t &= 126, \\
\int_{t=0}^{1} 163t\,\mathrm{d}t &= \left[\frac{163}{2}t^2\right]_{t=0}^{1} = \frac{163}{2}(1)^2 - \frac{163}{2}(0)^2 = \frac{163}{2}, \\
\int_{t=0}^{1} 8t^2\,\mathrm{d}t &= \left[\frac{8}{3}t^3\right]_{t=0}^{1} = \frac{8}{3}(1)^3 - \frac{8}{3}(0)^3 = \frac{8}{3}.
\end{aligned}$$

Combine the results:

$$\int_{\mathcal{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} = 126 + \frac{163}{2} + \frac{8}{3} = \boxed{\frac{1261}{6}}.$$

§36.4.2 Bare-hands for flux

For the flux integral, instead do the dot product

$$\begin{pmatrix} -(25t+20) \\ 4t^2+19t+13 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} = -2(25t+20) + 5\left(4t^2+19t+13\right) = 20t^2+45t+25.$$

Integrate term by term again:

$$\begin{aligned}
\int_{t=0}^{1} 25\,\mathrm{d}t &= 25, \\
\int_{t=0}^{1} 45t\,\mathrm{d}t &= \left[\frac{45}{2}t^2\right]_{t=0}^{1} = \frac{45}{2}(1)^2 - \frac{45}{2}(0)^2 = \frac{45}{2}, \\
\int_{t=0}^{1} 20t^2\,\mathrm{d}t &= \left[\frac{20}{3}t^3\right]_{t=0}^{1} = \frac{20}{3}(1)^3 - \frac{20}{3}(0)^3 = \frac{20}{3}.
\end{aligned}$$

Combine the results:

$$\int_{\mathcal{C}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}s = 25 + \frac{45}{2} + \frac{20}{3} = \frac{150}{6} + \frac{135}{6} + \frac{40}{6} = \boxed{\frac{325}{6}}.$$

Chapter 37. Practice midterm for 2D topics in Parts Golf, Hotel, India

This is a practice midterm that was given on November 13, 2024, covering topics in Part Golf, Hotel, and India, but only the 2D topics in these parts. Solutions are in Chapter 50.

Exercise 37.1. Another butterfly is fluttering in the xy plane with position $\mathbf{r}(t) = \langle \sin(t), \sin(t) \rangle$. Let $\mathcal{C}$ denote its trajectory between $0 \le t \le 2\pi$. Compute $\int_{\mathcal{C}} (x \, \mathrm{d}x)$ and $\int_{\mathcal{C}} (y \, \mathrm{d}x)$.

Exercise 37.2. Let $\mathcal{C}$ denote the unit circle $x^2 + y^2 = 1$ oriented counterclockwise, and consider the vector field $\mathbf{F}(x, y) = \langle x + 2y, 4x + 8y \rangle$. Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ and $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s$.

Exercise 37.3. Compute all real numbers k for which the following region has area π:

$$(kx + y)^2 + (x + ky)^2 \le \frac{1}{4}.$$

Exercise 37.4. Compute the center of mass of the region where $y \ge 0$ and $3x^2 \le y^2 \le 9 - x^2$, assuming constant density.

Exercise 37.5. Let $\mathcal{C}$ denote any path from $(0, 0)$ to (π, π). Determine the unique function $h(x)$ for which $\mathbf{F}(x, y) = \langle xy + \cos(x), h(x) + \cos(y) \rangle$ is conservative, and moreover $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 0$.

Exercise 37.6. Assume log is base $e \approx 2.718$. Use any method you want to compute

$$\int_{x=0}^{(e-1)^2} \log\left(\sqrt{x} + 1\right) \mathrm{d}x.$$

Recommended approach: view the integral as the area under a curve, then switch from vertical to horizontal slicing.

Part Juliett: Flux integrals of vector fields over a surface

For comparison, Part Juliett corresponds to §17.7, §18, §19, §21 of Poonen's notes.

Chapter 38. Flux

We now discuss (3D) flux, the final type of vector field integral that we haven't seen yet. This is the final cell in the poster Figure 75 that we haven't met yet.

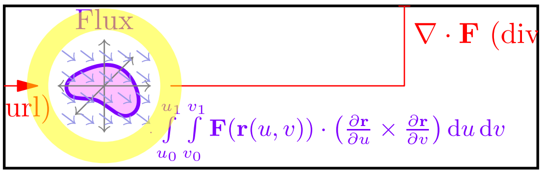

Figure 91: The flux integral for a surface circled in our poster Figure 75.

This chapter will be pretty reminiscent of Chapter 33. We'll start by giving a "bare-hands" definition of the flux through a parametrized surface. It will be usable, but pretty cumbersome, so in the next chapter Chapter 39 we'll immediately try to find ways to shortcut it. For 18.02, the methods available to you will be

- Bare-hands parametrization (covered here)
  - Even here, magic formulas can save you a lot of work — see Table 23.
- Shortcut: Transforming to a surface area integral (covered in Section 38.6)
- Shortcut: The divergence theorem, by converting to a 3D volume integral (covered in the next chapter Chapter 39)

§38.1 [TEXT] The definition of flux using bare-hands parametrization

Definition of flux

Let $\mathbf{r}(u,v) : \mathcal{R} \to \mathbb{R}^3$ parametrize an oriented surface $\mathcal{S}$ in $\mathbb{R}^3$. The flux of a vector field $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ through $\mathcal{S}$ is defined by

$$\iint_{\mathcal{R}} \mathbf{F}(\mathbf{r}(u,v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) \mathrm{d}u\,\mathrm{d}v.$$

(We'll explain what "oriented" means in the next section.)

Type signature

Flux requires two inputs: an *oriented surface* $\mathcal{S}$ and a *vector field* $\mathbf{F}$.

Yes, there's that hideous cross product again. Naturally, people have shorthand to make this easier to swallow: this time either

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S$$

is used to sweep everything under the carpet. That is, $\mathrm{d}\mathbf{S}$ and $\mathbf{n}\,\mathrm{d}S$ are both shorthands for the longer $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v$. We'll usually prefer $\mathbf{n}\,\mathrm{d}S$ in this book.

I promised you back in the surface area chapter (Chapter 29) that at some point you'd need the whole cross product and not just its magnitude, and here we are! In fact, the absolute value being gone is in some sense an *improvement*: I would argue $\left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle$ is less messy than $\sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$.

We'll do example calculations in a moment, but let me first talk about how to think about this, and also explain what the adjective "oriented".

§38.2 [TEXT] Aquatic interpretation of flux

Flip back to Figure 81 for a moment. Back when we were talking about the work integral $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$, I told you to visualize the work as adding the dot products of the force vectors with tangent vectors.

The interpretation here is similar to 2D flux. You should imagine the surface $\mathcal{S}$ as some membrane in the water; then the flux measures the rate at which water moves through it.

To make this picture complete, I need to tell you about orientation. Remember, back when we had work integrals, a curve $\mathcal{C}$ wasn't just a bunch of points; we also had to tell you which point was the "start" and which one was the "stop". In other words, work integrals operate on a curve with a *direction*.

Something similar happens with flux integrals over surfaces: in addition to the actual points, we need to specify an *orientation*. To be more precise, at every point P of the surface $\mathcal{S}$, the cross product from our parametrization could point in one of two opposite directions.

Definition of orienting a surface

To *orient* the surface $\mathcal{S}$ is to specify, at each point, which way you want the cross product of your parametrization to point.

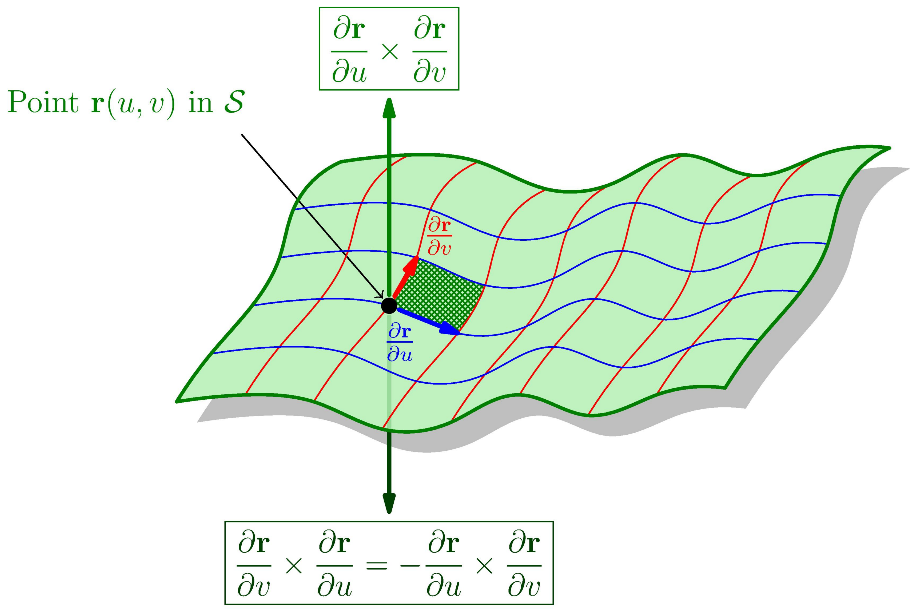

Figure 92: The normal vector from before, and its negation. Note that when we swap u and v, the vector flips the other way to the negative. Hence when parametrizing a surface, the order of u and v induces an orientation on the surface.

Algebraically, this corresponds to choosing the *order of u and v*; as if you flip the order of the two parameters it will negate the entire cross product:

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} = -\left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial u}\right).$$

Hence the flux will get negated too. This sign issue is disorienting because it wasn't present for work, where "start to stop" was pretty easy to think about; we'll give more examples momentarily.

Going back to our new flux integral, we need to visualize the dot products

$$\iint_{\mathcal{R}} \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) \mathrm{d}u\, \mathrm{d}v.$$

The $\mathbf{F}$ is still the force vector, and as we describe earlier, the vector

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$$

represents a normal vector to the surface at each point. We draw this in Figure 93.

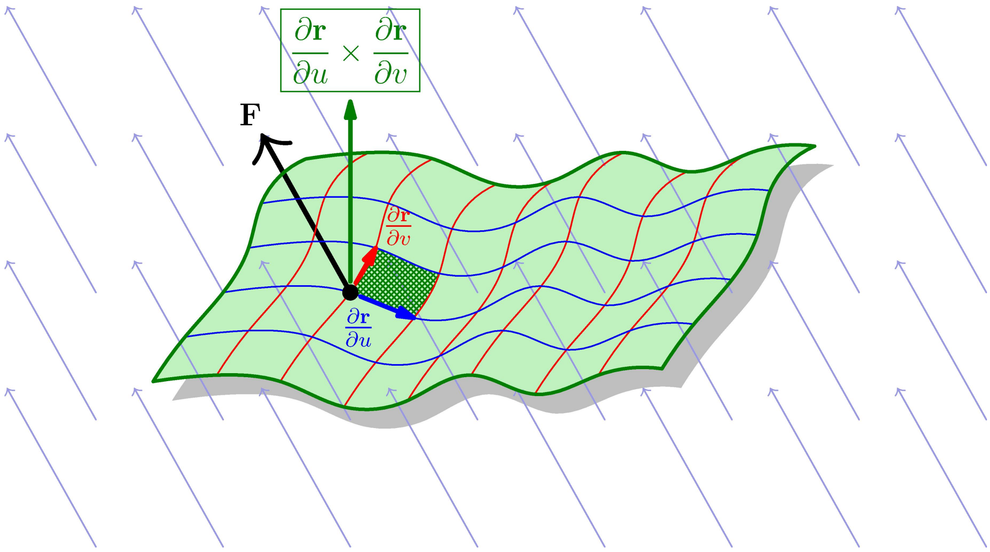

Figure 93: Picture of a parametrized surface sitting in a vector field $\mathbf{F}$. At each point, we take the dot product of the vector field $\mathbf{F}$ at that point (drawn in black here) with a normal vector on the surface given by the same cross product that we considered for surface area (drawn in green here). The flux can be thought of as the sum of all the dot products across the whole surface.

The dot product of Figure 93 should match your aquatic intuition. For our oriented surface, the dot product is large when the force is moving along the same direction as the normal vector. That matches our description of a water current puncturing the surface. On the other hand, if the force had been moving mostly parallel to the surface, then the dot products and hence flux are both close to zero.

§38.3 [TEXT] More on orientation

Here's another example of an orientation to make things less abstract.

Example: Orienting a sphere

Let's consider the sphere $x^2 + y^2 + z^2 = 1$ with $z > 0$. For each point P on the sphere, the normal vector to the sphere at P either points straight towards the center of from P, or away from the center of P.

What does this correspond to algebraically? We consider two possible ways to parametrize the sphere that differ only in the order.

- Let's imagine we used a spherical parametrization of the hemisphere as

$$\mathbf{r}(\varphi, \theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$$

  where $0 \le \varphi \le \pi$ and $0 \le \theta \le 2\pi$. If we grinded out the cross product, you would find that (see Chapter 30 to see this written out)

$$\frac{\partial \mathbf{r}}{\partial \varphi} \times \frac{\partial \mathbf{r}}{\partial \theta} = \sin\varphi \cdot (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi) = \sin\varphi \cdot \mathbf{r}(\varphi, \theta).$$

  At each point $P = \mathbf{r}(\varphi, \theta) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$ of the sphere, this points outwards (since $\sin\varphi \ge 0$), so this would be a parametrization of the sphere with all the cross products pointing out.

- But what if we had flipped the order of φ and θ? That is, suppose we used

$$\mathbf{r}(\theta, \varphi) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)$$

  where $0 \le \theta \le 2\pi$ and $0 \le \varphi \le \pi$ instead. Then the cross product will get negated:

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} = -\sin\varphi \cdot (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi) = -\sin\varphi \cdot \mathbf{r}(\varphi, \theta).$$

  And now at every point, the cross product points inside the sphere instead!

So which one of these orientations is "correct"? Well, that's why a convention is needed. It's just like when we computed work or flux integrals of circles in 2D, we had to say "counterclockwise" or "clockwise". For this sphere we have to say "outwards" or "inwards" or something like that so that whoever is computing the flux integral knows which way to take the cross product.

In general, for surfaces where inward vs outward has an obvious meaning, the convention is usually "outward". But not all surfaces have an obvious inward vs outward (for example, the xy-plane given by $z = 0$), and in those cases an exam question should tell you which one to use for that question.

Digression: Comparison to 2D flux

In 2D flux, we had a notion of "outside" vs "inside" even for curves $\mathcal{C}$ that weren't closed, because we had a notion of 90° clockwise vs 90° counterclockwise. We don't have this in 3D space, sadly, which is why we resort to normal vectors instead.

§38.4 [TEXT] Magic formulas for the cross product (reprise)

TL;DR: cross products are too annoying, so we pre-compute them all.

In Chapter 30 I gave you Table 14 which let you bypass the cross product step when calculating surface area, and it still works here. But I'm actually going to rewrite the table to connect it to the shorthand

$\mathbf{n}\,\mathrm{d}S$. In fact, people often split the shorthand $\mathbf{n}\,\mathrm{d}S$ into two parts: $\mathbf{n}$ is the unit vector in the *direction* of the cross product, while $\mathrm{d}S$ represents the absolute value with $\mathrm{d}u\,\mathrm{d}v$ tacked on. In symbols, this says

$$\mathbf{n} := \frac{\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}}{\left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right|} \quad \text{and} \quad \mathrm{d}S := \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\,\mathrm{d}v.$$

(So $\mathbf{n}\,\mathrm{d}S$ is indeed the full cross product, as the two absolute value things cancel.)

The reason people will separate it like this is to make the geometry a bit easier to think about. Remember from back in Chapter 6 that a cross product has two pieces of information: a *direction* (meant to give two right angles) and a *magnitude* (meant to interpret area). The point of separating the shorthand is to make these correspond to $\mathbf{n}$ and $\mathrm{d}S$ respectively.

Personally, I don't see the point of decomposing the information like this, since you need the entire cross product when you do calculation anyway. But a lot of people do it. So by popular request, here's a version of Table 14 that separates the components. I think this separation only really helps with the fourth and fifth rows, because back in Chapter 30 we described ways to remember $\mathrm{d}S$ geometrically for the cylinder and the sphere. (For the cylinder, $\mathrm{d}S \approx \frac{\mathrm{d}V}{\mathrm{d}r}$; for the sphere, $\mathrm{d}S \approx \frac{\mathrm{d}V}{\mathrm{d}\rho}$.) For the first and second rows, you should just remember the fifth column.

Surface	**Param's**	$\mathbf{n}$ **(unit vec)**	$\mathrm{d}S$	$\mathbf{n}\,\mathrm{d}S = \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\,\mathrm{d}u\,\mathrm{d}v$
$z = f(x, y)$	(x, y)	$\frac{\left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right\rangle}{\sqrt{1+\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2}}$	$\sqrt{1+\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2}\,\mathrm{d}x\,\mathrm{d}y$	$\left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right\rangle \mathrm{d}x\,\mathrm{d}y$
Level surface $g(x, y, z) = c$ over an xy-region	(x, y)	$\pm\frac{\nabla g}{\lvert\nabla g\rvert}$	$\frac{\lvert\nabla g\rvert}{\lvert\partial g/\partial z\rvert}\,\mathrm{d}x\,\mathrm{d}y$	$\frac{\nabla g}{\partial g/\partial z}\,\mathrm{d}x\,\mathrm{d}y$
Flat surface $z = c$	(x, y)	$\langle 0, 0, 1\rangle$	$\mathrm{d}x\,\mathrm{d}y$	$\langle 0, 0, 1\rangle\,\mathrm{d}x\,\mathrm{d}y$
Cylindrical coords with fixed R $\mathbf{r}(\theta, z) = (R\cos\theta, R\sin\theta, z)$	(θ, z)	$\langle \cos\theta, \sin\theta, 0\rangle$	$R\,\mathrm{d}\theta\,\mathrm{d}z$	$\langle R\cos\theta, R\sin\theta, 0\rangle\,\mathrm{d}\theta\,\mathrm{d}z$
Spherical coords with fixed R $\mathbf{r}(\varphi, \theta) = (R\sin\varphi\cos\theta, R\sin\varphi\sin\theta, R\cos\varphi)$	(φ, θ)	$\frac{1}{R}\cdot\mathbf{r}(\varphi, \theta)$ (if $0 \le \varphi \le \pi$)	$R^2\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta$ (if $0 \le \varphi \le \pi$)	$R\sin\varphi\cdot\mathbf{r}(\varphi, \theta)\,\mathrm{d}\varphi\,\mathrm{d}\theta$

Table 23: An alternate version of Table 14 written in $\mathbf{n}$ and $\mathrm{d}S$ notation. I think it's less elegant and you should just use the original Table 14, personally, but the tables are the same, so it doesn't matter which one you use.

Again, when actually doing flux calculation with bare hands, **you only need the fifth column**. And if you ever *do* need the third and fourth column for some other reason, they can be derived instantly from the fifth column anyways. So the third and fourth column are only helpful insomuch as they might make the formula for the cylinder and sphere easier to remember or more conceptually intuitive. But for practical calculation they are redundant.

§38.5 [RECIPE] Recipe for flux integrals with bare-hands parametrization

We go back to recipe format now.

Recipe for computing flux integrals with bare-hands parametrization

To compute the flux of $\mathbf{F}$ over a surface $\mathcal{S}$:

1. Get the cross product $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ for a parametrization $\mathbf{r}$ using the following checklist.
   - If you are using (x, y)-coordinates to parametrize (meaning $\mathcal{S}$ is $z = f(x, y)$ or a level surface), use the magic formulas in rows 1 or 2 of Table 23.
   - For a flat surface, it's easy (row 3 of Table 23).
   - If $\mathcal{S}$ is specifically given by cylindrical/spherical coordinates with fixed radius, use rows 4 or 5 of Table 23.
   - Otherwise, evaluate the cross product manually:
     - Pick a parametrization $\mathbf{r}(u, v) : \mathcal{R} \to \mathbb{R}^3$ of the surface $\mathcal{S}$. Sort of like in Section 12.7, you have some freedom in how you set the parametrization.
     - Compute $\frac{\partial \mathbf{r}}{\partial u}$ and $\frac{\partial \mathbf{r}}{\partial v}$ (both are three-dimensional vectors at each point).
     - Compute the cross product $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ as in Chapter 6.
2. Look at which way the cross product points. Does it point the direction you want? If not, negate the entire cross product (equivalently, swap the order of u and v) before going on.
3. Compute the dot product
$$\mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right).$$
   This gives you a number at every point on the parametrizing region $\mathcal{R}$.
4. Integrate the entire thing over $\mathcal{R}$ using any of the methods for double integrals (such as horizontal/vertical slicing, polar coordinates, change of variables, etc.).

Let's give one example corresponding to each row of Table 23.

Sample Question

Consider the surface $\mathcal{S}$ defined by $z = x^3 + y^3$ for $0 \le x \le 1$ and $0 \le y \le 1$, with the normal vector oriented upwards (i.e. with positive z-component). Let $\mathbf{F}(x, y, z) = \begin{pmatrix} 1 \\ 1 \\ z \end{pmatrix}$. Compute the flux of $\mathbf{F}$ through $\mathcal{S}$.

Solution. Our parametrization of the surface $\mathcal{S}$ is by definition
$$\mathbf{r}(x, y) = (x, y, x^3 + y^3)$$
for $0 \le x \le 1$ and $0 \le y \le 1$. Accordingly, we use the first row of Table 23 with $f(x, y) = x^3 + y^3$. Compute the partial derivatives
$$\frac{\partial f}{\partial x} = 3x^2, \quad \frac{\partial f}{\partial y} = 3y^2.$$
Then by using the first row of Table 23, if we we get that the cross product at each point is given by
$$\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = \begin{pmatrix} -3x^2 \\ -3y^2 \\ 1 \end{pmatrix}.$$

At this point we have to check whether this cross product points the direction specified in the problem, or if we need to negate everything and consider $\frac{\partial \mathbf{r}}{\partial y} \times \frac{\partial \mathbf{r}}{\partial x} = \begin{pmatrix} 3x^2 \\ 3y^2 \\ -1 \end{pmatrix}$ instead. The question wanted the normal vector to be oriented upwards, and since 1 is positive, the original we had is okay; we use

$$\mathbf{n}\,\mathrm{d}S = \begin{pmatrix} 3x^2 \\ 3y^2 \\ -1 \end{pmatrix} = \begin{pmatrix} -3x^2 \\ -3y^2 \\ 1 \end{pmatrix} \mathrm{d}x\,\mathrm{d}y.$$

Now, the vector field is given at each point (x, y) by

$$\mathbf{F}(\mathbf{r}(x,y)) = \begin{pmatrix} 1 \\ 1 \\ x^3 + y^3 \end{pmatrix}.$$

So we can compute the dot product

$$\begin{aligned} \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} \right) &= (1)(-3x^2) + (1)(-3y^2) + (x^3 + y^3)(1) \\ &= -3x^2 - 3y^2 + x^3 + y^3. \end{aligned}$$

Hence the flux requested is given by

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S = \int_{x=0}^{1} \int_{y=0}^{1} (-3x^2 - 3y^2 + x^3 + y^3)\,\mathrm{d}y\,\mathrm{d}x$$

which is straightforward to evaluate:

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S &= \int_{x=0}^{1} \int_{y=0}^{1} (x^3 - 3x^2 + y^3 - 3y^2)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{x=0}^{1} \left(\int_{y=0}^{1} (x^3 - 3x^2)\,\mathrm{d}x \right) \mathrm{d}y + \int_{y=0}^{1} \left(\int_{x=0}^{1} (y^3 - 3y^2)\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_{x=0}^{1} (x^3 - 3x^2\,\mathrm{d}x) + \int_{y=0}^{1} (y^3 - 3y^2\,\mathrm{d}y) \\ &= \left[\frac{x^4}{4} - x^3 \right]_{x=0}^{1} + \left[\frac{y^4}{4} - y^3 \right]_{y=0}^{1} \\ &= -\frac{3}{4} - \frac{3}{4} = \boxed{-\frac{3}{2}}. \end{aligned}$$

□

Sample Question

Consider the upper hemisphere of the sphere defined by $x^2 + y^2 + z^2 = 25$ with the unit normal vector oriented *downwards* towards the xy-plane. Calculate the flux of the vector field $\mathbf{F} = \begin{pmatrix} yz \\ xz \\ 0 \end{pmatrix}$ through this surface.

Solution. Our parametrization of $\mathcal{S}$ is going to be

$$\mathbf{r}(x,y) = \langle x^2, y^2, \sqrt{25-(x^2+y^2)} \rangle$$

across $x^2+y^2 \le 25$. If we wanted to use the first row of the table Table 23, we would use $f(x,y) = \sqrt{25-(x^2+y^2)}$. However, square roots are annoying and we'll use the second row instead by viewing this hemisphere as a chunk of the level surface

$$g(x,y,z) = x^2+y^2+z^2$$

for the value 25. Since $\nabla g = \langle 2x, 2y, 2z \rangle$ and $\frac{\partial g}{\partial z} = 2z$, our table gives

$$\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = \frac{\nabla g}{\frac{\partial g}{\partial z}} = \frac{\langle 2x, 2y, 2z \rangle}{2z} = \begin{pmatrix} \frac{x}{z} \\ \frac{y}{z} \\ 1 \end{pmatrix}.$$

Here $z = \sqrt{25-(x^2+y^2)}$.

At this point we have to check whether this cross product points the direction specified in the problem, or if we need to negate everything and consider $\frac{\partial \mathbf{r}}{\partial y} \times \frac{\partial \mathbf{r}}{\partial x} = \begin{pmatrix} -\frac{x}{z} \\ -\frac{y}{z} \\ -1 \end{pmatrix}$ instead. This time, the question specified the normal vector should point *downwards*, towards the xy-plane. So we had better use the negative one:

$$\mathbf{n}\,\mathrm{d}S = \frac{\partial \mathbf{r}}{\partial y} \times \frac{\partial \mathbf{r}}{\partial x} = \begin{pmatrix} -\frac{x}{z} \\ -\frac{y}{z} \\ -1 \end{pmatrix} \mathrm{d}x\,\mathrm{d}y.$$

Meanwhile, the force at each point of the parametrization is given by

$$\mathbf{F}(\mathbf{r})(x,y) = \begin{pmatrix} yz \\ xz \\ 0 \end{pmatrix}.$$

So the dot product is given by

$$\mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial y} \times \frac{\partial \mathbf{r}}{\partial x} \right) = yz \cdot \left(-\frac{x}{z} \right) + xz \cdot \left(-\frac{y}{z} \right) + 0 \cdot (-1) = -2xy.$$

Hence the flux we seek is

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S = \iint_{x^2+y^2 \le 25} -2xy \,\mathrm{d}x\,\mathrm{d}y.$$

But notice that the integrand $-2xy$ is an odd function in both x and y. Since the region $x^2+y^2 \le 25$ is symmetric with respect to both axes, we don't even have to bother changing to polar coordinates; we can just deduce directly that

$$\iint_{x^2+y^2 \le 25} -2xy \,\mathrm{d}x\,\mathrm{d}y = \boxed{0}.$$

□

Sample Question

Consider the plane $x = 3$ with the normal vector oriented in the $-x$ direction, and the vector field $\mathbf{F} = \langle e^x, e^y, e^z \rangle$. Compute the flux of $\mathbf{F}$ through the portion of the plane with $y^2+z^2 \le 25$.

Solution. Let $\mathcal{S}$ be the surface of the plane mentioned. We parametrize with the variables y and z;

$$\mathbf{r}(y, z) = (3, y, z)$$

across $y^2 + z^2 \leq 25$.

This cross product is the third (easiest) row of Table 23; you just get

$$\frac{\partial \mathbf{r}}{\partial y} \times \frac{\partial \mathbf{r}}{\partial z} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Before going on, we again have to check whether the normal vector points the correct way, or we should negate it and use $\frac{\partial \mathbf{r}}{\partial z} \times \frac{\partial \mathbf{r}}{\partial y} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$ instead. The problem wants the $-x$ direction, so indeed, we take the negated one here:

$$\mathbf{n}\,\mathrm{d}S = \frac{\partial \mathbf{r}}{\partial z} \times \frac{\partial \mathbf{r}}{\partial y} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

Meanwhile, the force vector at each point is just

$$\mathbf{F}(\mathbf{r}(y, z)) = \begin{pmatrix} e^3 \\ e^y \\ e^z \end{pmatrix}.$$

The dot product is then

$$\mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial z} \times \frac{\partial \mathbf{r}}{\partial y} \right) = \begin{pmatrix} e^3 \\ e^y \\ e^z \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = -e^3.$$

Hence, the flux we seek is

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S &= \iint_{y^2+z^2\leq 25} -e^3 \,\mathrm{d}y\,\mathrm{d}z \\ &= -e^3 \cdot \operatorname{Area}(y^2 + z^2 \leq 25) = \boxed{-25\pi e^3}. \end{aligned}$$

□

Sample Question

Let $\mathcal{S}$ be the portion of the cylinder $x^2 + y^2 = 49$ where $0 \leq z \leq 10$, with normal vector oriented outwards. Calculate the flux of $\mathbf{F} = \begin{pmatrix} 3x \\ 5y \\ e^z \end{pmatrix}$ through $\mathcal{S}$.

Solution. It's natural to parametrize this with cylindrical coordinates as

$$\mathbf{r}(\theta, z) = \langle 7\cos\theta, 7\sin\theta, z \rangle$$

for $0 \leq \theta \leq 2\pi$ and $0 \leq z \leq 10$. As this is a cylinder, we use the fourth row of Table 23 to get

$$\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial z} = \begin{pmatrix} 7\cos\theta \\ 7\sin\theta \\ 0 \end{pmatrix}.$$

As before we pause to see whether this points the right way or whether we need to instead use $\frac{\partial \mathbf{r}}{\partial z} \times \frac{\partial \mathbf{r}}{\partial \theta} = \begin{pmatrix} -7\cos\theta \\ -7\sin\theta \\ 0 \end{pmatrix}$. The question specifies to orient the normal vector outwards, so we use the former one:

$$\mathbf{n}\,\mathrm{d}S = \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial z} = \begin{pmatrix} 7\cos\theta \\ 7\sin\theta \\ 0 \end{pmatrix} \mathrm{d}\theta\,\mathrm{d}z.$$

Meanwhile, the force at each point is given by

$$\mathbf{F}(\mathbf{r}(\theta, z)) = \begin{pmatrix} 7 \cdot 3\cos\theta \\ 7 \cdot 5\sin\theta \\ e^z \end{pmatrix}.$$

Thus, the dot product $\mathbf{F} \cdot \mathbf{n}$ is:

$$\begin{aligned} \mathbf{F} \cdot \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial z} \right) &= \begin{pmatrix} 7 \cdot 3\cos\theta \\ 7 \cdot 5\sin\theta \\ e^z \end{pmatrix} \cdot \begin{pmatrix} 7\cos\theta \\ 7\sin\theta \\ 0 \end{pmatrix} \\ &= 49(3\cos^2\theta + 5\sin^2\theta). \end{aligned}$$

Hence, the flux we seek is

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S &= \int_{\theta=0}^{2\pi} \int_{z=0}^{10} (21\cos^2\theta + 35\sin^2\theta) \cdot 7\,\mathrm{d}z\,\mathrm{d}\theta \\ &= \int_{\theta=0}^{2\pi} 490(3\cos^2\theta + 5\sin^2\theta)\,\mathrm{d}\theta \\ &= 490 \int_{\theta=0}^{2\pi} (3\cos^2\theta + 5\sin^2\theta)\,\mathrm{d}\theta. \end{aligned}$$

Recall that:

$$\int_{\theta=0}^{2\pi} \cos^2\theta\,\mathrm{d}\theta = \int_{\theta=0}^{2\pi} \sin^2\theta\,\mathrm{d}\theta = \pi$$

by using $\cos^2\theta = \frac{1+\cos(2\theta)}{2}$ and $\sin^2\theta = \frac{1-\cos(2\theta)}{2}$. Hence,

$$490 \int_{\theta=0}^{2\pi} (3\cos^2\theta + 5\sin^2\theta)\,\mathrm{d}\theta = 490 \cdot (3\pi + 8\pi) = \boxed{3920\pi}.$$ □

For the final example, we actually use the same hemisphere again, but this time we use spherical coordinates, so you can compare the methods. (In my opinion, this is uglier, but some people prefer spherical coordinates anyway.)

Sample Question

Consider the upper hemisphere of the sphere defined by $x^2 + y^2 + z^2 = 25$ with the unit normal vector oriented *downwards* towards the xy-plane. Calculate the flux of the vector field $\mathbf{F} = \begin{pmatrix} yz \\ xz \\ 0 \end{pmatrix}$ through this surface.

Solution. We parametrize with spherical coordinates by writing

$$\mathbf{r}(\varphi, \theta) = (5\sin\varphi\cos\theta, 5\sin\varphi\sin\theta, 5\cos\varphi)$$

for $0 \le \varphi \le \frac{\pi}{2}$ and $0 \le \theta \le 2\pi$. In that case, the cross product according to Table 23 is

$$\left(\frac{\partial \mathbf{r}}{\partial \varphi}\right) \times \left(\frac{\partial \mathbf{r}}{\partial \theta}\right) = 5\sin\varphi \cdot \mathbf{r}(\varphi, \theta).$$

This points away from the sphere since $\sin\varphi \ge 0$, so we flip the order:

$$\left(\frac{\partial \mathbf{r}}{\partial \theta}\right) \times \left(\frac{\partial \mathbf{r}}{\partial \varphi}\right) = -5\sin\varphi \cdot \mathbf{r}(\varphi, \theta).$$

Meanwhile, we have

$$\mathbf{F}(\mathbf{r}(\varphi, \theta)) = \begin{pmatrix} 25\sin\varphi\cos\varphi\sin\theta \\ 25\sin\varphi\cos\varphi\cos\theta \\ 0 \end{pmatrix}$$

If we expand the entire dot product we now get

$$\begin{aligned} \mathbf{F} \cdot \left(\left(\frac{\partial \mathbf{r}}{\partial \theta}\right) \times \left(\frac{\partial \mathbf{r}}{\partial \varphi}\right)\right) &= (25\sin\varphi\cos\varphi\sin\theta) \cdot (-5\sin\varphi) \cdot (5\sin\varphi\cos\theta) \\ &\quad +(25\sin\varphi\cos\varphi\cos\theta) \cdot (-5\sin\varphi) \cdot (5\sin\varphi\sin\theta) \\ &= -1250(\sin^3\varphi\cos\varphi\sin\theta\cos\theta). \end{aligned}$$

In other words, we have

$$\begin{aligned} \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS &= -1250 \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\frac{\pi}{2}} \sin^3\varphi\cos\theta\sin\theta\, d\varphi\, d\theta \\ &= -1250 \left(\int_{\theta=0}^{2\pi} \sin\theta\cos\theta\, d\theta\right) \left(\int_{\varphi=0}^{\frac{\pi}{2}} \sin^3\varphi\cos\varphi\, d\varphi.\right) \end{aligned}$$

The latter integral is super annoying to evaluate, but the former integral is zero because $\sin\theta\cos\theta = \frac{1}{2}\sin(2\theta)$, so we don't have to worry about the $d\varphi$ integral at all; we just get $\boxed{0}$ as the answer. □

§38.6 [TEXT] Another trick: writing as surface area if $\mathbf{F} \cdot \mathbf{n}$ is constant

We give one more trick for avoiding the cross product that only works in certain situations, but when it does, it makes your life a lot easier. Let $\mathcal{S}$ be a surface parametrized by $\mathbf{r} : \mathcal{R} \to \mathbb{R}^3$, and as always let $\mathbf{n}$ be shorthand for the unit vector in the direction of $\left(\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right)$.

Let's compare the flux and surface area in both longhand and shorthand.

- In longhand, we have

$$\mathrm{SurfArea}(\mathcal{S}) = \iint_{\mathcal{R}} \left|\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right| du\, dv$$

$$\mathrm{Flux} = \iint_{\mathcal{R}} \mathbf{F} \cdot \left(\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right) du\, dv = \iint_{\mathcal{R}} (\mathbf{F} \cdot \mathbf{n}) \left|\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right| du\, dv.$$

  (Keep type safety in mind here: the absolute value is a number, and the $\cdot$ is dot product of vectors in $\mathbb{R}^3$.) What we've done for the flux is decompose the cross product $\left(\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v}\right)$ into $\mathbf{n}$ times its magnitude, which we can do (in general, *any* vector $\mathbf{w}$ equals $|\mathbf{w}|$ multiplied by its direction unit vector). In this way you can make flux look a little more like surface area.

- In shorthand, it's even more obvious:

$$\text{Flux} = \iint_{\mathcal{S}} (\mathbf{F} \cdot \mathbf{n})\, \mathrm{d}S \quad \text{and} \quad \text{SurfArea}(\mathcal{S}) = \iint_{\mathcal{S}} \mathrm{d}S.$$

However, this resemblance is mostly useless, *except* in one really particular circumstance: the case where it happens $\mathbf{F} \cdot \mathbf{n}$ is always equal to the same constant c for every point on the surface. If you are that lucky, then the resemblance can actually be put to use:

$$\text{Flux} = \iint_{\mathcal{R}} c \cdot \left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v = c \iint_{\mathcal{R}} \left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v = c \cdot \text{SurfArea}(\mathcal{S}).$$

Then if you know the surface are of $\mathcal{S}$, you don't have to do *any* integration. You just multiply the surface area by c.

Again, this particular trick is extremely specific. It will only happen if $\mathbf{F}$ and $\mathcal{S}$ have been cherry-picked so that $\mathbf{F} \cdot \mathbf{n}$ is constant, and if you write down a "random" vector field $\mathbf{F}$ there is absolutely no chance this occurs by luck. However, despite the brittleness of the technique, this trick is still popular for some homework and exam questions because no calculation is needed. Here are two examples of this with spheres.

Sample Question

Let $\mathcal{S}$ denote the sphere $x^2 + y^2 + z^2 = 17^2 = 289$ of radius 17. Let $\mathbf{F} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Compute the flux

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}S.$$

(Orient $\mathbf{S}$ outwards.)

Solution. The normal vector $\mathbf{n}$ at any point (x, y, z) on the surface of the sphere is a unit vector pointing in the direction of $\langle x, y, z \rangle$. Conveniently, the force vector $\mathbf{F}$ is a vector of magnitude 17 in the same direction! That is,

$$\mathbf{F} \cdot \mathbf{n} = (17\mathbf{n}) \cdot (\mathbf{n}) = 17.$$

Consequently,

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}S = 17\, \text{SurfArea}(\mathcal{S}) = 17 \cdot (4 \cdot 289)\pi = \boxed{4 \cdot 17^3 \pi}.$$

(In general, we know a sphere of radius R has surface area $4R^2\pi$.) □

Sample Question

Let $\mathcal{S}$ denote the sphere $x^2 + y^2 + z^2 = 17^2 = 289$ of radius 17. Let $\mathbf{G}$ be the force of gravity exerted by a point mass m at the origin. Compute the flux

$$\iint_{\mathcal{S}} \mathbf{G} \cdot \mathbf{n} \, \mathrm{d}S.$$

(Orient $\mathbf{S}$ outwards.)

Solution. This is just like the previous example except that the gravity $\mathbf{G}$ exerted has magnitude $\frac{Gm}{17^2}$ and points in the *opposite* direction as $\mathbf{n}$. That is,

$$\mathbf{G} \cdot \mathbf{n} = \left(- \left(\frac{Gm}{17^2} \right) \mathbf{n} \right) \cdot (\mathbf{n}) = -\frac{Gm}{289}.$$

Consequently,

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S = -\frac{Gm}{289} \cdot \operatorname{SurfArea}(\mathcal{S}) = \frac{-Gm}{17^2} \cdot (4 \cdot 17^2 \pi) = \boxed{-4\pi Gm}.$$

(In general, we know a sphere of radius R has surface area $R^2\pi$.) □

Note that the answer is independent of the radius! The 17 cancels out.

§38.7 [EXER] Exercises

Exercise 38.1. Calculate the flux of the vector field

$$\mathbf{F}(x, y, z) = \left\langle \frac{x}{3}, \frac{y}{4}, \frac{1}{5} \right\rangle$$

across the portion of the surface defined by

$$x^3 + y^4 = e^z, \quad 0 \le x \le 5, \quad 0 \le y \le 5$$

where the normal vector is oriented upwards.

Chapter 39. Shortcut for flux: the divergence theorem

§39.1 [TEXT] The divergence theorem

Remember back when we had Green's theorem, we could transform 2D scalar flux (which was a *line integral*) into an area integral:

$$\underbrace{\oint_{\mathcal{C}} (\mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s)}_{= \oint_{\mathcal{C}} (-q \, \mathrm{d}x + p \, \mathrm{d}y)} = \underbrace{\iint_{\mathcal{R}} \nabla \cdot \mathbf{F} \, \mathrm{d}A}_{= \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A}.$$

This was nice because parametrization was annoying and straight-up area integrals are simpler. In general, if you still have the poster, the green pictures are easier to deal with.

The divergence theorem will let you do the same thing, transforming a flux surface integral (which is the horrendous *surface integral* that has been haunting you for the last couple weeks) into a volume integral. Which is an even bigger profit — no parametrization, no cross product table, etc.

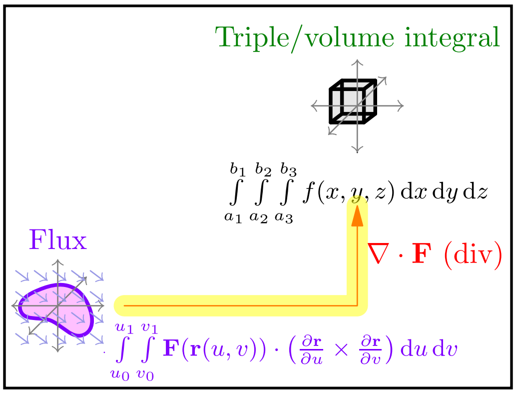

Figure 94: The highlighted arrow for the divergence theorem in our poster Figure 75.

Here's the result:

> **Memorize: Divergence theorem**
>
> Suppose a closed surface $\mathcal{S}$ encloses a compact solid $\mathcal{T}$, and $\mathbf{F}$ is defined everywhere in $\mathcal{T}$. Then
>
> $$\underbrace{\oiint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S}_{= \iint_{\mathcal{R}} \mathbf{F}(\mathbf{r}(u,v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \mathrm{d}u \, \mathrm{d}v} = \underbrace{\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V}_{= \iiint_{\mathcal{T}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z} \right) \mathrm{d}V}.$$

I've snuck in a new symbol $\oiint_{\mathcal{S}}$, but the extra circle is analogous to before. Just like how $\oint_{\mathcal{C}}$ was a reminder that $\mathcal{C}$ was a closed loop:

> **Definition of $\oiint$**
>
> $\oiint_{\mathcal{S}}$ means "$\iint_{\mathcal{S}}$ but with an extra optional reminder that $\mathcal{S}$ is a closed surface". (The reminder is optional, i.e. you are not obligated to add it even if $\mathcal{S}$ is closed.)

Also, note there's a fine-print requirement that $\mathcal{T}$ should be compact, i.e. it should not extend infinitely in any direction.

Remark: "Closed surface" = "holds water"

If you're unclear what "closed surface" means, a picture to keep in your head might be "holds water", i.e., you could imagine filling the interior of $\mathcal{S}$ with a water (that's the volume $\mathcal{T}$) and it shouldn't leak out. So the following are *not* closed surfaces:

- Curved part of hemisphere (e.g., bowl with no lid)
- Curved part of cylinder (e.g., straw)

But the following are closed surfaces:

- Cylinder including the two caps (e.g., water bottle)
- Sphere
- The six faces of a rectangular prism

We can jump straight into examples now!

Sample Question

Let $R > 0$ be given. Compute the flux of the vector field $\mathbf{F}(x, y, z) = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ through the closed surface of the sphere $\mathcal{S}$ defined by $x^2 + y^2 + z^2 = R^2$ oriented outward, using the Divergence Theorem.

Solution. The sphere $\mathcal{S}$ encloses a ball $\mathcal{T}$ of radius R. The divergence is given by

$$\nabla \cdot \mathbf{F} = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} + \frac{\partial r}{\partial z} = 1 + 1 + 1 = 3$$

Then the answer is

$$\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V = \iiint_{\mathcal{T}} 3 \, \mathrm{d}V = 3 \operatorname{Vol}(\mathcal{S}) = 3 \cdot \frac{4}{3} \pi R^3 = \boxed{4\pi R^3}. \qquad \square$$

Remark: Connection to surface area of sphere

This was also the first example we did with the surface area trick, where we found that the answer was $R \cdot \operatorname{SurfArea}(\mathcal{S})$ which is also $4\pi R^3$.

Actually, put another way: if you know the volume of the sphere is $\frac{4}{3}\pi R^3$ and the divergence theorem, then the surface area trick lets you derive the surface area formula of $4\pi R^2$.

Warning: Beware of undefined points of $\mathbf{F}$

You need to be careful to only apply the divergence theorem if the force is actually defined on the entire solid $\mathcal{T}$! Here's an example of what can go wrong.

Let $\mathcal{S}$ denote the sphere $x^2 + y^2 + z^2 = R^2$ of radius R again. Let $\mathbf{G}$ be the force of gravity exerted by a point mass m at the origin. In the last chapter we computed

$$\iint_{\mathcal{S}} \mathbf{G} \cdot \mathbf{n} \, \mathrm{d}S = -4\pi Gm$$

using the surface area trick.

However, if you compute the divergence $\nabla \cdot \mathbf{G}$, you'll actually find it's *zero* at every point — except the origin, where $\mathbf{G}$ is undefined because the gravity causes division-by-zero. (See Exercise 32.1.) If you blindly apply the divergence theorem and don't notice the issue with the origin, you would instead get the wrong answer $\iiint_{\mathcal{T}} 0 \, \mathrm{d}V = 0$, rather than the correct answer $-4\pi Gm$. (That said, see Exercise 39.2 for a safe usage.)

Sample Question

Let $a > 0$ be given. Calculate the flux of the vector field $\mathbf{F}(x, y, z) = \begin{pmatrix} x^2 \\ y^2 \\ z^2 \end{pmatrix}$ through the closed surface of the cube $\mathcal{S}$ bounded by $0 \le x, y, z \le a$ using the Divergence Theorem.

Solution. The divergence is

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial y}(y^2) + \frac{\partial}{\partial z}(z^2) = 2x + 2y + 2z.$$

Hence the flux turns into

$$\iiint_{\mathcal{T}} (2x + 2y + 2z) \, \mathrm{d}V = 2 \iiint_{\mathcal{T}} x \, \mathrm{d}V + 2 \iiint_{\mathcal{T}} y \, \mathrm{d}V + 2 \iiint_{\mathcal{T}} z \, \mathrm{d}V.$$

Due to the symmetry of the cube:

$$\iiint_{\mathcal{T}} x \, \mathrm{d}V = \iiint_{\mathcal{T}} y \, \mathrm{d}V = \iiint_{\mathcal{T}} z \, \mathrm{d}V = \frac{a^3}{2}.$$

If you can't see it by symmetry, you could also just explicitly calculate

$$\iiint_{\mathcal{T}} x \, \mathrm{d}V = \left(\int_{x=0}^{a} x \, \mathrm{d}x \right) \left(\int_{y=0}^{a} \mathrm{d}y \right) \left(\int_{z=0}^{a} \mathrm{d}z \right) = \frac{a}{2} \cdot a \cdot a = \frac{a^3}{2}.$$

In any case, we get an answer of

$$2 \cdot \frac{a^3}{2} + 2 \cdot \frac{a^3}{2} + 2 \cdot \frac{a^3}{2} = \boxed{3a^3}.$$

□

Sample Question

Compute the flux of the vector field $\mathbf{F}(x, y, z) = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix}$ through the closed surface $\mathcal{S}$ defined by $x^4 + (y-5)^6 + z^8 = 2025$.

Solution. The surface $\mathcal{S}$ is hard to describe, but it encloses *some* solid $\mathcal{T}$. However, if you compute the divergence, it is

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(yz) + \frac{\partial}{\partial y}(xz) + \frac{\partial}{\partial z}(xy) = 0 + 0 + 0 = 0.$$

So it doesn't even matter what the solid $\mathcal{T}$ is; the answer is just

$$\iiint_{\mathcal{T}} 0 \, \mathrm{d}V = \boxed{0}.$$

□

Sample Question

Compute the flux of the vector field $\mathbf{F}(x, y, z) = \begin{pmatrix} xy \\ yz \\ zx \end{pmatrix}$ through the closed surface $\mathcal{S}$ formed by the paraboloid $z = x^2 + y^2$ and its circular base $z = 0$, where $x^2 + y^2 \le 1$.

Solution. Let $\mathcal{T}$ denote the region enclosed by $\mathcal{S}$. The divergence is given by

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(yz) + \frac{\partial}{\partial z}(zx) = y + z + x$$

The region $\mathcal{T}$ is bounded by $z = x^2 + y^2$ and $z = 0$, within $x^2 + y^2 \le 1$. So the divergence theorem means we need to calculate

$$\iint_{x^2+y^2 \le 1} \int_{z=0}^{x^2+y^2} (x + y + z) \, \mathrm{d}z \, \mathrm{d}x \, \mathrm{d}y.$$

Naturally, this is best done using cylindrical coordinates. Writing $x = r\cos\theta$ and $y = r\sin\theta$, and remembering that

$$\mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \mathrm{d}V = r \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}z$$

then this becomes

$$\int_{r=0}^{1} \int_{\theta=0}^{2\pi} \int_{z=0}^{r^2} r(r\cos\theta + r\sin\theta + z) \, \mathrm{d}z \, \mathrm{d}\theta \, \mathrm{d}r.$$

But the integrals with θ in them are going to be zero by symmetry. For example, the first term is

$$\int_{r=0}^{1} \int_{\theta=0}^{2\pi} \int_{z=0}^{r^2} r^2 \cos\theta \, \mathrm{d}z \, \mathrm{d}r \, \mathrm{d}\theta = \left(\int_{r=0}^{1} \int_{z=0}^{r^2} r^2 \, \mathrm{d}z \, \mathrm{d}r \right) \underbrace{\left(\int_{\theta=0}^{2\pi} \cos\theta \, \mathrm{d}\theta \right)}_{=0} = 0.$$

Similarly, the contribution of $r\sin\theta$ is just zero as well. So we are just left with

$$\int_{r=0}^{1}\int_{\theta=0}^{2\pi}\int_{z=0}^{r^2} rz \,\mathrm{d}z\,\mathrm{d}\theta\,\mathrm{d}r = \left(\int_{r=0}^{1}\int_{z=0}^{r^2} rz\,\mathrm{d}z\,\mathrm{d}r\right)\left(\int_{\theta=0}^{2\pi}\mathrm{d}\theta\right).$$

Obviously $\int_{\theta=0}^{2\pi}\mathrm{d}\theta = 2\pi$. The double integral can be evaluated as

$$\begin{aligned}\int_{r=0}^{1} r\int_{z=0}^{r^2} z\,\mathrm{d}z\,\mathrm{d}r &= \int_{r=0}^{1} r\cdot\left[\frac{z^2}{2}\right]_{z=0}^{r^2}\mathrm{d}r \\ &= \int_{r=0}^{1}\frac{r^5}{2}\,\mathrm{d}r \\ &= \left[\frac{r^6}{12}\right]_{r=0}^{1} = \frac{1}{12}.\end{aligned}$$

Hence the final answer is

$$\frac{1}{12}\cdot 2\pi = \boxed{\frac{\pi}{6}}.$$

□

§39.2 [SIDENOTE] A picture for why the divergence theorem is true

The picture is actually exactly the same as Figure 90 from Section 35.4, our picture of Green's theorem for 2D flux! The divergence is still drawn as green explosions. The only change is in the dimensions:

- For Green's theorem for flux, we have a 1D path (purple square) enclosing a 2D region broken up into little grey squares.
- For divergence theorem, we have a 2D surface (purple box) enclosing a 3D regions broken up into little grey cubes.

And the rest of the analogy carries over: all the interior green arrows cancel except for those pushing directly against the purple faces of the cube, so there's the flux integral we wanted.

§39.3 [RECAP] All the methods for flux

Here's a complete recipe for flux, augmented with the two shortcuts we described.

Recipe for flux, with shortcuts

Suppose we need to calculate the flux of $\mathbf{F}$ through a surface $\mathcal{S}$.

1. If $\mathcal{S}$ is a closed region, use the divergence theorem to avoid parametrization:
$$\oiint_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S = \underbrace{\iiint_{\mathcal{T}} \nabla\cdot\mathbf{F}\,\mathrm{d}V}_{=\iiint_{\mathcal{T}}\left(\frac{\partial p}{\partial x}+\frac{\partial q}{\partial y}+\frac{\partial r}{\partial z}\right)\mathrm{d}V}.$$
2. If $\mathbf{F}\cdot\mathbf{n}$ happens to equal the same constant c everywhere (as described in Section 38.6), then output c times the surface area of $\mathcal{S}$, i.e.
$$\oiint_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S = c\,\mathrm{SurfArea}(\mathcal{S}).$$
3. Otherwise, fall back to the parametrization recipe described in Section 38.5. To describe it again here briefly:
   1. Get the cross product $\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}$ by either looking it up from Table 23 or by computing it by hand, for a parametrization $\mathbf{r}:\mathcal{R}\to\mathbb{R}^3$ of the surface $\mathcal{S}$.
   2. If necessary, negate the cross product to match the orientation of the surface specified in the question.
   3. Compute the dot product $\mathbf{F}\cdot\left(\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right)$.
   4. Integrate over the region $\mathcal{R}$ using any method for double integrals.

§39.4 [TEXT] Advanced technique: sealing regions

This is the 3D analog of Section 34.6: in some situations if you have a surface $\mathcal{S}$ which isn't closed, you can seal it by adding some part to the surface. The picture you can have in your head is that you have a bowl or something, and then you add a layer of plastic wrap on the bowl.

Sample Question

Let $\mathbf{F}$ be the vector field defined by:
$$\mathbf{F}(x,y,z) = \langle x+\tan z, y+e^z, 1\rangle.$$
Consider the hemisphere $\mathcal{S}$ defined by the equation:
$$x^2+y^2+z^2=1 \quad \text{with} \quad z\ge 0$$
oriented outward. Compute the flux of $\mathbf{F}$ through $\mathcal{S}$.

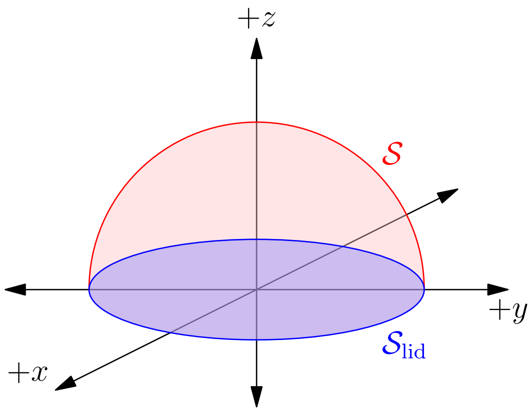

Figure 95: Sealing a bowl with a lid. Like microwaving food, though the bowl is upside-down.

Solution. Our picture is that $\mathcal{S}$ looks like an upside-down bowl. So we add a lid $\mathcal{S}_{\text{lid}}$ consisting of the disk $z = 0$ and $x^2 + y^2 \le 1$. This encloses a solid region $\mathcal{T}$, half a solid ball of radius 1, as in Figure 95.

The divergence of $\mathbf{F}$ is:

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(x + \tan z) + \frac{\partial}{\partial y}(y + e^z) + \frac{\partial}{\partial z}(1) = 1 + 1 + 0 = 2$$

which is constant. So the integral of the divergence over $\mathcal{T}$ is just

$$\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V = 2 \cdot \operatorname{Vol}(\mathcal{T}) = 2 \cdot \left(\frac{1}{2} \cdot \frac{4}{3}\pi \cdot 1^3\right) = \frac{4}{3}\pi.$$

Meanwhile, $\mathcal{S}_{\text{lid}}$ (which we orient downwards) is a flat surface, so its flux integral is easy to calculate: from Table 23 we choose $\mathbf{n} \, \mathrm{d}S = \langle 0, 0, -1 \rangle$ and hence

$$\iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S = \iint_{x^2+y^2 \le 1} \langle x + \tan 0, y + e^0, 1 \rangle \cdot \langle 0, 0, -1 \rangle \, \mathrm{d}x \, \mathrm{d}y = \iint_{x^2+y^2 \le 1} (-1) \, \mathrm{d}x \, \mathrm{d}y = -\pi.$$

So when we apply the divergence theorem, we get that

$$\underbrace{\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S}_{\text{Answer}} + \underbrace{\iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S}_{= -\pi} = \underbrace{\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V}_{= \frac{4}{3}\pi}.$$

Hence, we get the answer

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S = \frac{4}{3}\pi - (-\pi) = \boxed{\frac{7}{3}\pi}.$$

□

§39.5 [EXER] Exercises

Exercise 39.1. Let $\mathcal{S}$ be the part of the surface $z = e^{x^2+y^2}$ where $z \le e$, with normal vector oriented downwards. Let $\mathbf{F}(x, y, z) = \langle \cos(z^2) - x, \sin(z^2) - y, 2z \rangle$. Compute the flux of $\mathbf{F}$ through $\mathcal{S}$. (Recommended approach: sealing.)

Exercise 39.2. Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are two closed surfaces that don't intersect and such that $\mathcal{S}_2$ is contained inside $\mathcal{S}_1$. Orient both surfaces outwards. Let O be a point contained inside $\mathcal{S}_2$. Consider the force of gravity $\mathbf{G}$ exerted by a point mass of mass m at O. Show that

$$\oiint_{\mathcal{S}_1} \mathbf{G} \cdot \mathbf{n} \, \mathrm{d}S = \oiint_{\mathcal{S}_2} \mathbf{G} \cdot \mathbf{n} \, \mathrm{d}S.$$

Exercise 39.3 (*). Prove Green's theorem for flux by quoting the divergence theorem.

That is, suppose $\mathbf{F} = \left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right)$ is a vector field in $\mathbb{R}^2$ and $\mathcal{C}$ is a closed loop enclosing a region $\mathcal{R}$ counterclockwise. Find a way to use the divergence theorem to prove

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}\mathbf{s} = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A.$$

Chapter 40. Classical Stokes' Theorem

This topic was excluded from the final exam for 18.02 in Fall 2024, so this chapter is just a brief summary for interest.

All the red arrows in Figure 75 are special cases of what *mathematicians* refer to as generalized Stokes' theorem. Despite this, we confusingly call the final red arrow "Stokes' theorem" as well. I'll use the phrase "classical" to indicate this came first historically.

§40.1 [TEXT] The classical Stokes' theorem

Here's the statement of the result:

Definition of compatible orientations

Suppose $\mathcal{C}$ is a closed loop in $\mathbb{R}^3$ which is the boundary of an oriented surface $\mathcal{S}$. The orientation of $\mathcal{C}$ and $\mathcal{S}$ are *compatible* if, when walking along $\mathcal{C}$ in the chosen direction, with $\mathcal{S}$ to the left, the normal vector $\mathbf{n}$ is pointing up.

Classical Stokes' theorem

Let $\mathcal{C}$ be a closed loop in $\mathbb{R}^3$ parametrized by $\mathbf{r}_1(t)$. Suppose $\mathcal{S}$ is the boundary of an oriented surface $\mathcal{S}$ parametrized by $\mathbf{r}_2(u, v)$. Assume the orientation of $\mathcal{C}$ and $\mathcal{S}$ are compatible. Then

$$\underbrace{\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}_1}_{= \int_{t=\text{start}}^{\text{stop}} \mathbf{F} \cdot \mathbf{r}_1'(t)\,\mathrm{d}t} = \underbrace{\iint_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\,\mathrm{d}S.}_{= \iint_{u,v} (\operatorname{curl} \mathbf{F}) \cdot \left(\frac{\partial \mathbf{r}_2}{\partial u} \times \frac{\partial \mathbf{r}_2}{\partial v}\right) \mathrm{d}u\,\mathrm{d}v}$$

For a picture of why this theorem is true, one should just refer to Figure 90 again (from Section 35.4), which we used when explaining Green's theorem for flux. The classical Stokes theorem is the same picture, except that one should imagine the grid is superimposed onto the surface $\mathcal{S}$ (rather than laying flat in 2D). The same explanation should then carry over verbatim.

§40.2 [TEXT] Reasons to not be stoked about classical Stokes' theorem

Unlike the other big-name theorems we've seen (FTC, Green, and divergence theorem), the classical Stokes' theorem does not make for good exam questions, for a few reasons:

- Both sides require parametrization, so it's not as slick as FTC, Green, or divergence theorem, which were powerful because they let you skip the parametrization step.
- Surface integrals are more painful than line integrals, but there's no "anti-curl" procedure analogous to anti-gradient, so it doesn't help with surface integrals of a "random" vector field.

In other words, it doesn't provide a nice shortcut like the other theorems do.

Digression

A really clever student might imagine that maybe there's a situation where you have a line integral over a closed loop $\mathcal{C}$, you use classical Stokes theorem to change it to a surface integral, and then you use the divergence theorem to convert it to a volume integral, so maybe classical Stokes' theorem is good for something after all? But that kite won't fly: this could only work if $\mathcal{S}$ is a closed surface, but if $\mathcal{S}$ is a closed surface it's impossible for it to have a (nonempty) boundary $\mathcal{C}$. Hence this situation will never apply.

Similarly, imagine you have a surface integral of a vector field $\mathbf{F}$. Maybe you can try to find an anti-curl for it (i.e. a vector field $\mathbf{F}'$ such that $\nabla \times \mathbf{F}' = \mathbf{F}$), and then if $\mathbf{F}'$ is conservative hope you can use FTC? But that kite won't fly either: the curl of a conservative vector field is always $\mathbf{0}$, so this would only work if $\mathbf{F}$ was the zero vector field to begin with, and in that case you certainly don't need any help integrating it.

§40.3 [EXER] Exercises

Exercise 40.1. Prove Green's theorem for work by quoting classical Stokes' theorem.

That is, suppose $\mathbf{F} = \binom{p}{q}$ is a vector field in $\mathbb{R}^2$ and $\mathcal{C}$ is a closed loop enclosing a region $\mathcal{R}$ counterclockwise. Find a way to use classical Stokes' theorem to prove

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A.$$

Part Kilo: Practice questions

Chapter 41. Practice questions

Some practice questions from topics through the entire course. Solutions are in Chapter 52.

§41.1 Practice half-final

This was a cumulative practice exam given on December 18, 2024 covering the whole course. It was about half the length of the final exam (which was 14 questions long).

Exercise 41.1. Give an example of a complex number z whose real and imaginary part are both negative such that $z^3 = -1000i$. Write your answer in rectangular form.

Exercise 41.2. Compute the unique real number a for which the matrix $M = \begin{pmatrix} 1 & 1 \\ a & 6 \end{pmatrix}$ has an eigenvalue of 2. For this value of a, compute the other eigenvalue of M, and a (nonzero) eigenvector for that eigenvalue.

Exercise 41.3. The four points $(b, 0, 0)$, $(0, b, 0)$, $(0, 0, b)$, and $(2, 3, 6)$ lie on a plane $\mathcal{P}$. Compute b, and compute the distance from $(1, 2, 3)$ to $\mathcal{P}$.

Exercise 41.4. Let $f(x, y) = \cos(x) + \sin(y)$. Give an example of a saddle point of f, and an example of a local maximum of f. Pick either of these two points and sketch the level curve of f passing through it.

Exercise 41.5. Compute the maximum and minimum value of $x^2 + 2y^2 + 4x$ over the region $x^2 + y^2 \leq 9$.

Exercise 41.6. Use any method (recommended approach: change order of integration) to compute

$$\int_{x=0}^{1} \int_{y=x}^{1} \int_{z=y}^{1} e^{z^3} \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}x.$$

Exercise 41.7. Compute the real number c for which

$$\mathbf{F}(x, y, z) = \langle 7\cos(x), \cos(y)\cos(2z), c\sin(y)\sin(2z) \rangle$$

is conservative. For that c, compute the maximum possible value of a line integral $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ across all possible choices of some curve $\mathcal{C}$ in $\mathbb{R}^3$.

Exercise 41.8. Let

$$\mathbf{F}(x, y, z) = \langle x + e^y + z^3, e^x + y + z^3, z \rangle.$$

Let $\mathcal{S}$ be the surface defined by $x^2 + y^2 = 100$ and $7 \leq z \leq 9$, with normal vector oriented outwards (thus $\mathcal{S}$ is the curved part of a cylinder). Compute the divergence of $\mathbf{F}$. Then compute the flux of $\mathbf{F}$ through $\mathcal{S}$. (Recommended approach: add two "lids" to $\mathcal{S}$, calculate flux through the lids by hand, then use the divergence theorem.)

§41.2 Miscellaneous practice questions without solutions

Exercise 41.9. Let $\mathbf{v} = \begin{pmatrix} 8 \\ 9 \\ 10 \end{pmatrix}$. Suppose that

$$\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \begin{pmatrix} 3 \\ 5 \\ t \end{pmatrix}$$

for some real number t. Compute t, and compute all possibilities for the vector $\mathbf{w}$.

Exercise 41.10. Compute the unique 2×2 matrix A for which $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 3 and $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 4.

Exercise 41.11. Let $A = \begin{pmatrix} 4 & 3 \\ 4 & 8 \end{pmatrix}$.

- Compute the eigenvalues and eigenvectors for A.
- Compute the eigenvalues and eigenvectors for A^2.
- Compute the eigenvalues and eigenvectors for A^{100}.

Exercise 41.12. Let $\mathcal{S}$ be the level surface of $f(x,y,z) = e^x + e^{2y} + e^{3z}$ that passes through the origin. Compute all real numbers t such that the vector $\langle t+4, t+5, t+6 \rangle$ is tangent to $\mathcal{S}$ at the origin.

Exercise 41.13. Let $f(x,y,z)$ be a differentiable function, and let $g(x,y,z) = e^{f(x,y,z)}$. Let P be any point in $\mathbb{R}^3$. Suppose $f(P) = 2$, and $\nabla f(P)$ is a unit vector. Compute the magnitude of $\nabla g(P)$.

Exercise 41.14. Show that

$$f(x,y) = (x+y)^{100} - (x-y)^{100}$$

has exactly one critical point, and that critical point is a saddle point.

Exercise 41.15. Let $\mathcal{R}$ denote the region in the xy-plane cut out by $y = x+2$, $y = x + 20$, and $y = x^2$. Compute the area of $\mathcal{R}$.

Exercise 41.16. Compute

$$\int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} e^{-x^2-y^2} \, dy \, dx$$

by changing to polar coordinates. Then determine the value of $\int_{x=-\infty}^{+\infty} e^{-x^2} \, dx$.

Part Lima: Solutions

Chapter 42. Solutions to Part Alfa

§42.1 Solution to Exercise 2.1 (type safety)

Exercise 2.1. Let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ be vectors in $\mathbb{R}^3$. By using Table 1 (or skimming Section 4.1 briefly), determine whether each of the following expressions is a real number, a vector, or nonsense (type-error); there should be one of each.

- $(\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w}$
- $\mathbf{u} \cdot \mathbf{v} + \mathbf{w}$ (here order of operations is $\cdot$ before $+$)
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$

(The symbol $\cdot$ confusingly can refer to three different things: grade-school multiplication, scalar multiplication, or the dot product.)

- The expression $(\mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{w}$ is a **vector**, since
$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \text{scalar} \\ \text{scalar} \cdot \mathbf{w} &= \text{vector} \end{aligned}$$
- The expression $\mathbf{u} \cdot \mathbf{v} + \mathbf{w}$ is a **type-error** since
$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \text{scalar} \\ \text{scalar} + \mathbf{w} &= \text{undefined (cannot add scalar and vector)}. \end{aligned}$$
- The expression $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w})$ is a **real number** since
$$\begin{aligned} \mathbf{v} + \mathbf{w} &= \text{vector} \\ \mathbf{u} \cdot \text{vector} &= \text{scalar} \end{aligned}$$

§42.2 Solution to Exercise 3.1 (direction)

Exercise 3.1. Compute the unit vector along the direction of the vector
$$\begin{pmatrix} -0.0008\pi \\ -0.0009\pi \\ -0.0012\pi \end{pmatrix}.$$

The point of this example is to emphasize that you can scale out weird positive constants like 0.0001π; the vector
$$\mathbf{w} = \begin{pmatrix} -8 \\ -9 \\ -12 \end{pmatrix}$$
points in the same direction. So it's enough to find the unit vector in the direction of $\mathbf{w}$ which is

$$\frac{1}{|\mathbf{w}|}\mathbf{w} = \frac{1}{\sqrt{(-8)^2+(-9)^2+(-12)^2}}\begin{pmatrix}-8\\-9\\-12\end{pmatrix} = \frac{1}{\sqrt{64+81+144}}\begin{pmatrix}-8\\-9\\-12\end{pmatrix} = \frac{1}{\sqrt{289}}\begin{pmatrix}-8\\-9\\-12\end{pmatrix}$$
$$= \frac{1}{17}\begin{pmatrix}-8\\-9\\-12\end{pmatrix} = \boxed{\begin{pmatrix}-8/17\\-9/17\\-12/17\end{pmatrix}}.$$

(Note that $\langle 8/17, 9/17, 12/17\rangle$ is not a correct answer: that vector points in the opposite direction.)

§42.3 Solution to Exercise 3.2 ($\det(10A)$)

Exercise 3.2. If A is a 3×3 matrix with determinant 2, what values could $\det(10A)$ take?

I claim the answer is

$$\det(10A) = 10^3 \cdot \det A = \boxed{2000}.$$

Here are two ways to see this:

- To see it geometrically, consider the parallelepiped formed by the column vectors of A. If we scale each of its side lengths by 10, then the volume should increase by a factor of $10^3 = 1000$.
- To see it algebraically, in the formula for the determinant the point is that every term scales up by a factor of 10, and the products are three at a time.

  This might be easier to see from an example, so let's take

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

  as an example of a matrix with determinant 2:

  $$\det(A) = 3\cdot 9\cdot 1 - 5\cdot 5\cdot 1 = 2.$$

  Then

  $$10A = \begin{pmatrix}30&50&0\\50&90&0\\0&0&10\end{pmatrix}$$

  so

  $$\det(A) = 30\cdot 90\cdot 10 - 50\cdot 50\cdot 10 = 2000.$$

§42.4 Solution to Exercise 3.3 (coplanar)

Exercise 3.3. Compute the real number a for which the points $(0,0,0)$, $(1,0,1)$, $(0,1,2)$ and $(1,2,a)$ all lie on one plane.

Call the points $P_1 = (0,0,0)$, $P_2 = (1,0,1)$, $P_3 = (0,1,2)$, $P_4 = (1,2,a)$.

There are several approaches to this (including ones that use later material); the one using the material in this chapter is the following:

Idea

Four points are coplanar if the volume of the parallelepiped formed by the vectors connecting one point to the other three is zero. This condition is equivalent to the determinant of the matrix formed by these three vectors being zero.

Choose $P_1 = (0,0,0)$ as the reference point. Then, the vectors from P_1 to the other points are:

$$\mathbf{v}_1 = P_2 - P_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\mathbf{v}_2 = P_3 - P_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

$$\mathbf{v}_3 = P_4 - P_1 = \begin{pmatrix} 1 \\ 2 \\ a \end{pmatrix}$$

Then construct a 3×3 matrix using these vectors as columns:

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & a \end{pmatrix}.$$

The determinant of M is:

$$\det(M) = 1 \cdot \begin{vmatrix} 1 & 2 \\ 2 & a \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 2 \\ 1 & a \end{vmatrix} + 1 \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix}.$$

Compute each minor determinant:

$$\begin{vmatrix} 1 & 2 \\ 2 & a \end{vmatrix} = (1)(a) - (2)(2) = a - 4$$

$$\begin{vmatrix} 0 & 2 \\ 1 & a \end{vmatrix} = (0)(a) - (2)(1) = -2$$

$$\begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} = (0)(2) - (1)(1) = -1.$$

Substituting back:

$$\det M = 1 \cdot (a-4) - 0 \cdot (-2) + 1 \cdot (-1) = a - 4 - 1 = a - 5.$$

Hence $\det M = 0 \iff \boxed{a = 5}$.

§42.5 Solution to Exercise 4.1 (4d vectors)

Exercise 4.1. In four-dimensional space $\mathbb{R}^4$, the vectors $\langle 1,2,3,4 \rangle$ and $\langle 5,6,7,t \rangle$ are perpendicular. Compute t.

We need the dot product to be zero:

$$\begin{aligned} 0 &= \langle 1, 2, 3, 4 \rangle \cdot \langle 5, 6, 7, t \rangle \\ &= 1 \cdot 5 + 2 \cdot 6 + 3 \cdot 7 + 4 \cdot t \\ &= 38 + 4t \implies \boxed{t = -\frac{19}{2}}. \end{aligned}$$

§42.6 Solution to Exercise 4.2 (projection)

Exercise 4.2.
- Compute the vector projection of $\langle 123, 456, 789 \rangle$ in the direction of $\mathbf{e}_1$.
- Compute the scalar component and vector projection of $\mathbf{v} = \langle 1, 2, 3 \rangle$ along the direction of $\mathbf{w} = \langle -3000, -4000, 0 \rangle$.

The first part asks to compute the vector projection of $\langle 123, 456, 789 \rangle$ in the direction of $\mathbf{e}_1$. The answer is just

$$\boxed{\langle 123, 0, 0 \rangle}.$$

You could get this using the recipe if you wanted, but if you draw a picture the point is you're just projecting the vector $\langle 123, 456, 789 \rangle$ to the x-axis, which gives you its x-component.

For the second part, let

$$\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$
$$\mathbf{w} = \begin{pmatrix} -3000 \\ -4000 \\ 0 \end{pmatrix}.$$

Note that the factor of -1000 in $\mathbf{w}$ doesn't matter, since scaling $\mathbf{w}$ doesn't matter. We'll keep the -1000 around just for illustration reasons, but in practice an experienced student would just use $\mathbf{w} = \langle 3, 4, 0 \rangle$ instead.

We just follow the recipe in Section 4.4 directly. We first compute the dot product:

$$\mathbf{v} \cdot \mathbf{w} = (1)(-3000) + (2)(-4000) + (3)(0) = -3000 - 8000 + 0 = -11000$$

The magnitude is

$$|\mathbf{w}| = \sqrt{(-3000)^2 + (-4000)^2 + 0^2} = 5000.$$

The scalar component of $\mathbf{v}$ along $\mathbf{w}$ is given by:

$$\operatorname{comp}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{w}|} = \frac{-11000}{5000} = -\frac{11}{5}.$$

The vector projection of $\mathbf{v}$ along $\mathbf{w}$ is given by:

$$\begin{aligned}\operatorname{proj}_{\mathbf{w}}(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{w}}{|\mathbf{w}|}\left(\frac{\mathbf{w}}{|\mathbf{w}|}\right) &= -\frac{11}{5}\cdot\frac{\begin{pmatrix}-3000\\-4000\\0\end{pmatrix}}{5000} \\ &= -\frac{11}{5}\cdot\begin{pmatrix}-3/5\\-4/5\\0\end{pmatrix} \\ &= \boxed{\begin{pmatrix}33/25\\44/25\\0\end{pmatrix}}.\end{aligned}$$

§42.7 Solution to Exercise 4.3 (dot product 3)

Exercise 4.3. Let $\mathbf{w} = \langle 3, 4\rangle$. Compute all unit vectors $\mathbf{v}$ in $\mathbb{R}^2$ for which $\mathbf{v}\cdot\mathbf{w} = 3$.

Geometrically, we expect there to be two solutions: if θ is the angle between the two vectors, we need $\cos\theta = \frac{3}{5}$, and so there should be two vectors that work. See Figure 96 for a picture. (You might already guess one of the solutions — $\mathbf{w} = \langle 1, 0\rangle$ obviously works — but we'll pretend we didn't notice that.)

Translating the givens algebraically, we have the following system of equations:

$$\begin{aligned}\mathbf{v}\cdot\mathbf{w} = 3 &\implies 3x + 4y = 3 \\ |\mathbf{v}| = 1 &\implies x^2 + y^2 = 1\end{aligned}$$

From the dot product condition, solve for x:

$$3x + 4y = 3 \implies x = \frac{3-4y}{3}.$$

Substitute $x = \frac{3-4y}{3}$ into $x^2 + y^2 = 1$ and solve:

$$\begin{aligned}\left(\frac{3-4y}{3}\right)^2 + y^2 &= 1 \\ \iff \frac{25y^2 - 24y + 9}{9} &= 1 \\ \iff y(25y - 24) &= 0.\end{aligned}$$

Hence either $y = 0$ or $y = \frac{24}{25}$.

- If $y = 0$ we get $x = \frac{3-4(0)}{3} = 1$. Thus, the first unit vector is: $\mathbf{v}_1 = \langle 1, 0\rangle$.
- If $y = \frac{24}{25}$ we get $x = \frac{3-4\left(\frac{24}{25}\right)}{3} = -\frac{7}{25}$. Thus, the second unit vector is: $\mathbf{v}_2 = \langle -\frac{7}{25}, \frac{24}{25}\rangle$.

In conclusion the answer is

$$\boxed{\langle 1, 0\rangle \quad \text{and} \quad \left\langle -\frac{7}{25}, \frac{24}{25}\right\rangle}.$$

See Figure 96 for a picture of the two answers.

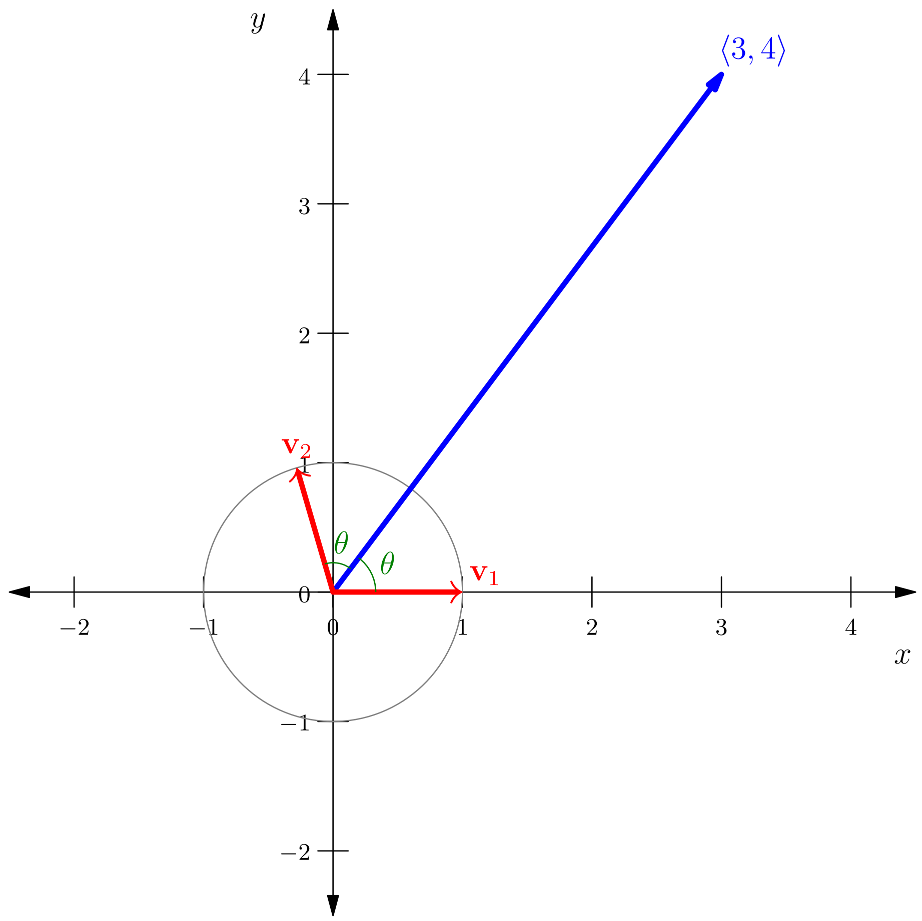

Figure 96: The two answers to Exercise 4.3, which are two unit vectors spaced at an equal angle away from $\langle 3, 4\rangle$.

§42.8 Solution to Exercise 4.4 (2 and 5)

Exercise 4.4 (*). Determine all possible values of $ax + by + cz$ over real numbers a, b, c, x, y, z satisfying $a^2 + b^2 + c^2 = 2$ and $x^2 + y^2 + z^2 = 5$.

Construct vectors $\mathbf{v} = \langle a, b, c\rangle$ and $\mathbf{w} = \langle x, y, z\rangle$ in $\mathbb{R}^3$. Then the problem statement is saying that $|\mathbf{v}| = \sqrt{2}$, $|\mathbf{w}| = \sqrt{5}$, and asks for all possible values of $\mathbf{v} \cdot \mathbf{w}$. But the geometric definition of the dot product says that

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\ |\mathbf{w}| \cos\theta = \sqrt{10}\cos\theta$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

Since θ could be any angle at all (as $\mathbf{v}$ and $\mathbf{w}$ are arbitrary $\mathbb{R}^3$ vectors), the possible values of $\sqrt{10}\cos\theta$ will range from in the interval $\boxed{\left[-\sqrt{10}, \sqrt{10}\right]}$ as $\cos\theta$ ranges from -1 to 1.

§42.9 Solution to Exercise 5.1 (faces of a cube)

Exercise 5.1. A cube is drawn somewhere in $\mathbb{R}^3$ (its faces are not parallel to the coordinate axes). Two of the faces of the cube are contained in the planes $x + 2y + 3z = 4$ and $5x + 6y + kz = 7$, respectively, for some real number k. Given this information, compute k.

The main observation is this:

> **Idea**
>
> The faces of the cube have orthogonal normal vectors.

And the normal vectors to the two planes are:

$$\mathbf{n}_1 = \langle 1, 2, 3 \rangle \mathbf{n}_2 = \langle 5, 6, k \rangle.$$

For the planes to be perpendicular, their normal vectors must satisfy:

$$\begin{aligned} 0 &= \mathbf{n}_1 \cdot \mathbf{n}_2 \\ &= (1)(5) + (2)(6) + (3)(k) = 5 + 12 + 3k = 17 + 3k. \end{aligned}$$

Solving for k gives $\boxed{k = -\frac{17}{3}}$.

§42.10 Solution to Exercise 5.2 (distance to two planes)

Exercise 5.2. The distance from a certain point P to the plane $3x + 4y + 12z = -1$ is 42. What are the possible distances from P to the plane $3x + 4y + 12z = 1000$?

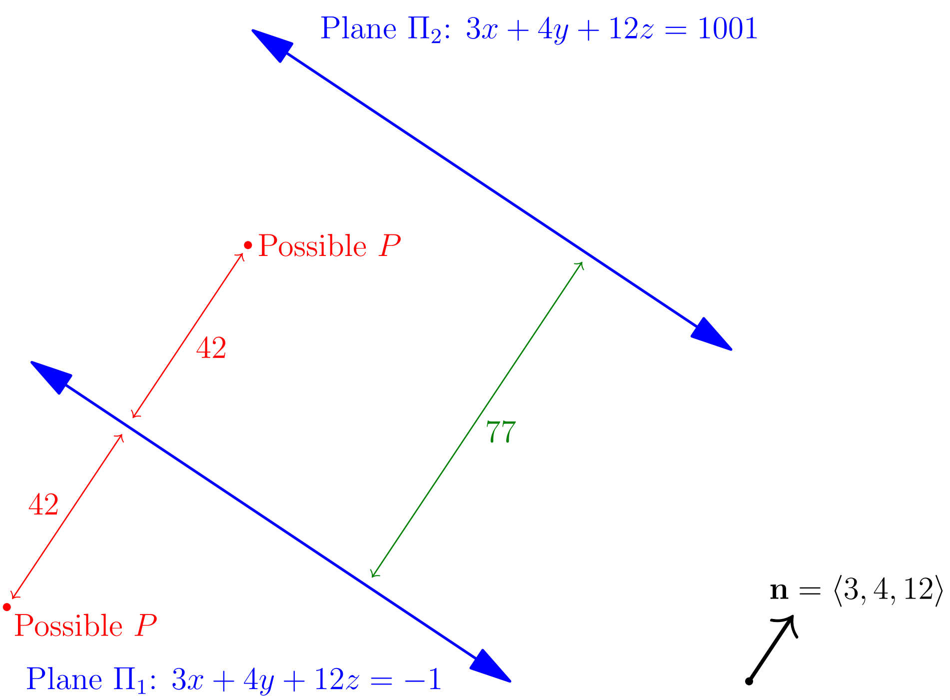

Figure 97: A cartoon showing the planes Π_1 and Π_2 and couple possible locations for the point P.

Denote the planes

$$\begin{aligned} \text{Plane } \Pi_1 &: \quad 3x + 4y + 12z = -1 \\ \text{Plane } \Pi_2 &: \quad 3x + 4y + 12z = 1000. \end{aligned}$$

See Figure 97. Let d be the desired distance.

The idea is that the planes are parallel, so there really should just be two possible answers. To do the algebra, first observe that (parallel) planes Π_1 and Π_2 have the same normal vector:

$$\mathbf{n} = \langle 3, 4, 12 \rangle$$

which has $|\mathbf{n}| = \sqrt{3^2 + 4^2 + 12^2} = 13$.

Now to compute the distance between Π_1 and Π_2, we consider an arbitrary point (x_0, y_0, z_0) on Π_2 (meaning $3x_0 + 4y_0 + 12z_0 = 1000$) and find the distance from it to Π_1. According to the recipe in Section 5.6, it equals

$$\begin{aligned} \text{distance from } (x_0, y_0, z_0) \text{ to } \Pi_1 &= \frac{|3x_0 + 4y_0 + 12z_0 + 1|}{|\mathbf{n}|} \\ &= \frac{|1000 + 1|}{13} = 77. \end{aligned}$$

Hence the answers are $d = 77 \pm 42$, that is $\boxed{d = 35}$ or $\boxed{d = 119}$.

> ***i* Remark**
>
> If you don't have the idea of looking at the distance between Π_1 and Π_2, you can still solve the problem by applying the recipe in Section 5.6 directly to P. Indeed, suppose $P = (x_1, y_1, z_1)$. Then
>
> $$\begin{aligned} 42 = \operatorname{dist}(P, \Pi_1) &= \frac{|3x_1 + 4y_1 + 12z_1 + 1|}{|\mathbf{n}|} \\ d = \operatorname{dist}(P, \Pi_2) &= \frac{|3x_1 + 4y_1 + 12z_1 - 1000|}{|\mathbf{n}|} \end{aligned}$$
>
> The first equation tells us that
>
> $$3x_1 + 4y_1 + 12z_1 = \pm 42 \cdot 13 - 1.$$
>
> The second equation tells us that
>
> $$d = \frac{|(\pm 42 \cdot 13 - 1) - 1000|}{13} = \frac{|\pm 42 \cdot 13 - 1001|}{13} = |\pm 42 - 77|$$
>
> and this gives the same answers.

§42.11 Solution to Exercise 6.1 (cross product 0)

Exercise 6.1. Suppose real numbers a and b satisfy

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 100 \\ a \\ b \end{pmatrix} = \mathbf{0}.$$

Compute a and b.

Two nonzero vectors have cross product 0 if and only if they're multiples of each other. Hence we get $\boxed{a = 200}$ and $\boxed{b = 300}$.

§42.12 Solution to Exercise 6.2 ($5\mathbf{w} \times 4\mathbf{v}$)

Exercise 6.2. Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^3$ for which $\mathbf{v} \times \mathbf{w} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Compute $5\mathbf{w} \times 4\mathbf{v}$.

Using properties of the cross product:

$$5\mathbf{w} \times 4\mathbf{v} = 20(\mathbf{w} \times \mathbf{v}) = -20(\mathbf{v} \times \mathbf{w}) = -20 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \boxed{\begin{pmatrix} -20 \\ -40 \\ -60 \end{pmatrix}}.$$

The fact that $5\mathbf{w} \times 4\mathbf{v} = 20(\mathbf{w} \times \mathbf{v})$ follows either directly from the geometric definition (the parallelogram has 20 times the area) or from looking at the algebraic definition (where the 4 and 5 factor out). Whereas the fact that $\mathbf{w} \times \mathbf{v} = -\mathbf{v} \times \mathbf{w}$ also follows directly from the right-hand rule.

§42.13 Solution to Exercise 6.3 ($|\mathbf{v} \times \mathbf{w}|^2 + (\mathbf{v} \cdot \mathbf{w})^2$)

Exercise 6.3. Let $\mathbf{v}$ and $\mathbf{w}$ be unit vectors in $\mathbb{R}^3$. Compute all possible values of

$$|\mathbf{v} \times \mathbf{w}|^2 + (\mathbf{v} \cdot \mathbf{w})^2.$$

Let θ be the angle between the vectors. Then the geometric definitions of the cross and dot products gives

$$|\mathbf{v} \times \mathbf{w}| = |\mathbf{v}|\ |\mathbf{w}|\ |\sin\theta| = 1 \cdot 1 \cdot |\sin\theta| = |\sin\theta|$$
$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}|\ |\mathbf{w}| \cos\theta = 1 \cdot 1 \cdot \cos\theta = \cos\theta.$$

Hence the answer is

$$|\sin\theta|^2 + (\cos\theta)^2 = \boxed{1}$$

by the Pythagorean theorem: there is only one possible value.

§42.14 Solution to Exercise 6.4 (solving for k)

Exercise 6.4. Suppose $\mathbf{v}$ is a vector in $\mathbb{R}^3$ and k is a real number such that

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ k \end{pmatrix}.$$

Compute k.

The point is that $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 5 \\ k \end{pmatrix}$ are supposed to be perpendicular; the vector $\mathbf{v}$ is otherwise completely irrelevant. For them to be perpendicular we need

$$0 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ k \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3k = 3k + 14 \implies \boxed{k = -\frac{14}{3}}$$

.

Chapter 43. Solutions to Part Bravo

§43.1 Solution to Exercise 7.2 (rotate and reflect)

Exercise 7.2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map that rotates each vector in $\mathbb{R}^2$ by 30° counterclockwise about the origin, then reflects around the line $y = x$. Write T as a 2×2 matrix.

We calculate the outputs of T on the basis vectors $\mathbf{e}_1 = \langle 1, 0 \rangle$ and $\mathbf{e}_2 = \langle 0, 1 \rangle$.

For $\mathbf{e}_1$, we first end up with

$$\mathbf{e}_1 = \langle 1, 0 \rangle \longrightarrow \langle \cos 30^\circ, \sin 30^\circ \rangle = \left\langle \frac{\sqrt{3}}{2}, \frac{1}{2} \right\rangle \longrightarrow \left\langle \frac{1}{2}, \frac{\sqrt{3}}{2} \right\rangle = T(\mathbf{e}_1)$$

(with the first arrow being the rotation and the second arrow being reflection).

For $\mathbf{e}_2$, we end up with

$$\mathbf{e}_2 = \langle 0, 1 \rangle \longrightarrow \langle \cos 120^\circ, \sin 120^\circ \rangle = \left\langle -\frac{1}{2}, \frac{\sqrt{3}}{2} \right\rangle \longrightarrow \left\langle \frac{\sqrt{3}}{2}, -\frac{1}{2} \right\rangle = T(\mathbf{e}_2).$$

Hence, the answer is

$$T = \boxed{\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}}.$$

Alternatively, one could obtain the same answer by multiplying the matrices corresponding to counterclockwise rotation and reflection around $y = x$, that is

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos 30^\circ & \cos 120^\circ \\ \sin 30^\circ & \sin 120^\circ \end{pmatrix}$$

would work out to the same thing. This is just an affirmation that Section 7.3 holds true: applying two transformations is the same as multiplying their corresponding matrices.

§43.2 Solution to Exercise 8.1 (birthday)

Exercise 8.1. Take your birthday and write it in eight-digit $Y_1Y_2Y_3Y_4$-M_1M_2-D_1D_2 format. Consider the two vectors

$$\mathbf{v}_1 = \begin{pmatrix} M_1M_2 \\ D_1D_2 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} Y_1Y_2 \\ Y_3Y_4 \end{pmatrix}.$$

For example, if your birthday was May 17, 1994 you would take $\mathbf{v}_1 = \binom{5}{17}$ and $\mathbf{v}_2 = \binom{19}{94}$.

- Compute the span of those two vectors in $\mathbb{R}^2$.
- Find a current or former K-pop idol who gets a different answer from you when they use their birthday.

Neither $\mathbf{v}_1$ nor $\mathbf{v}_2$ is zero, and for almost everyone the two vectors won't be a multiple of each other. So for most people the answer is that the span is $\boxed{\text{all of } \mathbb{R}^2}$.

In order to find a K-pop idol whose two vectors are linearly dependent (to get the answer "line" instead), we need to find a database of K-pop birthdays, and we need to know where to look in it. There are roughly two strategies you can adopt:

- For idols born before 2000, the only year that's viable is 1995 (because 19 is a prime greater than 12, the last two digits need to be a multiple of 19). The two days that work here are January 5 and February 10. As an example, Jo Sangho from former boy group Snuper was born on February 10, 1995:

$$\mathbf{v}_1 = \begin{pmatrix} 02 \\ 10 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 19 \\ 95 \end{pmatrix}.$$

- For idols born after 2000, good years to try would be 2004 or 2005. (The year 2004 has May 1 and October 2; the year 2005 has April 1, October 2, December 3.) As an example, Machida Riku from KJRGL was born on October 2, 2004:

$$\mathbf{v}_1 = \begin{pmatrix} 10 \\ 02 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 20 \\ 04 \end{pmatrix}.$$

§43.3 Solution to Exercise 8.2 (maximum perpendicular vectors)

Exercise 8.2. In $\mathbb{R}^5$, consider the vector $\mathbf{v} = \langle 1, 2, 3, 4, 5 \rangle$. Compute the maximum possible number of linearly independent vectors one can find which are all perpendicular to $\mathbf{v}$.

These vectors all need to lie in a hyperplane perpendicular to $\mathbf{v}$, which is a $\boxed{4}$-dimensional space. (The entries of the vector $\mathbf{v}$ are irrelevant besides $\mathbf{v}$ not being the zero vector.)

§43.4 Solution to Exercise 9.1 (four 2-by-2 matrices)

Exercise 9.1. Compute the eigenvalues and eigenvectors for the following matrices:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 1 \\ 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix}, \quad D = \begin{pmatrix} 6 & 1 \\ 0 & 6 \end{pmatrix}.$$

This is done just by following the recipe. Here are the answers.

Matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ To find the eigenvalues λ, we solve the characteristic equation:

$$\det(A - \lambda I) = 0$$

Where I is the identity matrix.

$$A - \lambda I = \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix}$$
$$\det(A - \lambda I) = (1-\lambda)^2 - (1)(1) = \lambda^2 - 2\lambda = \lambda(\lambda - 2).$$

Thus, the eigenvalues are:

$$\lambda_1 = 0, \quad \lambda_2 = 2$$

Now we compute the eigenvectors:

- **For** $\lambda_1 = 0$:

$$
\begin{aligned}
(A - 0I)\mathbf{v} = A\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x+y \\ x+y \end{pmatrix} = \mathbf{0} \\
x + y = 0 \implies y = -x &
\end{aligned}
$$

Thus, the eigenvectors corresponding to $\lambda_1 = 0$ are all the multiples of

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

- **For $\lambda_2 = 2$:**

$$
\begin{aligned}
(A - 2I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -x+y \\ x-y \end{pmatrix} = \mathbf{0} \\
-x + y = 0 \implies y = x &
\end{aligned}
$$

Thus, the eigenvector corresponding to $\lambda_2 = 2$ are all the multiples of

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Matrix $B = \left(\begin{smallmatrix} 5 & 1 \\ 2 & 4 \end{smallmatrix}\right)$ Solve

$$0 = \det(B - \lambda I) = \begin{vmatrix} 5-\lambda & 1 \\ 2 & 4-\lambda \end{vmatrix} = \lambda^2 - 9\lambda + 18 = (\lambda - 6)(\lambda - 3).$$

Thus, the eigenvalues are:

$$\lambda_1 = 6, \quad \lambda_2 = 3.$$

Now we compute the eigenvectors:

- **For $\lambda_1 = 6$:**

$$
\begin{aligned}
(B - 6I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -x+y \\ 2x-2y \end{pmatrix} = \mathbf{0} \\
-x + y = 0 \implies y = x &
\end{aligned}
$$

Thus, the eigenvectors corresponding to $\lambda_1 = 6$ are the multiples of

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- **For $\lambda_2 = 3$:**

$$
\begin{aligned}
(B - 3I)\mathbf{v} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 2x+y \\ 2x+y \end{pmatrix} = \mathbf{0} \\
2x + y = 0 \implies y = -2x &
\end{aligned}
$$

Thus, the eigenvectors corresponding to $\lambda_2 = 3$ are the multiples of

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Matrix $C = \left(\begin{smallmatrix} 9 & 0 \\ 0 & 9 \end{smallmatrix}\right)$ Note the matrix C is actually 9 times the identity matrix: hence the only eigenvalue is 9, and in fact *every* vector in $\mathbb{R}^2$ is an eigenvector.

Matrix $D = \begin{pmatrix} 6 & 1 \\ 0 & 6 \end{pmatrix}$ Solve

$$\det(D - \lambda I) = \begin{vmatrix} 6-\lambda & 1 \\ 0 & 6-\lambda \end{vmatrix} = (6-\lambda)^2 - (1)(0) = (6-\lambda)^2 = 0$$

Thus, the unique eigenvalue is:

$$\lambda = 6.$$

To find the eigenvector, solve

$$(D - 6I)\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix} = \mathbf{0}$$

$$y = 0$$

Thus, the eigenvectors satisfy $y = 0$. Therefore, the eigenvectors are all non-zero vectors of the form:

$$\mathbf{v} = \begin{pmatrix} x \\ 0 \end{pmatrix}.$$

§43.5 Solution to Exercise 9.2 (eigenvectors 5 and 7)

Exercise 9.2. Give an example of a 2×2 matrix T with four nonzero entries whose eigenvalues are 5 and 7. Then compute the corresponding eigenvectors.

We arbitrarily pick two vectors with nonzero entries to be the eigenvectors, say:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\mathbf{w} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then we seek a matrix T such that

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

$$T\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ -7 \end{pmatrix}.$$

At this point one could brute-force solve a system of equations with $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in four unknowns. On the other hand, it's more economical to add the two equations and get

$$T\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 12 \\ -2 \end{pmatrix} \implies T(\mathbf{e}_1) = \begin{pmatrix} 6 \\ -1 \end{pmatrix}$$

whereas if we subtract instead we get

$$T\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 12 \end{pmatrix} \implies T(\mathbf{e}_2) = \begin{pmatrix} -1 \\ 6 \end{pmatrix}.$$

This gives us one valid matrix:

$$T = \boxed{\begin{pmatrix} 6 & -1 \\ -1 & 6 \end{pmatrix}}.$$

And we already know the eigenvectors by construction: they are the multiples of $\mathbf{v}$ (for $\lambda = 5$) and the multiples of $\mathbf{w}$ (for $\lambda = 7$).

§43.6 Solution to Exercise 9.3 (6-by-6 matrix eigenvectors)

Exercise 9.3 (*). Compute the eigenvectors and eigenvalues of the 6×6 matrix

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{pmatrix}.$$

(You can do this question without using any determinants.)

The basic idea is that we can basically work with bare hands, not needing to resort to the determinants we saw before. Specifically, suppose we have a proposed nonzero eigenvector $\mathbf{v} = \langle x_1, x_2, x_3, x_4, x_5, x_6 \rangle$, with eigenvalue λ. Then we are hoping for

$$M\mathbf{v} = \begin{pmatrix} 5x_1 \\ -9x_2 \\ 5x_3 \\ 0 \\ 8x_5 \\ x_5 + 8x_6 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \\ \lambda x_4 \\ \lambda x_5 \\ \lambda x_6 \end{pmatrix} = \lambda \mathbf{v}.$$

Setting each component equal, we get six equations that say the following, in order:

1. Either $x_1 = 0$ or $\lambda = 5$.
2. Either $x_2 = 0$ or $\lambda = -9$.
3. Either $x_3 = 0$ or $\lambda = 5$.
4. Either $x_4 = 0$ or $\lambda = 0$.
5. Either $x_5 = 0$ or $\lambda = 8$.
6. We have $(\lambda - 8)x_6 = x_5$. In particular, if $x_6 \neq 0$ then either $\lambda = 8$ or $x_5 \neq 0$, but the previous item tells $x_5 \neq 0$ forces $\lambda = 8$ anyway.

Since at least one of x_i should be nonzero (since we always ignore $\mathbf{v} = \mathbf{0}$), it follows

$$\boxed{\lambda = -9, 0, 5, 8}$$

are the eigenvalues possible. And we can read off the corresponding eigenvectors from the above six numbered items:

- For $\lambda = -9$, the eigenvectors are $\langle 0, x_2, 0, 0, 0, 0 \rangle$ for any choice of x_2.
- For $\lambda = 0$, the eigenvectors are $\langle 0, 0, 0, x_4, 0, 0 \rangle$ for any choice of x_4.
- For $\lambda = 5$, the eigenvectors are $\langle x_1, 0, x_3, 0, 0, 0 \rangle$ for any choice of x_1 and x_3.
- For $\lambda = 8$, the eigenvectors are $\langle 0, 0, 0, 0, 0, x_6 \rangle$ for any choice of x_6. (The last equation above, when $\lambda = 8$, implies $x_5 = 0$.)

§43.7 Solution to Exercise 9.4 (computing M^{20})

Exercise 9.4 (*). Using the procedure described in Section 9.8, show that

$$\begin{pmatrix} 4 & 3 \\ 6 & 7 \end{pmatrix}^{20} = \begin{pmatrix} 33333333333333333334 & 33333333333333333333 \\ 66666666666666666666 & 66666666666666666667 \end{pmatrix}.$$

(Each number on the right-hand side is 20 digits.)

We'll follow the idea in Section 9.8: find a basis of eigenvectors of M and use that to compute powers of M.

As usual, to find the eigenvalues for M we work with

$$0 = \det M = \begin{vmatrix} 4-\lambda & 3 \\ 6 & 7-\lambda \end{vmatrix} = (4-\lambda)(7-\lambda) - 18 = \lambda^2 - 11\lambda + 10 = (\lambda-1)(\lambda-10)$$

so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 10$. Let's find the corresponding eigenvectors again.

- For the eigenvalue $\lambda_1 = 1$, we need
$$\begin{pmatrix} 4-1 & 3 \\ 6 & 7-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff x+y=0 \iff y=-x$$
so the eigenvectors are all the multiples of $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
- For the eigenvalue $\lambda_2 = 10$, we need
$$\begin{pmatrix} 4-10 & 3 \\ 6 & 7-10 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff -2x+y=0 \iff y=2x$$
so the eigenvectors are all the multiples of $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Hence, when raising to the 20th power, we should have

$$\begin{aligned} M^{20}\mathbf{v}_1 = 1^{20}\mathbf{v}_1 &\implies M^{20}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ M^{20}\mathbf{v}_2 = 10^{20}\mathbf{v}_2 &\implies M^{20}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 10^{20} \\ 2\cdot 10^{20} \end{pmatrix}. \end{aligned}$$

Hence, we've found M^{20} on two linearly independent vectors! As we showed in Section 9.8, this means we should be able to recover M^{20} on the basis vectors too.

To get the first column of M^{20}, we write

$$\begin{aligned} M^{20}(\mathbf{e}_1) &= M^{20}\left(\frac{2}{3}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) = \frac{2}{3}M^{20}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{1}{3}M^{20}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} (2+10^{20})/3 \\ (-2+2\cdot 10^{20})/3 \end{pmatrix} \\ &= \begin{pmatrix} 100000000000000000002/3 \\ 199999999999999999998/3 \end{pmatrix} = \begin{pmatrix} 33333333333333333334 \\ 66666666666666666666 \end{pmatrix} \end{aligned}$$

as needed for the first column.

To get the second column of M^{20}, we write

$$\begin{aligned} M^{20}(\mathbf{e}_2) &= M^{20}\left(\frac{1}{3}\begin{pmatrix}1\\2\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\-1\end{pmatrix}\right) = \frac{1}{3}M^{20}\begin{pmatrix}1\\2\end{pmatrix} - \frac{1}{3}M^{20}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}(10^{20}-1)/3\\(2\cdot 10^{20}+1)/3\end{pmatrix} \\ &= \begin{pmatrix}99999999999999999999/3\\200000000000000000001/3\end{pmatrix} = \begin{pmatrix}33333333333333333333\\66666666666666666667\end{pmatrix} \end{aligned}$$

as needed for the second column. This completes the solution.

Chapter 44. Solutions to Part Charlie

§44.1 Solution to Exercise 10.1 ($101 \cdot 401 \cdot 901$)

Exercise 10.1 (*). Without a calculator, give an example of an ordered pair (a, b) of integers satisfying

$$a^2 + b^2 = 101 \cdot 401 \cdot 901.$$

The idea is to consider the complex number

$$(10 + i)(20 + i)(30 + i) = (199 + 30i)(30 + i) = 5940 + 1099i.$$

Hence one possible choice is

$$\boxed{(a, b) = (5940, 1099).}$$

§44.2 Solution to Exercise 11.1 (projection onto plane)

Exercise 11.1. In $\mathbb{R}^3$, compute the projection of the vector $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ onto the plane $x + y + 2z = 0$.

Answer: $\begin{pmatrix} \frac12 \\ \frac32 \\ -1 \end{pmatrix}$.

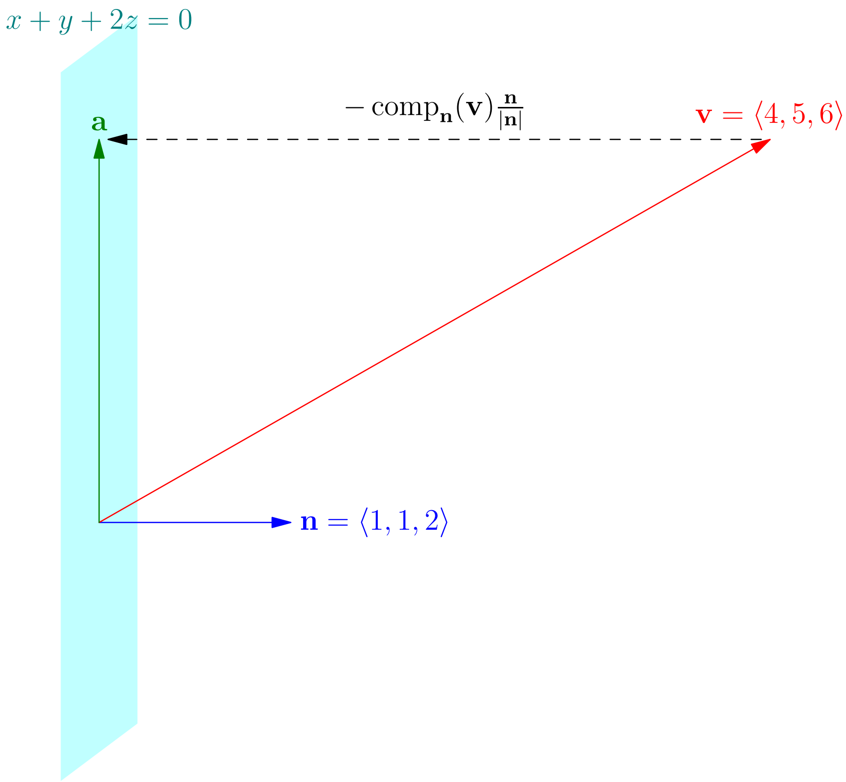

Figure 98: Projection onto a plane.

First approach using vector projection

Previously you had to calculate the distance from a vector to a plane. This problem only requires one step on top of that: you need to then translate by that multiple of the normal vector. See Figure 98, where $\mathbf{a}$ denotes the answer. To execute the calculation, let $\mathbf{v} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ and $\mathbf{n} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$. The scalar component is

$$\operatorname{comp}_{\mathbf{n}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{n}}{|\mathbf{n}|} = \frac{21}{\sqrt{6}}.$$

The vector projection is then

$$(\operatorname{comp}_{\mathbf{n}}(\mathbf{v}))\frac{\mathbf{n}}{|\mathbf{n}|} = \frac{21}{\sqrt{6}} \frac{\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}}{\sqrt{6}} = \begin{pmatrix} \frac{7}{2} \\ \frac{7}{2} \\ 7 \end{pmatrix}.$$

Then the desired projection is

$$\mathbf{v} - \operatorname{proj}_{\mathbf{n}}(\mathbf{v}) = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ -1 \end{pmatrix}.$$

Second approach using normal vectors only (no projection stuff)

A lot of you don't find vector projection natural (I certainly don't). So it might be easier to imagine shifting $\mathbf{v}$ by *some* multiple of $\mathbf{n} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and then work out which multiple it is.

Specifically, we're looking for[26] a real number $t \in \mathbb{R}$ such that the vector

$$\mathbf{a} = \mathbf{v} - t\mathbf{n} = \begin{pmatrix} 4-t \\ 5-t \\ 6-2t \end{pmatrix}$$

lies on the plane $x + y + 2z = 0$. But we can actually solve for t just by plugging this $\mathbf{a}$ into the equation of the plane:

$$(4-t) + (5-t) + 2(6-2t) = 0 \implies 21 - 6t = 0 \implies t = \frac{7}{2}.$$

Hence the answer

$$\mathbf{a} = \begin{pmatrix} 4 - \frac{7}{2} \\ 5 - \frac{7}{2} \\ 6 - 2(\frac{7}{2}) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \\ -1 \end{pmatrix}.$$

§44.3 Solution to Exercise 11.2 (geometric interpretation)

[26] In comparison to the first solution, the value of t is exactly

$$t = \frac{\operatorname{comp}_{\mathbf{n}}(\mathbf{v})}{|\mathbf{n}|}.$$

But the idea behind the second solution is that you don't *need to know* what the geometric formula of t is. You can just solve for t indirectly by asserting that $\mathbf{a}$ lies on $x + y + 2z = 0$.

Exercise 11.2 (*). Suppose A, B, C, D are points in $\mathbb{R}^3$. Give a geometric interpretation for this expression:

$$\frac{1}{6}\left|\overrightarrow{DA}\cdot\left(\overrightarrow{DB}\times\overrightarrow{DC}\right)\right|.$$

Answer: The quantity

$$\frac{1}{6}\left|\overrightarrow{DA}\cdot\left(\overrightarrow{DB}\times\overrightarrow{DC}\right)\right|.$$

equals the volume of the tetrahedron $ABCD$.

In general, the volume of the tetrahedron is $\frac{1}{6}$ the area of the parallelepiped formed by $\overrightarrow{DA}$, $\overrightarrow{DB}$, $\overrightarrow{DC}$. Se we will prove that

$$\left|\overrightarrow{DA}\cdot\left(\overrightarrow{DB}\times\overrightarrow{DC}\right)\right|$$

gives the volume of that parallelepiped. Here are two approaches for proving it.

First approach using coordinates

Let $D = (0,0,0)$, $A = (x_A, y_A, b_A)$, $B = (x_B, y_B, z_B)$, $C = (x_C, y_C, z_C)$. Then expanding the cross product gives

$$(x_A\mathbf{e}_1 + y_A\mathbf{e}_2 + z_A\mathbf{e_3})\cdot\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_B & y_B & z_B \\ x_C & y_C & z_C \end{vmatrix}.$$

If you think about what evaluating the determinant using the formula together with the dot product would give, you should find it's actually just

$$\begin{vmatrix} x_A & y_A & z_A \\ x_B & y_B & z_B \\ x_C & y_C & z_C \end{vmatrix}$$

which is the volume of the parallelepiped.

Second approach using geometric picture

The cross product $\overrightarrow{DB}\times\overrightarrow{DC}$ is a vector whose area is equal to the parallelogram formed by $\overrightarrow{DB}$ and $\overrightarrow{DC}$. The dot product of that cross product against $\overrightarrow{DA}$ is equal to the *height* of A to plane BCD times this area, and the volume is the height times the area. See the following picture from https://en.wikipedia.org/wiki/Triple_product (in the Wikipedia figure, $\mathbf{a}$ denotes our $\overrightarrow{DA}$, etc.).

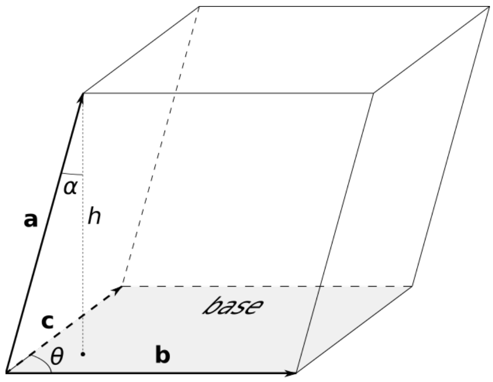

Figure 99: Triple product image taken from Wikipedia.

§44.4 Solution to Exercise 11.3 (determinant of projection)

Exercise 11.3 (*). Fix a plane $\mathcal{P}$ in $\mathbb{R}^3$ which passes through the origin. Consider the linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$ where $f(\mathbf{v})$ is the projection of $\mathbf{v}$ onto $\mathcal{P}$. Let M denote the 3×3 matrix associated to f. Compute the determinant of M.

Answer: 0, no matter which plane $\mathcal{P}$ is picked.

First approach using basis vectors

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be the three basis vectors. Then:

- The matrix M is formed by gluing $f(\mathbf{e}_1)$, $f(\mathbf{e}_2)$, $f(\mathbf{e}_3)$ together.
- I claim the vectors $f(\mathbf{e}_1)$, $f(\mathbf{e}_2)$, $f(\mathbf{e}_3)$ are linearly dependent. After all, they are all contained in the two-dimensional plane $\mathcal{P}$ by definition, and so three vectors in a plane can't be linearly independent.
- So the determinant is equal to zero (this theorem is one of the criteria we use to check whether vectors are linearly independent or not).

Second approach using eigenvectors

Let $\mathbf{n}$ be any nonzero normal vector to $\mathcal{P}$. Then $f(\mathbf{n}) = \mathbf{0}$, so $\mathbf{n}$ is an eigenvector with eigenvalue 0. Since the determinant is the product of the eigenvalues, the determinant must be 0 too.

Third approach using coordinate change

This approach requires you to know the fact that the determinant doesn't change if you rewrite the matrices in a new basis.

Let $\mathbf{n}$ be any nonzero normal vector to $\mathcal{P}$. Pick two more unit vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ perpendicular to $\mathbf{n}$ that span $\mathcal{P}$. Then $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{n}$ are linearly independent and spanning, i.e. a basis of $\mathbb{R}^3$. So we can change coordinates to use these instead.

We know that

$$\begin{aligned} M(\mathbf{b}_1) &= \mathbf{b}_1 \\ M(\mathbf{b}_2) &= \mathbf{b}_2 \\ M(\mathbf{n}_2) &= \mathbf{0}. \end{aligned}$$

If we wrote M as a matrix *in this new basis* $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{n} \rangle$ (rather than the usual basis), we would get the matrix

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which has determinant 0.

> ***i* Remark**
>
> In fact, if you also know that the trace doesn't change when you rewrite M in a different basis, this approach shows the trace M is always exactly $1 + 1 + 0 = 2$ as well, no matter which plane $\mathcal{P}$ is picked.

§44.5 Solution to Exercise 11.4 (perpendicular unit vectors)

Exercise 11.4 (*). Let $\mathbf{a}$ and $\mathbf{b}$ be two perpendicular unit vectors in $\mathbb{R}^3$. A third vector $\mathbf{v}$ in $\mathbb{R}^3$ lies in the span of $\mathbf{a}$ and $\mathbf{b}$. Given that $\mathbf{v} \cdot \mathbf{a} = 2$ and $\mathbf{v} \cdot \mathbf{b} = 3$, compute the magnitudes of the cross products $\mathbf{v} \times \mathbf{a}$ and $\mathbf{v} \times \mathbf{b}$.

Answer: $|\mathbf{a} \times \mathbf{v}| = 3$ and $|\mathbf{b} \times \mathbf{v}| = 2$.

Since $\mathbf{v}$ is contained in the span of $\mathbf{a}$ and $\mathbf{b}$, we can just pay attention to the plane spanned by these two perpendicular unit vectors. So the geometric picture is that $\mathbf{v}$ can be drawn in a rectangle with $\mathbf{a}$ and $\mathbf{b}$ as a basis, as shown. Because $\mathbf{v} \cdot \mathbf{a} = 2$ and $\mathbf{v} \cdot \mathbf{b} = 3$, this rectangle is 2 by 3.

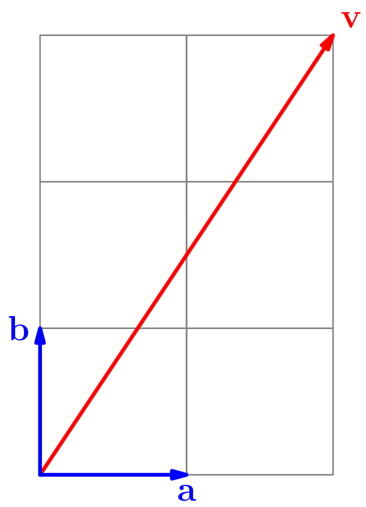

Figure 100: Plotting $\mathbf{v}$ in the span of $\mathbf{a}$ and $\mathbf{b}$.

Now the magnitude of the cross product $\mathbf{a} \times \mathbf{v}$ is supposed to be equal to the area of the parallelogram formed by $\mathbf{a}$ and $\mathbf{v}$, which is 3 (because this parallelogram has base $|\mathbf{a}| = 1$ and height $|\mathbf{v} \cdot \mathbf{b}| = 3$). Similarly, $\mathbf{b} \times \mathbf{v}$ has magnitude 2.

§44.6 Solution to Exercise 11.5 (trace of matrix)

Exercise 11.5. Compute the trace of the 2×2 matrix M given the two equations

$$M \begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \end{pmatrix} \quad \text{and} \quad M \begin{pmatrix} 5 \\ 9 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}.$$

Answer: 0.

There are several approaches possible. The first two show how to find the four entries of the matrix M. The latter sidestep this entirely and show that the matrix is actually always trace 0.

First approach: brute force

Like in the pop quiz in my R04 notes, we will try to work out $M\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $M\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. We're looking for constants c_1 and c_2 such that $c_1\left(\begin{smallmatrix}4\\7\end{smallmatrix}\right) + c_2\left(\begin{smallmatrix}5\\9\end{smallmatrix}\right) = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$.

- Solving the system of equations $4c_1 + 5c_2 = 1$ and $7c_1 + 9c_2 = 0$ using your favorite method gives coefficients $c_1 = 9$ and $c_2 = -7$, i.e.

  $$9\begin{pmatrix}4\\7\end{pmatrix} - 7\begin{pmatrix}5\\9\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}.$$

  This lets us get

  $$M\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = 9M\left(\begin{pmatrix}4\\7\end{pmatrix}\right) - 7M\left(\begin{pmatrix}5\\9\end{pmatrix}\right) = 9\begin{pmatrix}5\\9\end{pmatrix} - 7\begin{pmatrix}4\\7\end{pmatrix} = \begin{pmatrix}17\\32\end{pmatrix}.$$

- By solving the analogous system we can find the identity

  $$-5\begin{pmatrix}4\\7\end{pmatrix} + 4\begin{pmatrix}5\\9\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix},$$

  and hence:

  $$M\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = -5M\left(\begin{pmatrix}4\\7\end{pmatrix}\right) + 4M\left(\begin{pmatrix}5\\9\end{pmatrix}\right) = -5\begin{pmatrix}5\\9\end{pmatrix} + 4\begin{pmatrix}4\\7\end{pmatrix} = \begin{pmatrix}-9\\-17\end{pmatrix}.$$

Gluing these together

$$M = \begin{pmatrix}17 & -9\\32 & -17\end{pmatrix}.$$

The trace is thus $17 + (-17) = 0$.

Second approach: inverse matrices

We can collate the two given equations into saying that

$$M\begin{pmatrix}4 & 5\\7 & 9\end{pmatrix} = \begin{pmatrix}5 & 4\\9 & 7\end{pmatrix}.$$

Hence one could also recover M by multiplying by the inverse matrix:

$$M = \begin{pmatrix}5 & 4\\9 & 7\end{pmatrix}\begin{pmatrix}4 & 5\\7 & 9\end{pmatrix}^{-1} = \begin{pmatrix}5 & 4\\9 & 7\end{pmatrix}\frac{1}{4\cdot 9 - 7\cdot 5}\begin{pmatrix}9 & -5\\-7 & 4\end{pmatrix} = \begin{pmatrix}17 & -9\\32 & -17\end{pmatrix}.$$

(Of course, we get the same entries for M as the last approach.) Again the trace is $17 + (-17) = 0$.

Third approach: Guessing eigenvectors and eigenvalues

Let $\mathbf{b}_1 = \left(\begin{smallmatrix}4\\7\end{smallmatrix}\right)$ and $\mathbf{b}_2 = \left(\begin{smallmatrix}5\\9\end{smallmatrix}\right)$. Adding and subtracting the given equations gives

$$\begin{aligned}M(\mathbf{b}_1 + \mathbf{b}_2) &= \mathbf{b}_1 + \mathbf{b}_2\\ M(\mathbf{b}_1 - \mathbf{b}_2) &= -(\mathbf{b}_1 - \mathbf{b}_2).\end{aligned}$$

So $\mathbf{b}_1 \pm \mathbf{b}_2$ are eigenvectors with eigenvalues ± 1. Since M is a 2×2 matrix there are at most two eigenvalues: we found them all!

The trace of M is the sum of the eigenvalues. Call in the answer $1 + (-1) = 0$.

Fourth approach: Change coordinates

This approach requires you to know the fact that the trace doesn't change if you rewrite the matrices in a new basis.

Since $\mathbf{b}_1 = \begin{pmatrix} 4 \\ 7 \end{pmatrix}$ and $\mathbf{b}_2 = \begin{pmatrix} 5 \\ 9 \end{pmatrix}$ are a basis of $\mathbb{R}^2$, we can change coordinates to use the $\mathbf{b}_i$. In that case,

$$M(\mathbf{b}_1) = \mathbf{b}_2 \quad \text{and} \quad M(\mathbf{b}_2) = \mathbf{b}_1.$$

If we wrote M as a matrix *in this new basis* $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$ (rather than the usual basis), we would get the matrix

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which has trace $0 + 0 = 0$.

§44.7 Solution to Exercise 11.6 (complex triangle)

Exercise 11.6. There are three complex numbers z satisfying $z^3 = 5 + 6i$. Suppose we plot these three numbers in the complex plane. Compute the area of the triangle they enclose.

Answer: $\frac{3\sqrt{3}}{4} \sqrt[3]{61}$.

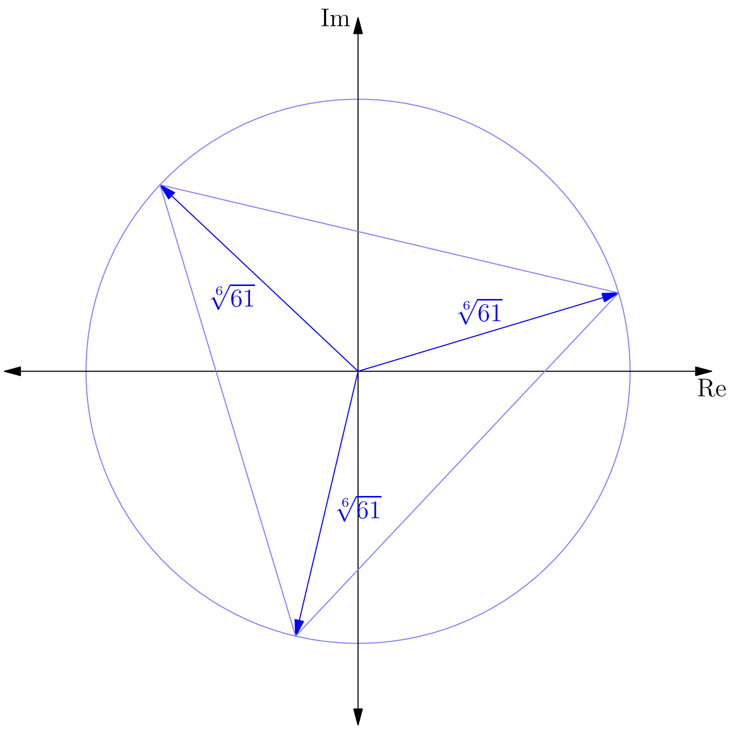

Figure 101: Three solutions to $z^3 = 5 + 6i$

We start by converting the complex number $5 + 6i$ into polar form. The modulus r of $5 + 6i$ is:

$$r = |5 + 6i| = \sqrt{5^2 + 6^2} = \sqrt{25 + 36} = \sqrt{61}.$$

The argument θ is some random angle we won't use the exact value of: $\theta = \arg(5+6i) = \tan^{-1}\left(\frac{6}{5}\right)$.

Now to find the cube roots of $z^3 = 5 + 6i$, we use the polar form:

$$z = \sqrt[6]{61}\left(\cos\left(\frac{\theta + 2k\pi}{3}\right) + i\sin\left(\frac{\theta + 2k\pi}{3}\right)\right)$$

for $k = 0, 1, 2$. This gives us three roots corresponding to the different values of k.

This looks like an equilateral triangle centered around the origin, where each spoke coming from the origin has magnitude s, where

$$s = \sqrt[6]{61}.$$

See Figure 101.

If we cut up the equilateral triangle by the three arrows above, we get three small isosceles triangles with a 120° angle at the apex. The area of each triangle is going to be $\frac{s^2}{2}\sin(120°)$.

So this gives a final answer of

$$3 \cdot \frac{\sqrt[3]{61}}{2} \cdot \sin(120°) = \frac{3\sqrt{3}}{4}\sqrt[3]{61}.$$

Chapter 45. Solutions to Part Delta

§45.1 Solution to Exercise 12.1 (parabola arc length)

Exercise 12.1. Compute the arc length of the part of the parabola $y = x^2 - x - 12$ between $(-3, 0)$ and $(4, 0)$.

You will probably need the following antiderivative fact not commonly seen in 18.01:

$$\int \sqrt{u^2+1}\,\mathrm{d}u = \frac{u}{2}\sqrt{u^2+1} + \frac{\log\left(u+\sqrt{u^2+1}\right)}{2} + C.$$

We take the parametrization

$$\mathbf{r}(t) = \left(t, t^2 - t - 12\right) \quad -3 \le t \le 4.$$

The derivative is

$$\mathbf{r}'(t) = \langle 1, 2t-1 \rangle \quad -3 \le t \le 4.$$

Hence, the arc length in the problem is given by

$$L = \int_{t=-3}^{4} \sqrt{1+(2t-1)^2}\,\mathrm{d}t.$$

To use the hint in the exercise, we perform the u-substitution

$$u = 2t - 1 = 2\left(t - \frac{1}{2}\right) \implies \mathrm{d}u = 2\,\mathrm{d}t \implies \mathrm{d}t = \frac{du}{2}.$$

When $t = -3$ we get $u = -7$ and when $t = 4$ we get $u = 7$. Thus we get

$$L = \frac{1}{2}\int_{u=-7}^{7} \sqrt{u^2+1}\,\mathrm{d}u.$$

Now using the hint we get that

$$\begin{aligned}
L &= \frac{1}{4}\left[u\sqrt{u^2+1} + \log\left(u+\sqrt{u^2+1}\right)\right]_{u=-7}^{7} \\
&= \frac{1}{4}\left(14\sqrt{50} + \log\left(7+\sqrt{50}\right) - \log\left(-7+\sqrt{50}\right)\right)) \\
&= \frac{35}{2}\sqrt{2} + \frac{1}{4}\log\left(\frac{7+5\sqrt{2}}{-7+5\sqrt{2}}\right) \\
&= \frac{35}{2}\sqrt{2} + \frac{1}{4}\log\left(\frac{\left(7+5\sqrt{2}\right)^2}{\left(7+5\sqrt{2}\right)\left(-7+5\sqrt{2}\right)}\right) \\
&= \frac{35}{2}\sqrt{2} + \frac{1}{4}\log\left(\left(7+5\sqrt{2}\right)^2\right) \\
&= \boxed{\frac{35}{2}\sqrt{2} + \frac{1}{2}\log\left(7+5\sqrt{2}\right)}.
\end{aligned}$$

§45.2 Solution to Exercise 12.2 (teacups)

Exercise 12.2. At an amusement park, a teacup ride consists of teacups rotating clockwise around a fixed center while each individual teacup rotates counterclockwise. (See Figure 27 if you've never seen one of these before.) The teacup ride is specified in $\mathbb{R}^2$ as follows:

- The teacup ride revolves around $(0,0)$ with radius R and angular velocity ω_{ride} *clockwise*.
- Each individual teacup rotates *counterclockwise* with angular velocity ω_{cup} and radius r.
- Initially, at $t = 0$, the center of the teacup is at $(R, 0)$, and a toddler is positioned at the rightmost point on the edge of the teacup relative to its center.

Compute the *velocity* vector of the toddler at time t.

We will first parametrize the motion of the toddler and then compute the distance traveled after one full revolution of the ride.

1. The teacup center rotates clockwise with angular velocity ω_{ride} in a circular path of radius R around a fixed center. The position of the teacup center as a function of time t is:

$$\mathbf{C}(t) = \begin{pmatrix} R\cos(\omega_{\text{ride}}t) \\ -R\sin(\omega_{\text{ride}}t) \end{pmatrix}.$$

   This describes the circular motion of the teacup center around the fixed center of the ride, with the negative sign on the sine term indicating clockwise rotation.

2. The toddler is sitting on the edge of the teacup, which rotates counterclockwise with angular velocity ω_{cup} and radius r. Initially, at $t = 0$, the toddler is positioned at $(r, 0)$ relative to the center of the teacup. The position of the toddler relative to the center of the teacup is:

$$\mathbf{T}_{\text{relative}}(t) = \begin{pmatrix} r\cos(\omega_{\text{cup}}t) \\ r\sin(\omega_{\text{cup}}t) \end{pmatrix}.$$

   This describes the counterclockwise circular motion of the toddler relative to the center of the teacup.

3. To find the total position of the toddler as a function of time, we sum the position of the teacup center $\mathbf{C}(t)$ and the position of the toddler relative to the teacup $\mathbf{T}_{\text{relative}}(t)$. The total position of the toddler is:

$$\mathbf{T}(t) = \mathbf{C}(t) + \mathbf{T}_{\text{relative}}(t).$$

   Substituting the expressions for $\mathbf{C}(t)$ and $\mathbf{T}_{\text{relative}}(t)$, we get:

$$\mathbf{T}(t) = \begin{pmatrix} R\cos(\omega_{\text{ride}}t) \\ -R\sin(\omega_{\text{ride}}t) \end{pmatrix} + \begin{pmatrix} r\cos(\omega_{\text{cup}}t) \\ r\sin(\omega_{\text{cup}}t) \end{pmatrix}.$$

   Simplifying, we have:

$$\mathbf{T}(t) = \begin{pmatrix} R\cos(\omega_{\text{ride}}t) + r\cos(\omega_{\text{cup}}t) \\ -R\sin(\omega_{\text{ride}}t) + r\sin(\omega_{\text{cup}}t) \end{pmatrix}.$$

   This gives the parametrization of the toddler's position as a function of time.

The velocity vector $\mathbf{T}'(t)$ is the derivative of the position vector $\mathbf{T}(t)$ with respect to time:

$$\boxed{\mathbf{T}'(t) = \begin{pmatrix} -R\omega_{\text{ride}}\sin(\omega_{\text{ride}}t) - r\omega_{\text{cup}}\sin(\omega_{\text{cup}}t) \\ -R\omega_{\text{ride}}\cos(\omega_{\text{ride}}t) + r\omega_{\text{cup}}\cos(\omega_{\text{cup}}t) \end{pmatrix}}.$$

§45.3 Solution to Exercise 12.3 (helicopter)

Exercise 12.3. A helicopter in $\mathbb{R}^3$ is moving upward with constant speed 5 in the $+z$ direction while its rotor blades are spinning with *clockwise* angular velocity $\frac{\pi}{3}$ and radius 2 in the horizontal plane. Let P be a point on the tip of the blade, initially at $(r, 0, 0)$.

- Parametrize the motion of a point on the tip of one of the blades as a function of time, assuming the helicopter starts at height $z = 0$ and the blade points along the positive x-axis at $t = 0$.
- Calculate the distance traveled by P from time $t = 0$ to time $t = 18$.

Let's first parametrize $\mathbf{P}(t)$:

1. Since the helicopter is moving upward with constant speed $v = 5$, the height of the helicopter at time t is given by:
$$z(t) = vt = 5t.$$
2. The point P is on the tip of the rotor blade, which is spinning clockwise with angular velocity $\omega = \frac{\pi}{3}$. In the horizontal plane, the position of P relative to the center of the rotor can be parametrized as:
$$\mathbf{v}(t) = \begin{pmatrix} r\cos(\omega t) \\ -r\sin(\omega t) \end{pmatrix},$$
where $r = 2$ is the radius of the blade, and the negative sign in the y-coordinate reflects the clockwise rotation. Thus, the position of P in the xy-plane is:
$$\mathbf{v}(t) = \begin{pmatrix} 2\cos\left(\frac{\pi}{3}t\right) \\ -2\sin\left(\frac{\pi}{3}t\right) \end{pmatrix}.$$
3. The total position of the point P as a function of time is the combination of the upward motion in the z-direction and the rotational motion in the xy-plane. Thus, the position of P is:
$$\boxed{\mathbf{P}(t) = \begin{pmatrix} 2\cos\left(\frac{\pi}{3}t\right) \\ -2\sin\left(\frac{\pi}{3}t\right) \\ 5t \end{pmatrix}}.$$

As for the distance, we first compute the velocity vector by differentiating:
$$\mathbf{P}'(t) = \frac{d}{dt}\begin{pmatrix} 2\cos\left(\frac{\pi}{3}t\right) \\ -2\sin\left(\frac{\pi}{3}t\right) \\ 5t \end{pmatrix} = \begin{pmatrix} -\frac{2\pi}{3}\sin\left(\frac{\pi}{3}t\right) \\ -\frac{2\pi}{3}\cos\left(\frac{\pi}{3}t\right) \\ 5 \end{pmatrix}.$$

The speed is the magnitude of the velocity vector:
$$|\mathbf{P}'(t)| = \sqrt{\left(-\frac{2\pi}{3}\sin\left(\frac{\pi}{3}t\right)\right)^2 + \left(-\frac{2\pi}{3}\cos\left(\frac{\pi}{3}t\right)\right)^2 + 5^2}.$$

Using the trigonometric identity $\sin^2(\theta) + \cos^2(\theta) = 1$, this simplifies to:

$$\begin{aligned} |\mathbf{P}'(t)| &= \sqrt{\left(\frac{2\pi}{3}\right)^2 + 5^2} = \sqrt{\frac{4\pi^2}{9} + 25} \\ &= \sqrt{\frac{4\pi^2}{9} + \frac{225}{9}} = \sqrt{\frac{4\pi^2 + 225}{9}} = \frac{\sqrt{4\pi^2 + 225}}{3} \end{aligned}$$

which is a constant! Hence the total distance traveled is simply

$$\text{Distance} = \int_0^{18} \frac{\sqrt{4\pi^2 + 225}}{3} \, \mathrm{d}t = 18 \cdot \frac{\sqrt{4\pi^2 + 225}}{3} = \boxed{6\sqrt{4\pi^2 + 225}}.$$

§45.4 Solution to Exercise 12.4 (clock)

Exercise 12.4 (*) (AMC 10A 2015)**.** In Figure 28, there's a circular clock with radius 20 cm and a circular disk of radius 10 cm externally tangent at the 12 o'clock position. The disk has an arrow painted that points directly up and rolls clockwise. At what point on the clock face will the disk be tangent when the arrow is next pointing in the upward vertical direction?

The answer is $\boxed{\text{4 o'clock}}$! In other words, the red disk makes *three* complete revolutions around the blue block when it goes all the way around, not just two.

This is a variation on what's called the **coin rotation paradox** (see Wikipedia), where $R = 20$ and $r = 10$. See the description there for details and an animation (when $R = r$ and $R = 3r$).

Chapter 46. Solutions to Part Echo

§46.1 Solution to Exercise 13.1 (five level curve drawings)

Exercise 13.1. Draw 2D level curves for some values for the following functions:

- $f(x,y) = \frac{3}{2}x + y$
- $f(x,y) = xy$
- $f(x,y) = \sin(x^2 + y^2)$
- $f(x,y) = e^{y-x^2}$
- $f(x,y) = \max(x,y)$ (i.e. f outputs the larger of its two inputs, so $f(3,5) = 5$ and $f(2,-9) = 2$, for example).

In what follows, c always denotes the value we're drawing the level curve.

For $f(x,y) = \frac{3}{2}x + y$ The level curves of $\frac{3}{2}x + y$ will be straight lines with slope $-\frac{3}{2}$ whose y-intercept is the point $(0,c)$. See Figure 102.

For $f(x,y) = xy$ When $c \neq 0$, the shape of $xy = c$ is a hyperbola $y = \frac{c}{x}$. For the exceptional value $c = 0$, the shape $xy = 0$ is the union of the axes. See Figure 103.

For $f(x,y) = \sin(x^2 + y^2)$ The level curve is only nonempty when $-1 \le c \le 1$. For these c, we obtain a bunch of concentric circles whose radii r satisfy $\sin\sqrt{r} = c$. For example, when $c = 0$, we get circles of radius 0, $\sqrt{\pi}$, $\sqrt{2\pi}$, and so on. See Figure 104.

For $f(x,y) = e^{y-x^2}$ The level curve is only nonempty when $c > 0$. The level curve is the parabola $y = x^2 + \log(c)$. See Figure 105.

For $f(x,y) = \max(x,y)$ The curve consists of what look like rotated L-shapes, as shown in the figure. See Figure 106.

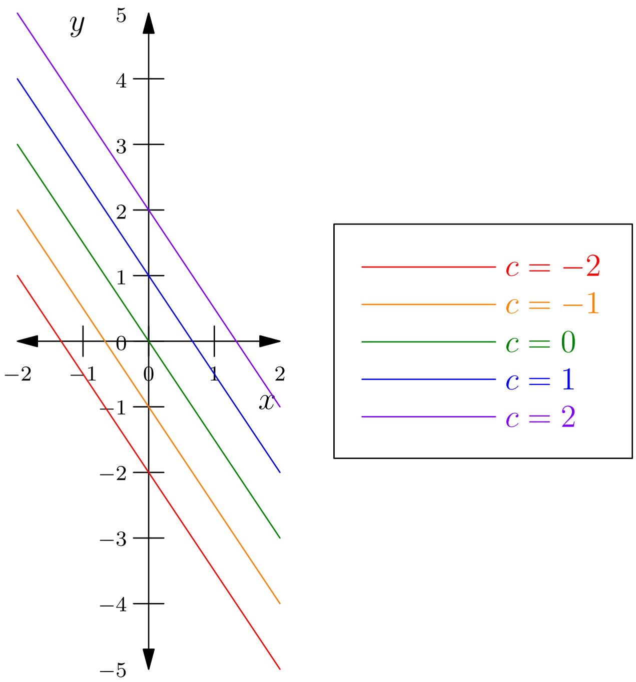

Figure 102: Level curves of $f(x,y) = \frac{3}{2}x + y$

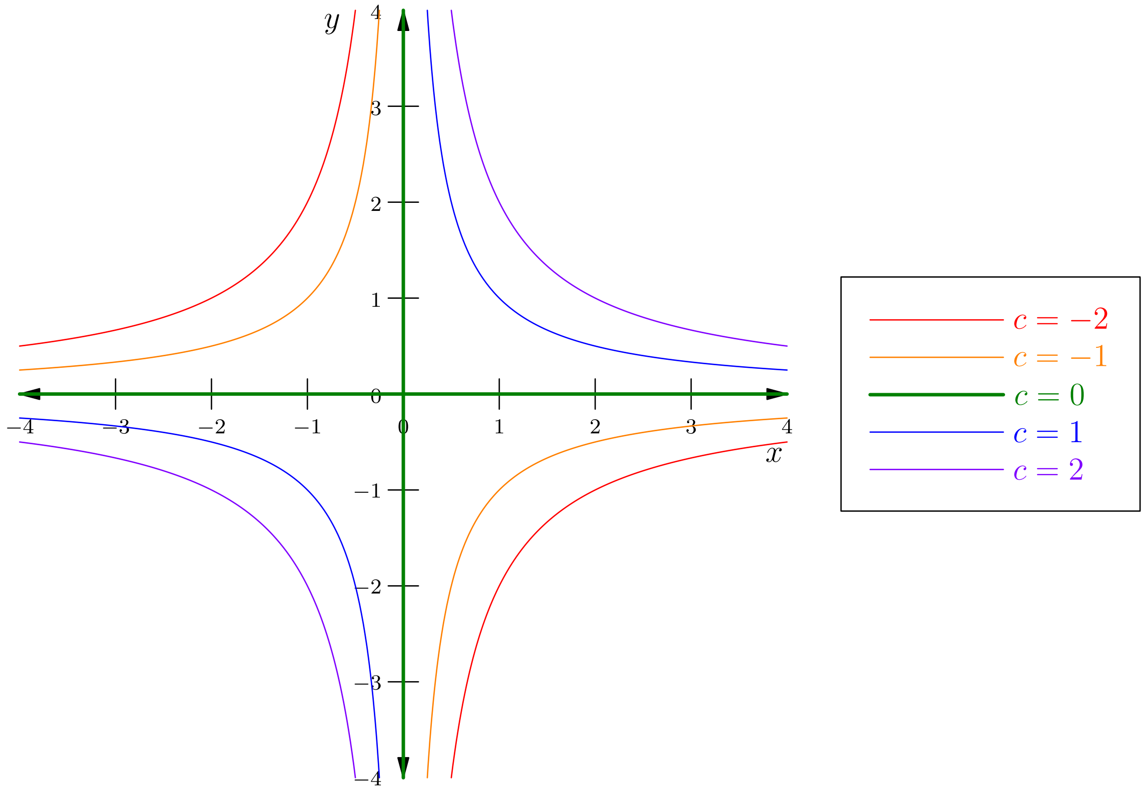

Figure 103: Level curves of $f(x, y) = xy$

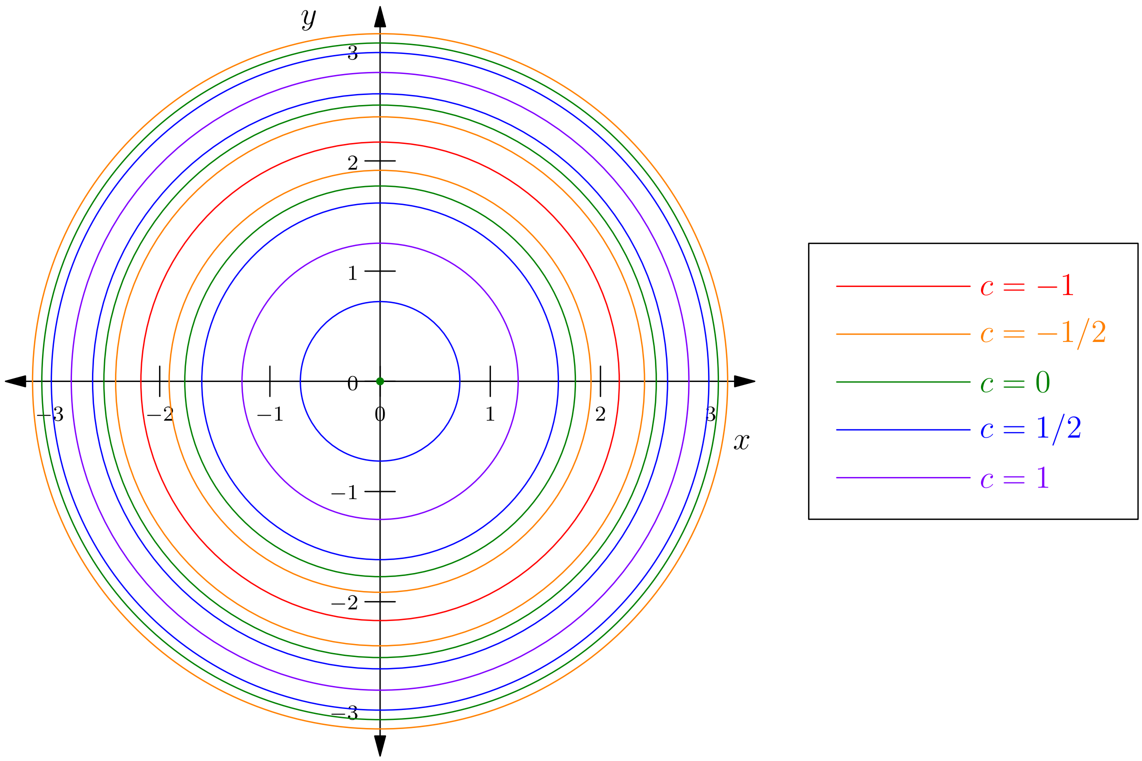

Figure 104: Level curves of $f(x, y) = \sin(x^2 + y^2)$

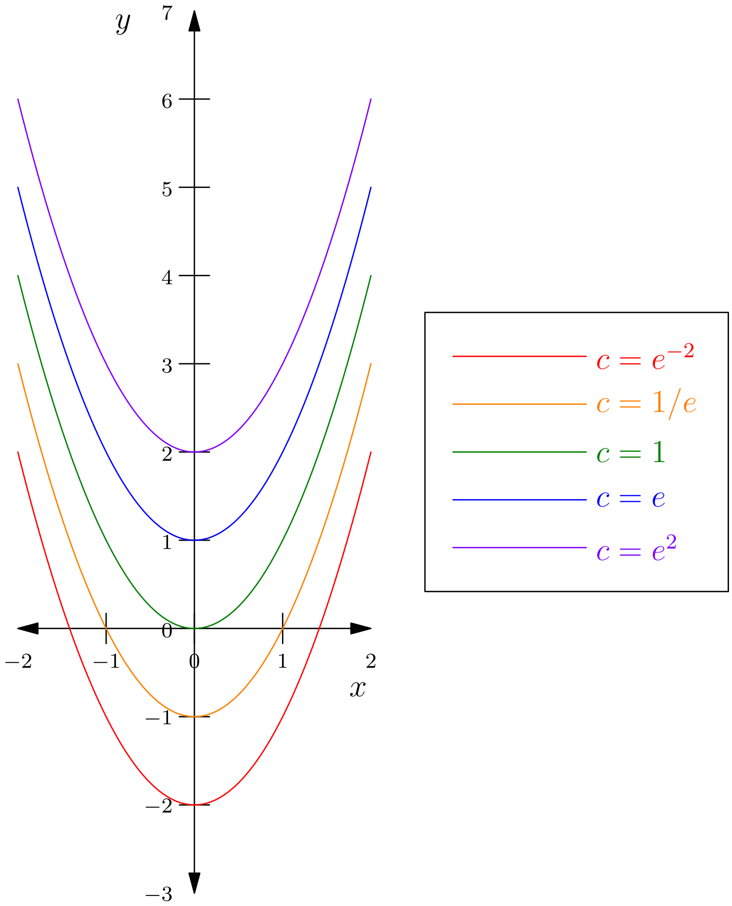

Figure 105: Level curves of $f(x, y) = e^{y-x^2}$

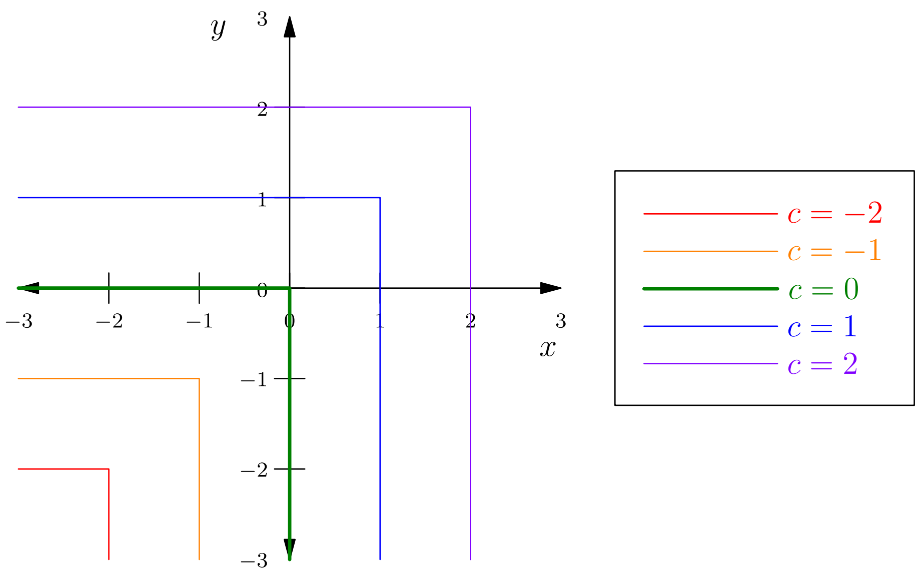

Figure 106: Level curves of $f(x, y) = \max(x, y)$

§46.2 Solution to Exercise 13.2 (level curve with seven points)

Exercise 13.2 (*). Give an example of a polynomial function $f(x, y)$ for which the level curve for the value 100 consists of exactly seven points.

This is quite tricky. The following function should work:

$$f(x, y) = 100 + (x^2 + (y - 1)^2)(x^2 + (y - 2)^2)...(x^2 + (y - 7)^2).$$

Then $f = 100$ if only if the product of the seven quadratics we gave was zero, i.e. we have $x^2 + (y - k)^2 = 0$ for some $k = 1, ..., 7$. But that can only happen when $(x, y) = (0, k)$.

In other words, this level curve for 100 consists of only seven points: $(0, 1)$, $(0, 2)$, $(0, 3)$, $(0, 4)$, $(0, 5)$, $(0, 6)$, and $(0, 7)$.

§46.3 Solution to Exercise 14.1 (partial derivative practice)

Exercise 14.1. Compute all the partial derivatives of the following functions, defined for $x, y, z > 0$:

- $f(x, y, z) = \frac{x}{y} + \frac{y}{z} + \frac{z}{x}$
- $f(x, y, z) = \sin(xyz)$
- $f(x, y, z) = x^y + y^z + z^x$.

This is direct calculation and doesn't require any trick.

- For $f(x, y, z) = \frac{x}{y} + \frac{y}{z} + \frac{z}{x}$ we compute f_x:

  $$f_x = \frac{1}{y} - \frac{z}{x^2}.$$

  For the same reason:

  $$f_y = \frac{1}{z} - \frac{x}{y^2}, \quad f_z = \frac{1}{x} - \frac{y}{z^2}.$$

- For $f(x, y, z) = \sin(xyz)$ we compute f_x:

  $$f_x = yz\cos(xyz).$$

  For the same reason:

  $$f_y = xz\cos(xyz), \quad f_z = xy\cos(xyz).$$

- For $f(x, y, z) = x^y + y^z + z^x$ we compute f_x:

  $$f_x = yx^{y-1} + (\log z)z^x.$$

  For the same reason:

  $$f_y = zy^{z-1} + (\log x)x^y, \quad f_z = xz^{x-1} + (\log y)y^z.$$

§46.4 Solution to Exercise 15.1 (tangent to sphere)

Exercise 15.1. Compute the equation of the tangent plane to the sphere $x^2 + y^2 + z^2 = 14$ at the point $(1, 2, 3)$.

The equation of the given sphere is:

$$x^2 + y^2 + z^2 = 14.$$

To find the equation of the tangent plane at the point $(1, 2, 3)$, we first compute the gradient of the function

$$F(x, y, z) = x^2 + y^2 + z^2 - 14.$$

The gradient is:

$$\nabla F = \langle 2x, 2y, 2z \rangle.$$

Evaluating at $(1, 2, 3)$:

$$\nabla F(1, 2, 3) = \langle 2, 4, 6 \rangle.$$

Hence the tangent plane should be given by

$$2x + 4y + 6z = c$$

for some number c. In order for this to pass through $(1, 2, 3)$, we take

$$c = 2 \cdot 1 + 4 \cdot 2 + 6 \cdot 3 = 28.$$

Hence the answer is

$$\boxed{2x + 4y + 6z = 28}$$

. Or one could write this as $x + 2y + 3z = 14$ if you don't like the unneeded factor of 2.

§46.5 Solution to Exercise 15.2 (given level curve is a circle)

Exercise 15.2. The level curve of a certain differentiable function $f(x, y)$ for the value -7 turns out to be a circle of radius 2 centered at $(0, 0)$.

- Give an example of one such function f.
- What are all possible vectors that $\nabla f(1.2, -1.6)$ could be?
- Do linear approximation to estimate $f(1.208, -1.594)$ starting from the point $(1.2, -1.6)$.

Examples of functions

For an example, one natural choice for $f(x, y)$ is:

$$f(x, y) = x^2 + y^2 - 11.$$

This satisfies:

$$f(x, y) = -7 \quad \text{if and only if} \quad x^2 + y^2 = 4,$$

which defines a circle of radius 2 centered at the origin. There are other examples, such as

$$f(x, y) = 100(x^2 + y^2) - 407$$

or

$$f(x, y) = e^{x^2+y^2} - (e^4 + 7)$$

and so on.

Possible gradients

For the second part, let P denote the point $(1.2, -1.6)$. Then P lies on this circle. However, from high school geometry (no calculus involved), the tangent to the circle at P is the line $3(x - 1.2) + 4(y - 1.6) = 0$. (See Figure 107.) The gradient needs to be some perpendicular to this line, so it must be some vector in the direction of $(1.2, -1.6)$. That is, $\nabla f(1.2, -1.6)$ could be[27] **any vector in the same direction as $(3, 4)$**.

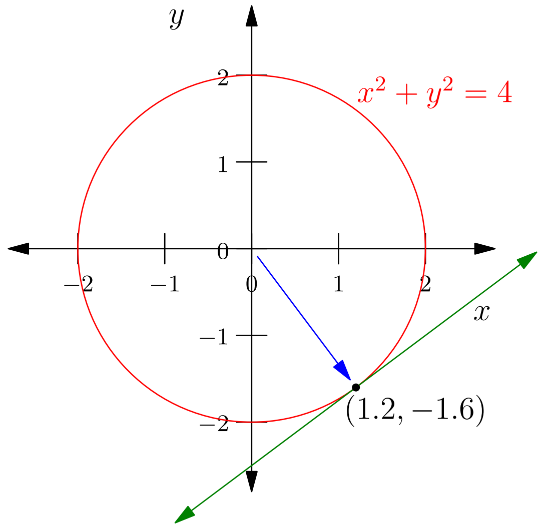

Figure 107: The gradient should be perpendicular to the tangent line (green) to the red circle at P.

Linear approximation

For the linear approximation step, we have

$$f(1.208, -1.594) \approx f(1.2, -1.6) + (0.08, 0.06) \cdot \nabla f(1.2, -1.6).$$

But the vectors $(0.08, 0.06)$ and $\nabla f(1.2, -1.6)$ are perpendicular.

§46.6 Solution to Exercise 15.3 (preview of anti-gradients)

Exercise 15.3. For each part, either give an example of $f : \mathbb{R}^2 \to \mathbb{R}$ or show that none exist.

- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle x, y \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle 100x, y \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle y, x \rangle$?
- Can you find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla f(x, y) = \langle 100y, x \rangle$?

For the first three exercises, one just takes the following:

$$\begin{aligned} f_1(x, y) &= \frac{x^2}{2} + \frac{y^2}{2} \\ f_2(x, y) &= 50x^2 + \frac{y^2}{2} \\ f_3(x, y) &= xy. \end{aligned}$$

The fourth task is impossible: no such function could exist. One will see this more clearly in later chapters, but even directly now, one might be able to realize that from $f_x = 100y$ one should have

$$f(x, y) = 100xy + (\text{stuff with only } y \text{ in it})$$

[27]Technically, one ought to show that every vector could occur for some function, but for 18.02 I won't dwell on that. Anyone who knows what I'm talking about should be able to fill in this step for me :P

and there's no way one can differentiate that with respect to y and get just x. Again, read Chapter 16 for more details.

§46.7 Solution to Exercise 15.4 (level curve through $(0, d)$)

Exercise 15.4 (*). Let a, b, c, d be nonzero real numbers and let

$$f(x, y) = ae^{x+y} + be^{x-y}.$$

Suppose the level curve of f for the value c is tangent to the line $y = 5x$ at the origin, and also passes through $(0, d)$. Compute d.

The gradient of f is given by

$$\nabla f = \begin{pmatrix} ae^{x+y} + be^{x-y} \\ ae^{x+y} - be^{x-y} \end{pmatrix}.$$

At the origin $(x, y) = (0, 0)$ we get

$$\nabla f(0, 0) = \begin{pmatrix} a + b \\ a - b \end{pmatrix}.$$

The condition that the level curve is tangent to the line $y = 5x$ at the origin means the gradient vector $\nabla f(0, 0)$ must be a multiple of the normal vector to $-5x + y = 0$, which is $\left(\begin{smallmatrix} -5 \\ 1 \end{smallmatrix}\right)$. In other words, we should have

$$\frac{a + b}{a - b} = -5 \implies 3a = 2b \implies b = \frac{3}{2}a.$$

We also know that $c = f(0, 0) = f(0, d)$. We compute

$$\begin{aligned} f(0, 0) &= ae^{0+0} + be^{0-0} = a + b = a + \frac{3}{2}a = \frac{5}{2}a \\ f(0, d) &= ae^{d} + be^{-d} = \left(e^{d} + \frac{3}{2}e^{-d}\right)a. \end{aligned}$$

As $a \neq 0$, we conclude

$$e^{d} + \frac{3}{2}e^{-d} = \frac{5}{2} \iff (e^{d} - 1)\left(e^{d} - \frac{3}{2}\right) = 0$$

and so $\boxed{d = \log\left(\frac{3}{2}\right)}$.

§46.8 Solution to Exercise 16.1 (anti-gradient practice)

Exercise 16.1. Suppose $f(x, y)$ is a differentiable function and that

$$\nabla f(x, y) = \begin{pmatrix} x^2 + axy + 2y^2 + y + 1 \\ x^2 + x + bxy + y^2 + 2 \end{pmatrix}$$

for some constants a and b. Compute the constants a and b, and determine f.

We are given the gradient of a function $f(x, y)$:

$$\nabla f(x,y) = \begin{pmatrix} x^2 + axy + 2y^2 + y + 1 \\ x^2 + x + bxy + y^2 + 2 \end{pmatrix}.$$

To compute a and b, we compute

$$\begin{aligned} \frac{\partial}{\partial y}(x^2 + axy + 2y^2 + y + 1) &= ax + 4y + 1 \\ \frac{\partial}{\partial x}(x^2 + x + bxy + y^2 + 2) &= 2x + 1 + by. \end{aligned}$$

These need to be equal for all (x, y) so we require $\boxed{(a,b) = (2,4)}$.

To recover f, we write

$$\begin{aligned} f(x,y) &= \int \frac{\partial f}{\partial x}\,\mathrm{d}x = \int (x^2 + 2xy + 2y^2 + y + 1)\,\mathrm{d}x \\ &= \frac{x^3}{3} + x^2y + 2xy^2 + xy + x + C_1(y). \\ f(x,y) &= \int \frac{\partial f}{\partial y}\,\mathrm{d}y = \int (x^2 + x + 4xy + y^2 + 2)\,\mathrm{d}y \\ &= x^2y + xy + 2xy^2 + \frac{y^3}{3} + 2y + C_2(x). \end{aligned}$$

Stitching these together to get the final expression for $f(x,y)$ as:

$$f(x,y) = \boxed{\frac{x^3}{3} + x^2y + 2xy^2 + \frac{y^3}{3} + xy + x + 2y + C}$$

for any constant C.

Chapter 47. Solutions to Part Foxtrot

§47.1 Solution to Exercise 17.1 (critical points of a 2-variable function)

Exercise 17.1. Compute the critical point(s) of $f(x,y) = x^3 + 2y^3 - 6xy$ and classify them as local minimums, local maximums, or saddle points.

Finding the critical points

The first-order partial derivatives are:

$$f_x = \frac{\partial f}{\partial x} = 3x^2 - 6y, \quad f_y = \frac{\partial f}{\partial y} = 6y^2 - 6x.$$

To find the critical points, we solve the system:

$$3x^2 - 6y = 0, \quad 6y^2 - 6x = 0.$$

Rewriting the equations:

$$x^2 = 2y, \quad y^2 = x.$$

From $y^2 = x$, substitute into $x^2 = 2y$:

$$\left(y^2\right)^2 = 2y \implies 0 = y(y^3 - 2).$$

Hence either $y = 0$ or $y = \sqrt[3]{2}$. These correspond to $x = 0$ and $x = \sqrt[3]{4}$.

Classifying the critical points

Hence the two critical points are

$$\boxed{(x,y) = (0,0) \quad \text{and} \quad (x,y) = \left(\sqrt[3]{4}, \sqrt[3]{2}\right).}$$

To classify them, compute the second-order derivatives. Then in the notation of Section 17.6,

$$A = f_{xx} = 6x, \quad C = f_{yy} = 12y, \quad B = f_{xy} = -6.$$

- At $(0,0)$, we have

$$AC - B^2 = 72(0)(0) - 36 = -36 < 0.$$

  So $\boxed{(0,0) \text{ is a saddle point}}$.
- At $\left(\sqrt[3]{4}, \sqrt[3]{2}\right)$, we have

$$AC - B^2 = 72\left(\sqrt[3]{4}\right)\left(\sqrt[3]{2}\right) - 36 = 108 > 0.$$

  Since $A, C > 0$, and $AC - B^2 > 0$, it follows $\boxed{\left(\sqrt[3]{4}, \sqrt[3]{2}\right) \text{ is a local minimum}}$.

§47.2 Solution to Exercise 17.2 (critical points of a 3-variable function)

Exercise 17.2. Compute the critical point(s) of $f(x,y,z) = x^2 + y^3 + z^4$ and classify them as local minimums, local maximums, or saddle points.

Finding the critical points

The first-order partial derivatives are:

$$f_x = \frac{\partial f}{\partial x} = 2x, \quad f_y = \frac{\partial f}{\partial y} = 3y^2, \quad f_z = \frac{\partial f}{\partial z} = 4z^3.$$

To find the critical points, we solve the system:

$$2x = 0, \quad 3y^2 = 0, \quad 4z^3 = 0.$$

Solving for each variable:

$$x = 0, \quad y^2 = 0 \Rightarrow y = 0, \quad z^3 = 0 \Rightarrow z = 0.$$

Thus, the only critical point is:

$$(0, 0, 0).$$

Classifying the critical points

Since this is a 3-variable function, we cannot classify it using the second derivative test. However, one can tell just by looking at the function that it is neither a local minimum or maximum. One simple way to do so is to note that for any small $\varepsilon > 0$ we have

$$f(0, \varepsilon, 0) = \varepsilon^3 > 0$$
$$f(0, -\varepsilon, 0) = -\varepsilon^3 < 0$$

In other words, there are always points near the origin $(0,0,0)$ which are both larger than $f(0,0,0) = 0$ and smaller than $f(0,0,0) = 0$. Hence $\boxed{(0,0,0) \text{ is a saddle point}}$.

§47.3 Solution to Exercise 17.3 (every point is a saddle point)

Exercise 17.3 (*). Does there exist a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ such that *every* point in $\mathbb{R}^2$ is a saddle point?

No, it's not possible.

We will prove that the following result:

Theorem 47.1. *Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is a function for which every point is a critical point. Then f must be the constant function.*

In particular, every point of f will be both a local minimum or local maximum. This means that f has no saddle points at all.

Proof. Consider any two points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$. We prove that $f(x_1, y_1) = f(x_2, y_2)$.

Define the function $g : [0, 1] \to \mathbb{R}$ along the line segment from (x_1, y_1) to (x_2, y_2) by

$$g(t) = f(x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1)).$$

This function g represents f restricted to the straight-line path between the two points.

Since f is differentiable, g is also differentiable on $(0, 1)$, and the derivative of g can be computed using the chain rule:

$$g'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Since $x = x_1 + t(x_2 - x_1)$ and $y = y_1 + t(y_2 - y_1)$, we have

$$\frac{dx}{dt} = x_2 - x_1, \quad \frac{dy}{dt} = y_2 - y_1.$$

Thus,

$$g'(t) = \frac{\partial f}{\partial x}(x_2 - x_1) + \frac{\partial f}{\partial y}(y_2 - y_1).$$

By assumption, every point is a critical point, meaning $\frac{\partial f}{\partial x} = 0$ and $\frac{\partial f}{\partial y} = 0$ everywhere. Therefore,

$$g'(t) = 0 \quad \text{for all} \quad t \in (0, 1).$$

Hence, g must be a constant function. So $g(0) = g(1)$ implies $f(x_1, y_1) = f(x_2, y_2)$ as needed. □

§47.4 Solution to Exercise 17.4 (every lattice point is a saddle point)

Exercise 17.4 (*). Give an example of a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$ with the following property: every lattice point (x, y) (i.e. a point where both x and y are integers) is a saddle point, and there are no other saddle points. For example, $(2, -7)$, $(100, 100)$, and $(-42, -13)$ should be saddle points, but $(\frac{1}{2}, 0)$, $(\pi, -\sqrt{2})$, and $(\sqrt{7}, \sqrt{11})$ should not be.

Inventing the function

The function

$$\boxed{f(x, y) = \cos((x - y)\pi) - \cos((x + y)\pi) = \sin(\pi x)\sin(\pi y)}$$

works. The idea is that a saddle point occurs if and only if one of these two conditions holds:

- when the cosines reach their maximum value, meaning $x - y$ and $x + y$ are even integers;
- when the cosines reach their minimum value, meaning $x - y$ and $x + y$ are odd integers.

Verifying the function works

Inventing the example above (which is one of many) is the main difficulty of the exercise. For completeness, let's just verify that the guess above does in fact work. From now on, use

$$f(x, y) = \sin(\pi x)\sin(\pi y).$$

The first-order partial derivatives are:

$$f_x = \frac{\partial f}{\partial x} = \pi\cos(\pi x)\sin(\pi y),$$
$$f_y = \frac{\partial f}{\partial y} = \pi\sin(\pi x)\cos(\pi y).$$

To find critical points, we set $f_x = 0$ and $f_y = 0$:

$$\pi\cos(\pi x)\sin(\pi y) = 0, \quad \pi\sin(\pi x)\cos(\pi y) = 0.$$

Since $\pi \neq 0$, the equations reduce to:

$$\cos(\pi x)\sin(\pi y) = 0, \quad \sin(\pi x)\cos(\pi y) = 0.$$

Solving these equations:

- $\cos(\pi x) = 0$ when $x = m + \frac{1}{2}$ for $m \in \mathbb{Z}$.
- $\sin(\pi y) = 0$ when $y = n$ for $n \in \mathbb{Z}$.
- $\sin(\pi x) = 0$ when $x = m$ for $m \in \mathbb{Z}$.
- $\cos(\pi y) = 0$ when $y = n + \frac{1}{2}$ for $n \in \mathbb{Z}$.

A critical point must satisfy both conditions simultaneously. The only common solutions occur in two cases.

- Each lattice point (m, n), where m, n are integers, is a critical point.
- Each point $\left(m + \frac{1}{2}, n + \frac{1}{2}\right)$, where m, n are integers, is a critical point.

We classify with the second derivative test. Compute

$$\begin{aligned} A = f_{xx} &= -\pi^2 \sin(\pi x)\sin(\pi y) \\ C = f_{yy} &= -\pi^2 \sin(\pi x)\sin(\pi y) \\ B = f_{xy} &= \pi^2 \cos(\pi x)\cos(\pi y). \end{aligned}$$

Now we check the cases:

- At any lattice point (m, n), we have $A = C = 0$ and $B = \pm\pi^2$, so $AC - B^2 = -\pi^4 < 0$. So every lattice point is indeed a saddle point.
- At any point of the form $\left(m + \frac{1}{2}, n + \frac{1}{2}\right)$, we have $A = C = \pm\pi^2$ and $B = 0$, so $AC - B^2 = \pi^4 < 0$. Hence there are no saddle points here.

§47.5 Solution to Exercise 19.1 (geometry optimization)

Exercise 19.1. Let ABC be the triangle in the xy-plane with vertices $A = (0, 12)$, $B = (-5, 0)$, $C = (9, 0)$. For what point P in the plane is the sum

$$PA^2 + PB^2 + PC^2$$

as small as possible?

For $P = (x, y)$ we let

$$f(x, y) = PA^2 + PB^2 + PC^2.$$

See Figure 108.

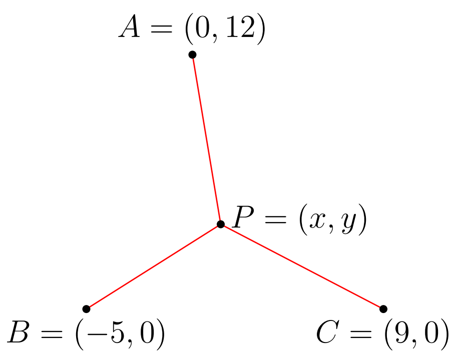

Figure 108: A triangle ABC and a point P connected to its three vertices

We first want to change f into an expression like we're used to. We do this using the Pythagorean theorem as follows:

$$\begin{aligned} PA^2 &= (x-0)^2 + (y-12)^2 = x^2 + y^2 - 24y + 144, \\ PB^2 &= (x+5)^2 + (y-0)^2 = x^2 + 10x + 25 + y^2, \\ PC^2 &= (x-9)^2 + (y-0)^2 = x^2 - 18x + 81 + y^2. \end{aligned}$$

Summing these expressions:

$$\begin{aligned} f(x,y) &= \left(x^2 + y^2 - 24y + 144\right) + \left(x^2 + 10x + 25 + y^2\right) + \left(x^2 - 18x + 81 + y^2\right) \\ &= 3x^2 + 3y^2 - 8x - 24y + 250. \end{aligned}$$

We are optimizing f over the entire space $\mathcal{R} = \mathbb{R}^2$. Let's follow the recipe in Section 19.2:

0. There is no boundary, but we have limit cases if $x \to \pm\infty$ or $y \to \pm\infty$.
1. To find the critical points, we compute the partial derivatives:

   $$f_x = \frac{\partial f}{\partial x} = 6x - 8 = 0 \implies x = \frac{4}{3}. f_y = \frac{\partial f}{\partial y} = 6y - 24 = 0 \implies y = 4.$$

   Thus, the only critical point is:

   $$P = \left(\frac{4}{3}, 4\right).$$
2. There are no boundary points to consider.
3. If either $x \to \pm\infty$ or $y \to \pm\infty$, then the quadratic terms $3x^2 + 3y^2$ dominate and cause $f(x,y) \to +\infty$. Hence f can take arbitrarily large values.

Putting this together, the point $P = \left(\frac{4}{3}, 4\right)$ is the unique point minimizing $PA^2 + PB^2 + PC^2$.

Remark

If ABC is replaced by a different triangle, it turns out that the best point P is the *center of mass* of the three points A, B, C. In other words, if $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_3, y_3)$ the answer will work out to

$$P = \left(\frac{x_1 + x_2 + x_3}{3}, \frac{y_1 + y_2 + y_3}{3}\right).$$

(This point is called the *centroid* or *gravity center* of ABC and is often denoted by the letter G.)

§47.6 Solution to Exercise 19.2 (sine optimization)

Exercise 19.2. Compute the minimum possible value of $x + y$ given that $\sin(x) + \sin(y) = 1$ and $x, y \geq 0$.

0. The region $\mathcal{R}$ has boundary whenever $x = 0$ or $y = 0$. It also has limit cases when $x \to +\infty$ or $y \to +\infty$.
1. We find the LM-critical points. The gradients are

   $$\nabla f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$\nabla g = \begin{pmatrix} \cos(x) \\ \cos(y) \end{pmatrix}.$$

From the first two equations, we have:

$$\lambda = \frac{1}{\cos(x)} \quad \text{and} \quad \lambda = \frac{1}{\cos(y)}.$$

Equating these expressions for λ, we get:

$$\frac{1}{\cos(x)} = \frac{1}{\cos(y)} \implies \cos(x) = \cos(y).$$

Thus, we get an LM-critical point whenever

$$\begin{aligned} \cos(x) &= \cos(y) \\ \sin(x) + \sin(y) &= 1 \end{aligned}$$

Note that $\sin(x)^2 = 1 - \cos(x)^2 = 1 - \cos(y)^2 = \sin(y)^2$. Since $\sin(x) + \sin(y) = 1 \neq 0$, we conclude $\sin(x) = \sin(y) = \frac{1}{2}$. Together with $\cos(x) = \cos(y)$, that means x and y must differ by a multiple of 2π.

Since we want $x + y$ to be as small as possible, we might as well take $x = y$. The smallest x for which $\sin(x) = \frac{1}{2}$ is $x = \frac{\pi}{6}$. So of all the LM-critical points, the lowest value occurs when

$$f\left(\frac{\pi}{6}, \frac{\pi}{6}\right) = \frac{\pi}{6} + \frac{\pi}{6} = \frac{\pi}{3}.$$

2. In the limit cases if either $x \to +\infty$ or $y \to +\infty$ then $f \to +\infty$.
3. Suppose $x = 0$. Then $\sin(y) = 1$. So this part of the boundary consists of the points $(0, \pi)$, $(0, 3\pi)$, $(0, 5\pi)$, …. All of these have $x + y \geq \pi$, so they do worse than the point $\left(\frac{\pi}{6}, \frac{\pi}{6}\right)$ from before.

   Similarly, if $y = 0$, we get boundary points $(\pi, 0)$, $(3\pi, 0)$, $(5\pi, 0)$, …. Again all of these have $x + y \geq \pi$, so they do worse than the point $\left(\frac{\pi}{6}, \frac{\pi}{6}\right)$ from before.

In conclusion the minimum possible value occurs at

$$f\left(\frac{\pi}{6}, \frac{\pi}{6}\right) = \frac{\pi}{3}.$$

§47.7 Solution to Exercise 19.3 (optimization with absolute value)

Exercise 19.3 (Suggested by Ting-Wei Chao)**.** Compute the global minimum of the function

$$f(x, y) = |x^2 + y^2 - 25| - 3x - 4y.$$

This problem shows a case where ∇f does not exist at certain points: the derivative of $|x|$ only exists when $x \neq 0$.

To avoid this issue, we split into cases based on the sign of the term inside the absolute value. We will split[28] into two cases $x^2 + y^2 \leq 25$ and $x^2 + y^2 \geq 25$. In each case we execute the procedure.

[28]If you wanted to, you could split the cases a bit differently. For example, you could do $x^2 + y^2 \leq 25$ and $x^2 + y^2 > 25$ so the cases don't overlap. Or you could split into three cases with $x^2 + y^2 < 25$, $x^2 + y^2 = 25$, $x^2 + y^2 > 25$. However,

Case where $x^2 + y^2 \le 25$

We follow Section 19.2.

0. The region $x^2 + y^2 \le 25$ is a two-dimensional region with no limit cases but whose boundary is $x^2 + y^2 = 25$.
1. We seek all points with $\nabla f = \mathbf{0}$. In the region $x^2 + y^2 \le 25$, we have

$$f(x, y) = 25 - x^2 - y^2 - 3x - 4y.$$

   Solving $\nabla f = \mathbf{0}$ gives

$$\begin{aligned} \frac{\partial f}{\partial x} &= -2x - 3 = 0, \\ \frac{\partial f}{\partial y} &= -2y - 4 = 0. \end{aligned}$$

   Therefore, it gives the critical point $(x, y) = \left(-\frac{3}{2}, -2\right)$. This point is indeed in the region $x^2 + y^2 < 25$, so this is a critical point.
2. We apply Lagrange multipliers (Section 19.4) on the boundary $x^2 + y^2 = 25$. Let $g(x, y) = x^2 + y^2$.
    0. This is a one-dimensional region with no boundary and no limit cases.
    1. We search for all LM-critical points. First note that ∇g is never zero on this boundary. The system of equations is

$$\begin{aligned} -3 &= \lambda \cdot 2x, \\ -4 &= \lambda \cdot 2y \\ x^2 + y^2 &= 25. \end{aligned}$$

       It's clear that λ, x, y must be nonzero. Hence, the first two equations together imply $\frac{x}{y} = \frac{3}{4}$. Hence, we get two critical points $(3, 4)$ and $(-3, -4)$ in this case.
    2. There are no boundary points to consider.
    3. There are no limit points to consider.
3. There are no limit cases.

Case where $x^2 + y^2 \ge 25$

We follow Section 19.2 again.

0. The region $x^2 + y^2 \ge 25$ is a two-dimensional region with limit cases when $x \to \pm\infty$ and $y \to \pm\infty$ but whose boundary is $x^2 + y^2 = 25$.
1. In the region $x^2 + y^2 \ge 25$, we have

$$f(x, y) = x^2 + y^2 - 25 - 3x - 4y.$$

   Solving $\nabla f = \mathbf{0}$ gives

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x - 3 = 0, \\ \frac{\partial f}{\partial y} &= 2y - 4 = 0. \end{aligned}$$

I think doing it with $x^2 + y^2 \le 25$ and $x^2 + y^2 \ge 25$ is cleanest, so you do not need to consider limit cases when $x^2 + y^2$ approaches 25. The boundary $x^2 + y^2 = 25$ is shared, so you only need to do it once.

This gives the point $\left(\frac{3}{2}, 2\right)$, but this point doesn't lie inside $x^2 + y^2 \geq 25$, so we don't need to consider it.

2. We need to apply Lagrange multipliers (Section 19.4) on the boundary $x^2 + y^2 = 25$. However, we did this already in the earlier case where $x^2 + y^2 \leq 25$. So we can just repeat the same calculation verbatim here.
3. When $|x| \to +\infty$ or $|y| \to +\infty$, the square terms dominate and $f \to +\infty$. Hence we get that f is bounded above.

Putting things together

We are searching for the global minimum of f. Aggregating the critical points, we check

$$\begin{aligned} f\left(\frac{3}{2}, 2\right) &= \frac{25}{4}, \\ f(3, 4) &= -25, \\ f(-3, -4) &= 25. \end{aligned}$$

Therefore, the global minimum is $\boxed{f(3, 4) = -25}$.

§47.8 Solution to Exercise 20.1 (butterfly)

Exercise 20.1. A butterfly is fluttering in the xy plane with position given by $\mathbf{r}(t) = \langle \cos(t), \cos(t) \rangle$, starting from time $t = 0$ at $\mathbf{r}(0) = \langle 1, 1 \rangle$.

- Compute the speed of the butterfly at $t = \frac{\pi}{3}$.
- Compute the arc length of the butterfly's trajectory from $t = 0$ to $t = 2\pi$.
- Sketch the butterfly's trajectory from $t = 0$ to $t = 2\pi$ in the xy plane.

Sketch of the trajectory

We start actually by sketching the trajectory first (even though this was the last part), since that will make it easier to see what's going on in future parts. See Figure 109. The trajectory described by $\mathbf{r}(t) = \langle \cos(t), \cos(t) \rangle$ traces out a straight line in the xy-plane because both the x- and y-coordinates are equal for all t. Specifically, the butterfly's motion follows the line $y = x$, with $t \in [0, 2\pi]$ producing oscillations between $x = 1$ and $x = -1$.

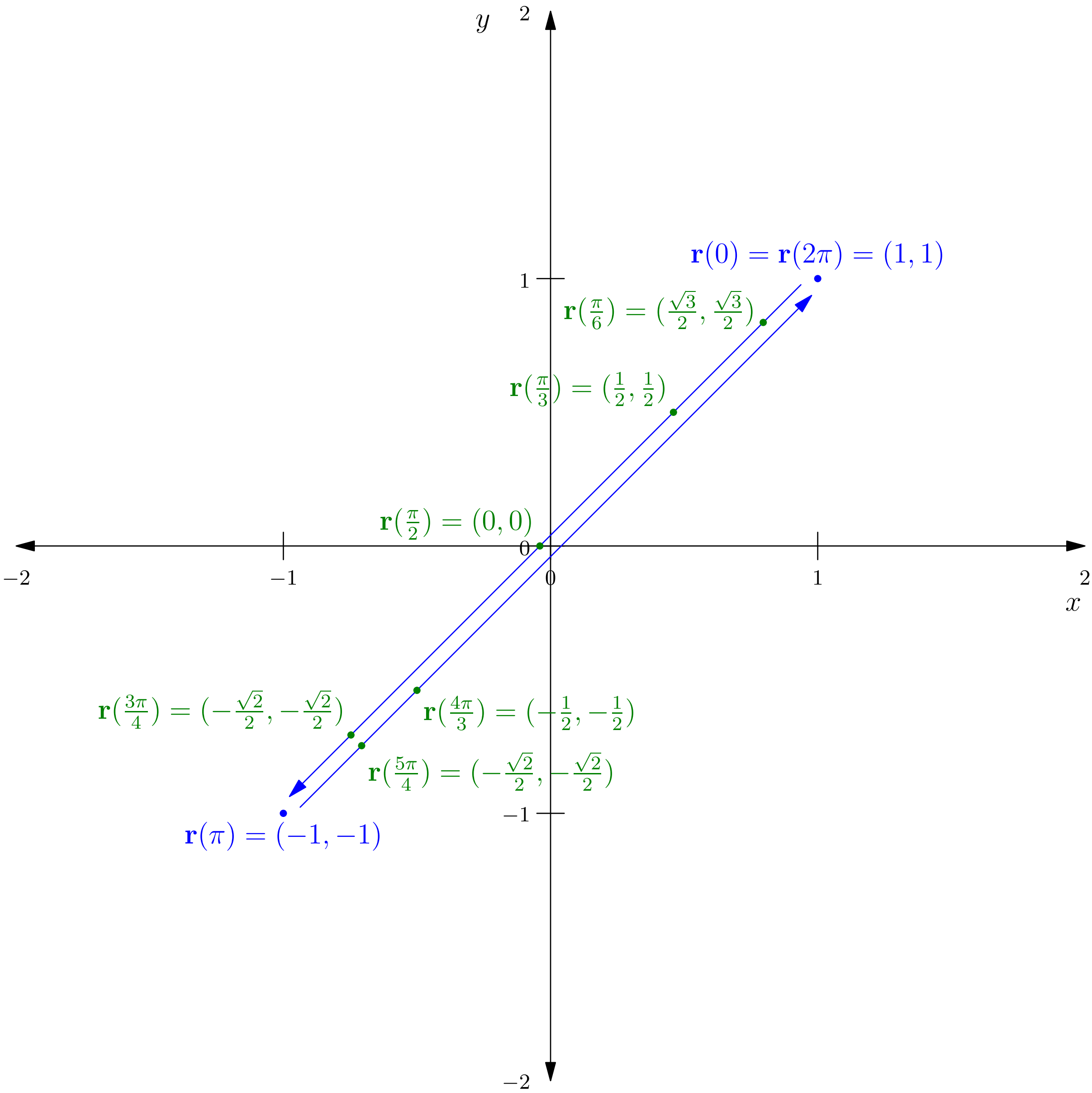

Figure 109: Butterfly fluttering along the plane. A few more examples of points in the trajectory are marked in green for illustration, but the blue endpoints are the important ones. The green points are a little offset to show both parts of the trajectory, e.g. $\mathbf{r}(\frac{\pi}{2}) = (0, 0)$ is drawn a little bit left of where it should be.

The trajectory is a straight line from $(1, 1)$ to $(-1, -1)$ and back following the line $y = x$.

Speed of the butterfly at $t = \frac{\pi}{3}$

The speed of the butterfly is given by the magnitude of its velocity vector, which is the derivative of $\mathbf{r}(t)$ with respect to time t.

First, compute the velocity $\mathbf{r}'(t)$:

$$\mathbf{r}'(t) = \frac{d}{\mathrm{d}t}\langle \cos(t), \cos(t) \rangle = \langle -\sin(t), -\sin(t) \rangle.$$

(This has direction along the line $y = x$, which is what we expect.)

The speed at any time t is the magnitude of the velocity vector:

$$\text{Speed} = |\mathbf{r}'(t)| = \sqrt{(-\sin(t))^2 + (-\sin(t))^2} = \sqrt{2\sin^2(t)} = \sqrt{2}|\sin(t)|.$$

At $t = \frac{\pi}{3}$, we have:

$$\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}.$$

Thus, the speed at $t = \frac{\pi}{3}$ is:

$$\text{Speed} = \sqrt{2} \cdot \frac{\sqrt{3}}{2} = \frac{\sqrt{6}}{2}.$$

Arc length of the butterfly's trajectory from $t = 0$ to $t = 2\pi$

Note that from the sketch of the trajectory, we can actually find the arc length with no calculus at all. Indeed, "arc length" is a misnomer because the "arc" is just two line segments!

From the Pythagorean theorem, distance from $(1,1)$ to $(-1,-1)$ is

$$\sqrt{(1-(-1))^2 + (1-(-1))^2} = \sqrt{4+4} = 2\sqrt{2}.$$

So the total distance is

$$2\sqrt{2} + 2\sqrt{2} = 4\sqrt{2}.$$

Of course, one could also use the arc length formula, and we show how to do so. The arc length of the trajectory is given by the integral of the speed:

$$L = \int_{t=\text{start time}}^{\text{stop time}} |\mathbf{r}'(t)| \, \mathrm{d}t.$$

We just saw that $|\mathbf{r}'(t)| = \sqrt{2}|\sin(t)|$. Therefore, the arc length from $t = 0$ to $t = 2\pi$ is:

$$L = \int_0^{2\pi} \sqrt{2}|\sin(t)| \, \mathrm{d}t.$$

> ⚠ **Warning**
>
> Don't forget about the absolute value! In general, for real X, we have $\sqrt{X^2} = |X|$. If you forget the absolute value here, you'll end up getting 0 as the answer, which doesn't make sense because the butterfly certainly traveled more than 0 distance. Remember, speed (absolute value of velocity vector) should always be nonnegative.

Because of the absolute value, we can break the integral into two parts. On the interval $[0, \pi]$, $\sin(t) \geq 0$, and on the interval $[\pi, 2\pi]$, $\sin(t) \leq 0$, so

$$L = \sqrt{2}\left(\int_0^{\pi} \sin(t) \, \mathrm{d}t + \int_{\pi}^{2\pi} -\sin(t) \, \mathrm{d}t\right).$$

See Figure 110 for an illustration of this integral.

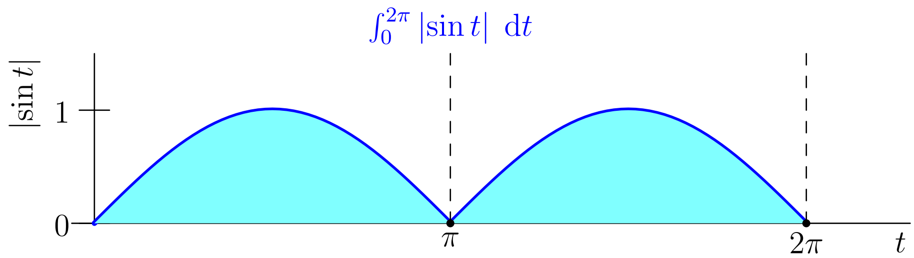

Figure 110: The integral $\int_0^{2\pi} |\sin(t)| \, dt$ is two copies of the first hump $\int_0^{\pi} \sin(t) \, dt$ (which doesn't have an absolute value on it).

Both integrals are the same, so we compute one and multiply by 2:

$$\int_0^{\pi} \sin(t) \, dt = [-\cos(t)]_0^{\pi} = -\cos(\pi) + \cos(0) = 1 + 1 = 2.$$

Thus, the total arc length is:

$$L = \sqrt{2} \cdot 2 \cdot 2 = 4\sqrt{2}.$$

§47.9 Solution to Exercise 20.2 (tangent to level curve)

Exercise 20.2. Let $k > 0$ be a fixed real number and let $f(x, y) = x^3 + ky^2$. Assume that the level curve of f for the value 21 passes through the point $P = (1, 2)$. Compute the equation of the tangent line to this level curve at the point P.

The first task is to recover the value of k which wasn't given in the statement. First, substitute the point $(1, 2)$ into the function $f(x, y)$:

$$f(1, 2) = 1^3 + k(2^2) = 1 + 4k.$$

We are told that $f(1, 2) = 21$, so we set the equation equal to 21:

$$1 + 4k = 21 \implies 4k = 20 \implies k = 5.$$

Thus, the function is:

$$f(x, y) = x^3 + 5y^2.$$

Now that we know f, we can compute the gradient by taking the partial derivatives:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 \\ 10y \end{pmatrix}.$$

Now evaluate the gradient at $P = (1, 2)$:

$$\nabla f(1, 2) = \begin{pmatrix} 3(1)^2 \\ 10(2) \end{pmatrix} = \begin{pmatrix} 3 \\ 20 \end{pmatrix}.$$

The gradient is always normal to the tangent line, so the tangent line must be of the form

$$3x + 20y = t$$

for some number t. This line passes through $(1, 2)$ so we can get

$$t = 3 \cdot 1 + 20 \cdot 2 = 43.$$

Hence the line requested is

$$3x + 20y = 43.$$

§47.10 Solution to Exercise 20.3 (approximation of $f = x^{5y}$)

Exercise 20.3. Let $f(x, y) = x^{5y}$ for $x, y > 0$. Use linear approximation to estimate $f(1.001, 3.001)$ starting from the point $(1, 3)$.

We are given the function:

$$f(x, y) = x^{5y}$$

and are asked to estimate $f(1.001, 3.001)$ using linear approximation, starting from the point $(1, 3)$, at which

$$f(1, 3) = 1.$$

We start by computing ∇f.

- To get the partial derivative with respect to x, use the power rule and chain rule:
$$\frac{\partial f}{\partial x} = 5yx^{5y-1}.$$
- For the partial derivative with respect to y, we treat x as a constant:
$$\frac{\partial f}{\partial y} = x^{5y}\log(x) \cdot 5.$$

Thus, the gradient of $f(x, y)$ is:

$$\nabla f(x, y) = \langle 5yx^{5y-1}, 5x^{5y}\log(x)\rangle.$$

The gradient at $(1, 3)$ is thus

$$\nabla f(1, 3) = \begin{pmatrix} 15 \\ 0 \end{pmatrix}.$$

The linear approximation of $f(1.001, y)$ near the point $(1, 3)$ can be expressed in terms of the gradient dot the displacement:

$$f(1.001, 3.001) \approx f(1, 3) + \nabla f(1, 3) \cdot \begin{pmatrix} 0.001 \\ 0.001 \end{pmatrix} = 1 + \begin{pmatrix} 15 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0.001 \\ 0.001 \end{pmatrix} = 1.015.$$

§47.11 Solution to Exercise 20.4 (cosine-quartic critical points)

Exercise 20.4. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = \cos(\pi x) + \frac{y^4}{4} - \frac{y^3}{3} - y^2.$$

- Compute all the critical points and classify them as saddle point, local minimum, or local maximum.
- Compute the global minimums and global maximums of f, if they exist.

Although this is stated as an 18.02 problem, it can actually be solved basically only using 18.01 methods. We'll still present the solution from an 18.02 perspective, but we'll comment many times on places where just 18.01 methods would have been sufficient.

Finding the critical points

To find the critical points, we first compute the gradient. The partial derivatives are

$$f_x(x,y) = \frac{\partial}{\partial x}\left(\cos(\pi x) + \frac{y^4}{4} - \frac{y^3}{3} - y^2\right) = -\pi\sin(\pi x).$$
$$f_y(x,y) = \frac{\partial}{\partial y}\left(\cos(\pi x) + \frac{y^4}{4} - \frac{y^3}{3} - y^2\right) = y^3 - y^2 - 2y.$$

Hence

$$\nabla f(x,y) = \begin{pmatrix} -\pi\sin(\pi x) \\ y^3 - y^2 - 2y \end{pmatrix}.$$

Setting this equal to $\mathbf{0}$ lets us solve each equation individually:

- $-\pi\sin(\pi x) = 0$ is true whenever x is an integer.
- To solve $y^3 - y^2 - 2y = 0$, factor the equation:

  $$0 = y(y^2 - y - 2) = y(y-2)(y+1) = 0.$$

So there are infinitely many critical points! The critical points occur when x is any integer and $y = -1$, $y = 0$, $y = 2$. See Figure 111.

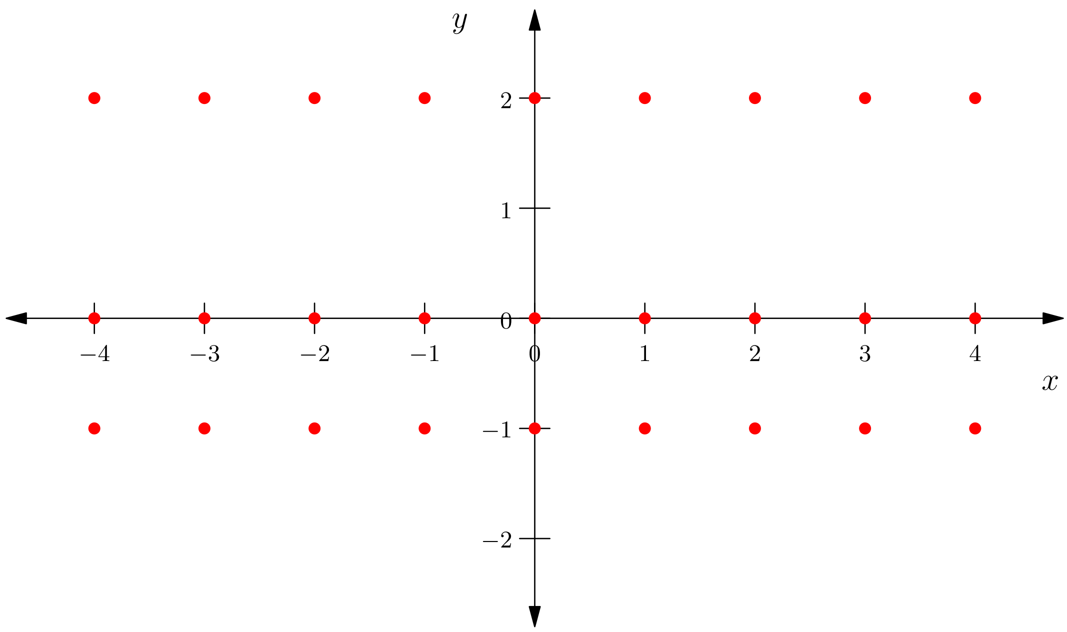

Figure 111: Plot of the critical points of the function in the xy-plane as red dots.

Classification using second derivative test

We now classify each of the points using the second derivative test. Calculate the second derivatives needed:

$$
\begin{aligned}
A &= f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = -\pi^2 \cos(\pi x), \\
B &= f_{xy}(x,y) = \frac{\partial^2 f}{\partial x \partial y} = 0, \\
C &= f_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = 3y^2 - 2y - 2.
\end{aligned}
$$

- We have $A = -\pi^2$ if x is odd and $A = \pi^2$ if x is even.
- We always have $B = 0$.
- We have

$$
C = \begin{cases} 3(-1)^2 - 2(-1) - 2 = 3 & \text{if } y = -1 \\ 3(0)^2 - 2(0) - 2 \quad = -2 & \text{if } y = 0 \\ 3(2)^2 - 2(2) - 2 \quad = 6 & \text{if } y = 2. \end{cases}
$$

We summarize all six cases in the table below. For each entry in the table we also compute $AC - B^2$ and then specify the answer based on the second derivative test.

	$x = \dots, -4, -2, 0, 2, 4\dots$ **is even**	$x = \dots, -3, -1, 1, 3\dots$ **is odd**
$y = -1$	$(A,B,C) = (-\pi^2, 0, 3)$ $AC - B^2 = -3\pi^2 < 0$ gives **saddle pt**	$(A,B,C) = (\pi^2, 0, 3)$ $AC - B^2 = 3\pi^2 > 0$ gives **local min**
$y = 0$	$(A,B,C) = (-\pi^2, 0, -2)$ $AC - B^2 = 2\pi^2 > 0$ gives **local max**	$(A,B,C) = (\pi^2, 0, -2)$ $AC - B^2 = -2\pi^2 < 0$ gives **saddle pt**
$y = 2$	$(A,B,C) = (-\pi^2, 0, 6)$ $AC - B^2 = -6\pi^2 < 0$ gives **saddle pt**	$(A,B,C) = (\pi^2, 0, 6)$ $AC - B^2 = 6\pi^2 > 0$ gives **local min**

Another approach without the second derivative test

You can get the same classification by just looking at the given function too. The point is that the function splits nicely into two halves: if define the one-variable functions

$$
\begin{aligned}
a(x) &:= \cos(\pi x) \\
b(y) &:= \frac{y^4}{4} - \frac{y^3}{3} - y^2
\end{aligned}
$$

then

$$
f(x,y) = a(x) + b(y).
$$

In that case, the following result is true:

- A point $P = (x,y)$ is a critical point of $f(x,y)$ if x is a critical point of $a(x)$ and y is a critical point of $b(y)$.
- If so then, the point P is...
  - a local minimum of f if x is a local minimum of $a(x)$ and y is a local minimum of $b(y)$.
  - a local maximum of f if x is a local maximum of $a(x)$ and y is a local maximum of $b(y)$.
  - a saddle point otherwise.

If you have a good conceptual understanding of saddle points, this should be obvious. It's essentially Figure 38 from Section 17.3.

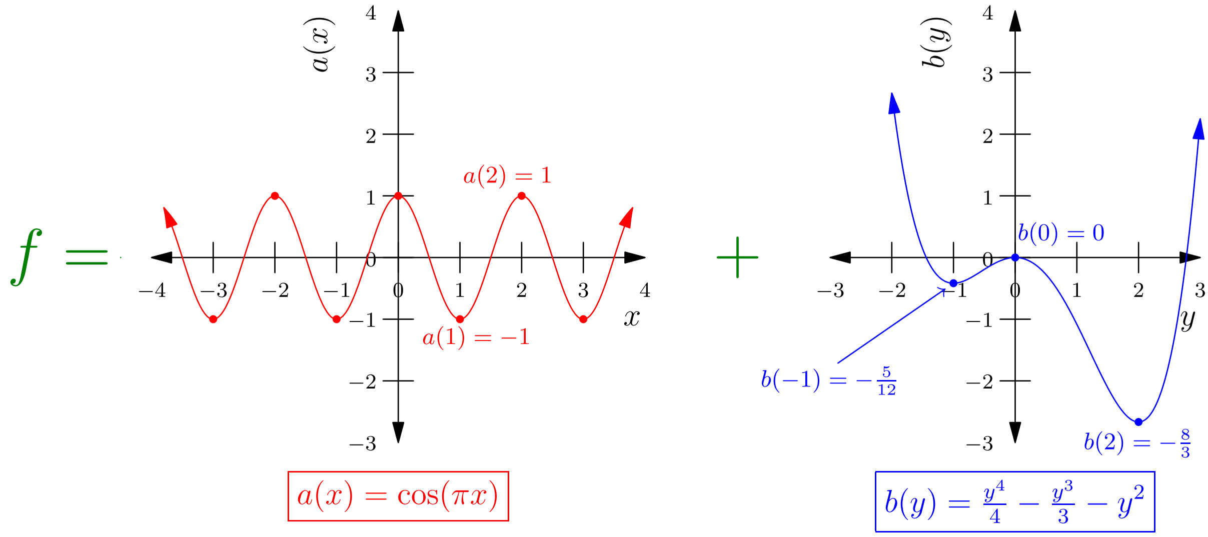

Figure 112: The function f is just the sum of two independent functions, which can be optimized independently.

This gives us the same table as above, since:

- The critical points of $a(x) = \cos(\pi x)$ are $x = -2, -1, 0, 1, 2, \ldots$. The minimums are the odd integers when the cos value reaches -1, the maximums are the even integers when the cos value reaches $+1$.
- The critical points of $b(y) = \frac{y^4}{4} - \frac{y^3}{3} - y^2$ are the roots of $b'(y) = y^3 - y^2 - 2y = y(y+1)(y-2)$, which are the same $y = -1, 0, 2$ we saw before. See Figure 112. There are local minimums at $y = -1$ and $y = 2$ and a local maximum at $y = 0$.

The global minimums and maximums

First, we evaluate f on every critical point. This is easiest to do if we use the a and b notation from before and compute

$$\begin{aligned} a(\text{even}) &= \cos(\pi \cdot \text{even}) = 1 \\ a(\text{odd}) &= \cos(\pi \cdot \text{odd}) = -1 \\ b(-1) &= \frac{(-1)^4}{4} - \frac{(-1)^3}{3} - (-1)^2 = -\frac{5}{12} \\ b(0) &= \frac{(0)^4}{4} - \frac{(0)^3}{3} - (0)^2 = 0 \\ b(2) &= \frac{(2)^4}{4} - \frac{(2)^3}{3} - (2)^2 = -\frac{8}{3} \end{aligned}$$

Then we get the six values shown in Table 25.

	$x = \ldots, -4, -2, 0, 2, 4\ldots$ **is even**	$x = \ldots, -3, -1, 1, 3\ldots$ **is odd**
$y = -1$	$f(\text{even}, -1) = 1 - \frac{5}{12} = \frac{7}{12}$	$f(\text{odd}, -1) = -1 - \frac{5}{12} = -\frac{17}{12}$
$y = 0$	$f(\text{even}, 0) = 1 + 0 = 1$	$f(\text{odd}, 0) = -1 + 0 = -1$
$y = 2$	$f(\text{even}, 2) = 1 - \frac{8}{3} = -\frac{5}{3}$	$f(\text{odd}, 2) = -1 - \frac{8}{3} = -\frac{11}{3}$

Table 25: Values of f at the critical points

There are no inequality constraints at all, so we just think about limit cases $x \to \pm\infty$ or $y \to \pm\infty$.

When $y \to \pm\infty$, the quartic $b(y) = \frac{y^4}{4} - \frac{y^3}{3} - y^2$ explodes to infinity. This implies already there cannot be any global maximum.

In the case where $x \to \pm\infty$, the cosine term of $f(x, y)$ will oscillate between -1 and 1, with period 2π. So there are no new smaller values of f that can be obtained here.

Another way to see the global minimums and maximums

Because

$$f(x, y) = a(x) + b(y)$$

the global minimum of f should be the sum of the global minimums of a and b, and likewise the global maximum of f should be the sum of the global maximums of a and b. So we could have also just used 18.01 methods on a and b individually, as in Figure 112. That is:

- Because $\min a(x) = -1$ and $\min b(y) = -\frac{8}{3}$, the global minimum is $-\frac{11}{3}$.
- Because $\max a(x) = 1$ and $\min b(y) = +\infty$, there is no global maximum.

Remember, this only works because we could easily divorce $f(x, y)$ into a function in x plus a function in y. For most functions $f(x, y)$ like xy or $e^x \sin(y)$, this approach is not going to fly.

§47.12 Solution to Exercise 20.5 (LM practice)

Exercise 20.5. Compute the minimum and maximum possible value of $x + 2y + 2z$ over real numbers x, y, z satisfying $x^2 + y^2 + z^2 \le 100$.

Let $f(x, y, z) = x + 2y + 2z$. Let $\mathcal{R}$ denote the region $x^2 + y^2 + z^2 \le 100$ (a ball of radius 10) and let $\mathcal{S}$ denote the boundary $x^2 + y^2 + z^2 = 100$ (a sphere of radius 10). We follow the steps we described in the recipe in Section 19.2 and Section 19.4.

0. $\mathcal{R}$ is three-dimensional and has no limit cases but a two-dimensional boundary $\mathcal{S}$. (Because of the condition $x^2 + y^2 + z^2 \le 100$ and all the squares being nonnegative, none of the variables can go to $\pm\infty$.)

1. We calculate all the critical points of the objective function $f(x, y, z) = x + 2y + 2z$. The gradient is

$$\nabla f = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

   So there are no critical points, because this gradient is never 0.

2. The boundary in $\mathcal{S}$ is a sphere, and it cannot easily be handled. We pull out Lagrange multipliers and follow the recipe all the way through again.

   0. $\mathcal{S}$ is two-dimensional and has no limit cases or boundary.

   1. We search for LM-critical points by letting $g(x, y, z) = x^2 + y^2 + z^2$, so $\mathcal{S}$ is the level surface $g = 100$. Calculate the gradient of g:

$$\nabla g = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}.$$

      Recall that an LM-critical point is one for which $g(P) = 100$ and *either*

$$\nabla f(P) = \lambda \nabla g(P) \quad \text{OR} \quad \nabla(P) = 0.$$

This gradient ∇g could be $\mathbf{0}$ at $(x, y, z) = (0, 0, 0)$, but this point does not lie on $\mathcal{S}$, so we disregard it.

In the main case $\nabla f = \lambda g$, we seek points such that

$$\begin{aligned} 1 &= \lambda \cdot 2x \\ 2 &= \lambda \cdot 2y \\ 2 &= \lambda \cdot 2z. \end{aligned}$$

Our strategy is to kill every variable *except* λ, by writing

$$\begin{aligned} x &= \frac{1}{2\lambda} \\ y &= \frac{1}{\lambda} \\ z &= \frac{1}{\lambda}. \end{aligned}$$

Plugging this back into the constraint equation $x^2 + y^2 + z^2 = 100$ and simplifying gives

$$\begin{aligned} \left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 &= 100 \\ \iff \frac{9}{4\lambda^2} &= 100 \\ \iff \lambda^2 &= \frac{9}{400} \\ \iff \lambda &= \pm\frac{3}{20}. \end{aligned}$$

Putting these two values of λ in gives $(x, y, z) = \left(\frac{10}{l}, \frac{20}{3}, \frac{20}{3}\right)$ and $(x, y, z) = \left(-\frac{10}{3}, -\frac{20}{3}, -\frac{20}{3}\right)$ These are the two LM-critical points. Evaluating this gives

$$\begin{aligned} f\left(\frac{10}{3}, \frac{20}{3}, \frac{20}{3}\right) &= \frac{10}{3} + 2 \cdot \frac{20}{3} + 2 \cdot \frac{20}{3} = 30 \\ f\left(-\frac{10}{3}, -\frac{20}{3}, -\frac{20}{3}\right) &= -\frac{10}{3} + 2 \cdot -\frac{20}{3} + 2 \cdot -\frac{20}{3} = -30. \end{aligned}$$

2. There are no boundary cases to consider for $\mathcal{S}$.
3. There are no limit cases to consider for $\mathcal{S}$.

In conclusion, the maximum value is 30 and the minimum value is -30, at the points $\left(\frac{10}{3}, \frac{20}{3}, \frac{20}{3}\right)$ and $\left(-\frac{10}{3}, -\frac{20}{3}, -\frac{20}{3}\right)$ we found earlier.

3. There are no limit cases to consider for $\mathcal{R}$.

Digression

Note that in fact one can note *a priori* that any maximum or minimum should occur on the sphere. One way to see this is that if one takes a point strictly inside $\mathcal{R}$ like $P = (6, 8, 0)$, one can always increase the absolute value of f by scaling P until it lies on the sphere (e.g. $(60, 80, 0)$). Hence there is no loss of generality in assuming maximums and minimums lie on $\mathcal{S}$. So if one is observant enough they can skip straight to the LM on $\mathcal{S}$, ignoring the region $\mathcal{R}$ entirely.

§47.13 Solution to Exercise 20.6 (tangent plane)

Exercise 20.6. Consider the level surface of $f(x, y, z) = (x-1)^2 + (y-1)^3 + (z-1)^4$ that passes through the origin $O = (0, 0, 0)$. Let $\mathcal{H}$ denote the tangent plane to this surface at O. Give an example of two nonzero tangent vectors to this surface at O whose span is $\mathcal{H}$.

The gradient of the function f

$$\nabla f = \begin{pmatrix} 2(x-1) \\ 3(y-1)^2 \\ 4(z-1)^3 \end{pmatrix}$$

and so the gradient at the origin is

$$\nabla f(0,0,0) = \begin{pmatrix} -2 \\ 3 \\ -4 \end{pmatrix}.$$

The tangent plane $\mathcal{H}$ consists of those vectors which are normal to $\left(\begin{smallmatrix} -2 \\ 3 \\ -4 \end{smallmatrix}\right)$. This plane is two-dimensional. So, to find two vectors spanning $\mathcal{H}$, according to the "buy two get one free" result from we just need to give any two linearly independent (i.e. not multiples of each other) vectors which are both perpendicular to $\left(\begin{smallmatrix} -2 \\ 3 \\ -4 \end{smallmatrix}\right)$.

There are many valid choices. One such example might be $\left(\begin{smallmatrix} 3 \\ 2 \\ 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 \\ 4 \\ 3 \end{smallmatrix}\right)$. These two vectors are clearly not multiples of each other, and

$$\begin{pmatrix} -2 \\ 3 \\ -4 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = (-2) \cdot 3 + 3 \cdot 2 + (-4) \cdot 0 = 0$$

$$\begin{pmatrix} -2 \\ 3 \\ -4 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 4 \\ 3 \end{pmatrix} = (-2) \cdot 0 + 3 \cdot 4 + (-4) \cdot 3 = 0$$

so they are indeed tangent vectors contained in $\mathcal{H}$.

Chapter 48. Solutions to Part Golf

§48.1 Solution to Exercise 22.1 (practice with slicing)

Exercise 22.1. Let $\mathscr{R}$ be the region between the curves $y = \sqrt{x}$ and $y = x^3$. Compute $\iint_{\mathscr{R}} x^{100}y^{200}\,\mathrm{d}x\,\mathrm{d}y$ using both horizontal and vertical slicing.

Let $\mathscr{R}$ be the region bounded by the curves $y = \sqrt{x}$ and $y = x^3$. We wish to compute the integral

$$I = \iint_{\mathscr{R}} x^{100}y^{200}\,\mathrm{d}x\,\mathrm{d}y.$$

To determine the limits of integration, we find the intersection points by solving $\sqrt{x} = x^3$ Squaring both sides gives $x = x^6 \implies x(x^5 - 1) = 0$, so $x = 0$ or $x = 1$. That is, the intersection points are $(0, 0)$ and $(1, 1)$. A sketch of the region is shown in Figure 113.

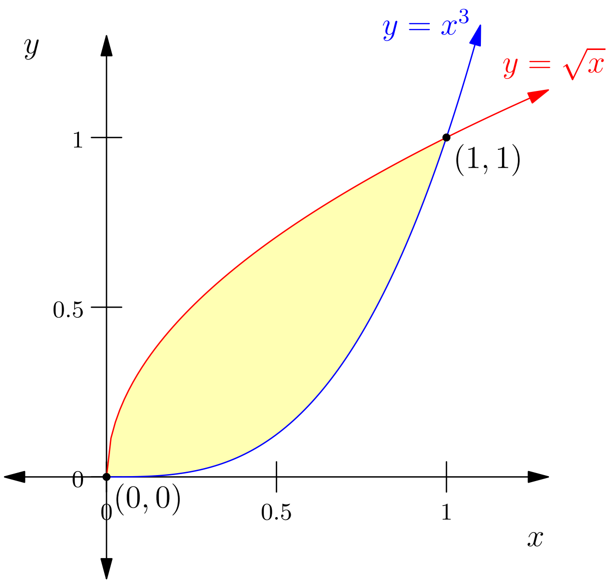

Figure 113: The region between $y = \sqrt{x}$ and $y = x^3$.

From the figure, we can describe the region $\mathscr{R}$ with the inequalities

$$0 \le x^3 \le y \le \sqrt{x} \le 1.$$

Integrating with x inside and y outside

The values of y go from 0 to 1. For each fixed y, the values of x range from

$$y^2 \le x \le y^{1/3}$$

so the integral is:

$$I = \int_{y=0}^{1} \int_{x=y^2}^{y^{1/3}} x^{100}y^{200}\,\mathrm{d}x\,\mathrm{d}y.$$

Evaluating the inner integral:

$$\int_{x=y^2}^{y^{1/3}} x^{100}\,\mathrm{d}x = \left[\frac{x^{101}}{101}\right]_{x=y^2}^{y^{1/3}} = \frac{y^{101/3}}{101} - \frac{y^{202}}{101}.$$

Now, integrating over y:

$$\begin{aligned} I &= \int_{y=0}^{1} \left(\frac{y^{101/3}}{101} - \frac{y^{202}}{101} \right) y^{200} \, dy \\ &= \frac{1}{101} \left(\int_{y=0}^{1} y^{701/3} \, dy - \int_{y=0}^{1} y^{402} \, dy \right) \\ &= \frac{1}{101} \left(\frac{3}{704} - \frac{1}{403} \right) = \boxed{\frac{5}{283712}}. \end{aligned}$$

Integrating with y inside and x outside

The values of x go from 0 to 1. For a fixed x, the values of y range from

$$x^3 \le y \le x^{1/2}$$

so the integral is

$$I = \int_{x=0}^{1} \int_{y=x^3}^{x^{1/2}} x^{100} y^{200} \, dy \, dx.$$

Evaluating the inner integral:

$$\int_{y=x^3}^{\sqrt{x}} y^{200} \, dy = \left[\frac{y^{201}}{201} \right]_{y=x^3}^{y=\sqrt{x}} = \frac{x^{201/2}}{201} - \frac{x^{603}}{201}.$$

Now, integrating over x:

$$\begin{aligned} I &= \int_{x=0}^{1} x^{100} \left(\frac{x^{201/2}}{201} - \frac{x^{603}}{201} \right) dx \\ &= \frac{1}{201} \left(\int_{x=0}^{1} x^{4012} \, dx - \int_{x=0}^{1} x^{703} \, dx \right) \\ &= \frac{1}{201} \left(\frac{2}{403} - \frac{1}{704} \right) = \boxed{\frac{5}{283712}}. \end{aligned}$$

§48.2 Solution to Exercise 22.2 (center of mass of a region)

Exercise 22.2. Let $\mathcal{R}$ be the region between the curves $y = \sqrt{x}$ and $y = x^2$. Assume $\mathcal{R}$ has constant density. Calculate its center of mass.

The region is really similar to the one in the preceding exercise, and can be described as

$$0 \le x^2 \le y \le \sqrt{x} \le 1$$

for the same reason, as shown in Figure 114. (It's exactly the same as the last exercise except x^3 was changed to x^2, so one just replaces all the 3's with 2's.) Note that the region is symmetric around the line $y = x$, so *a priori* we should expect our answer to lie on $y = x$ as well.

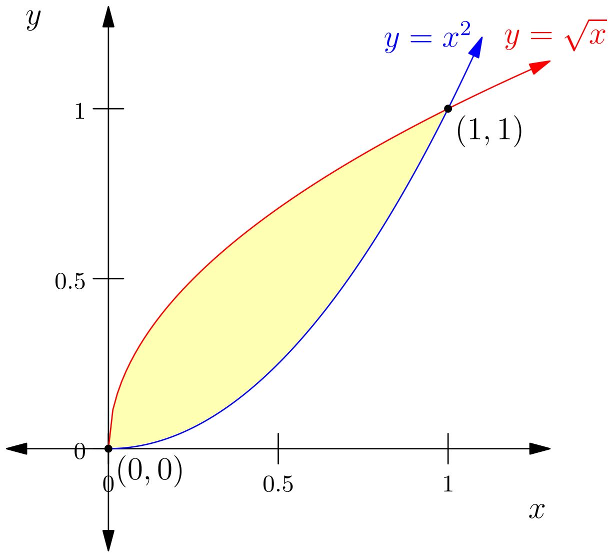

Figure 114: The region between $y = \sqrt{x}$ and $y = x^2$.

First, to compute the area of $\mathcal{R}$, we can write

$$\operatorname{Area}(\mathcal{R}) = \int_{x=0}^{1} \int_{y=x^2}^{\sqrt{x}} 1 \, \mathrm{d}y \, \mathrm{d}x = \int_{x=0}^{1} \left(\sqrt{x} - x^2\right) \mathrm{d}x = \left[\frac{2}{3}x^{3/2} - \frac{1}{3}x^3\right]_{x=0}^{1} = \frac{1}{3}.$$

The x-coordinate of the center of mass is therefore given by

$$\begin{aligned} \bar{x} = \frac{1}{\operatorname{Area}(\mathcal{R})} \int_{\mathcal{R}} x \, \mathrm{d}A &= \frac{1}{1/3} \int_{x=0}^{1} \int_{y=x^2}^{\sqrt{x}} x \, \mathrm{d}y \, \mathrm{d}x \\ &= 3 \int_{x=0}^{1} x\left(\sqrt{x} - x^2\right) \mathrm{d}x \\ &= 3\left[\frac{2}{5}x^{\frac{5}{2}} - \frac{1}{4}x^4\right]_{x=0}^{1} = \frac{9}{20}. \end{aligned}$$

As for the y-coordinate, we expect $\bar{y} = \bar{x}$ from the symmetry of the region, and indeed

$$\begin{aligned} \bar{y} = \frac{1}{\operatorname{Area}(\mathcal{R})} \int_{\mathcal{R}} y \, \mathrm{d}A &= \frac{1}{1/3} \int_{x=0}^{1} \int_{y=x^2}^{\sqrt{x}} y \, \mathrm{d}y \, \mathrm{d}x \\ &= 3 \int_{x=0}^{1} \left[\frac{y^2}{2}\right]_{y=x^2}^{\sqrt{x}} \mathrm{d}x \\ &= \frac{3}{2} \int_{x=0}^{1} \left(x - x^4\right) \mathrm{d}x \\ &= \frac{3}{2}\left[\frac{1}{2}x^2 - \frac{1}{5}x^5\right]_{x=0}^{1} = \frac{9}{20}. \end{aligned}$$

Thus, the center of mass of the region is:

$$\boxed{\left(\frac{9}{20}, \frac{9}{20}\right)}.$$

§48.3 Solution to Exercise 22.3 (double integral with 5th root)

Exercise 22.3. Evaluate the double integral:

$$\int_{y=0}^{1} \int_{x=y}^{\sqrt[5]{y}} \frac{xy^2}{1-x^{12}} \,\mathrm{d}x \,\mathrm{d}y.$$

Writing as a region, this is

$$\mathcal{R} = \begin{cases} 0 \le y \le 1 \\ y \le x \le \sqrt[5]{y} \end{cases}.$$

The values of x could range anywhere in $0 \le x \le 1$. For a fixed x, the value y needs to satisfy four conditions: $0 \le y \le 1$ and also $x^5 \le y \le x$. But in fact

$$0 \le x^5 \le y \le x \le 1$$

so we can compress this to just:

$$\mathcal{R} = \begin{cases} 0 \le x \le 1 \\ x^5 \le y \le x \end{cases}.$$

Thus, the new limits of integration become:

$$\int_{x=0}^{1} \int_{y=x^5}^{x} \frac{xy^2}{1-x^{12}} \,\mathrm{d}y \,\mathrm{d}x$$

We now compute the inner integral with respect to y:

$$\int_{y=x^5}^{x} y^2 \,\mathrm{d}y = \left[\frac{y^3}{3} \right]_{y=x^5}^{y=x}$$

Substituting the limits of integration:

$$\frac{1}{3}\left(x^3 - \left(x^5\right)^3\right) = \frac{1}{3}(x^3 - x^{15})$$

Now substitute this result into the outer integral:

$$\int_{x=0}^{1} \frac{x}{1-x^{12}} \cdot \frac{1}{3}(x^3 - x^{15}) \,\mathrm{d}x$$

Simplifying:

$$\begin{aligned} \frac{1}{3}\int_{x=0}^{1} \frac{x}{1-x^{12}}(x^3 - x^{15}) \,\mathrm{d}x &= \frac{1}{3}\int_{x=0}^{1} \frac{x^4 - x^{16}}{1-x^{12}} \,\mathrm{d}x \\ &= \frac{1}{3}\int_{x=0}^{1} x^4 \,\mathrm{d}x \\ &= \frac{1}{3}\left(\frac{1}{5} - 0\right) = \frac{1}{15}. \end{aligned}$$

§48.4 Solution to Exercise 22.4 (rational integral)

Exercise 22.4 (*). Prove that

$$\int_{x=0}^{999^5} \sqrt[3]{\sqrt[5]{x}+1}$$

is a rational number.

For brevity, let $N := 999^5$.

At face value, this looks like an 18.01 integral, but we know from 18.01 that this integral is actually measuring the area under some curve. The idea is that, to avoid having to deal with the hideous roots, we are going to use horizontal slicing for the region under the curve shown in the figure.

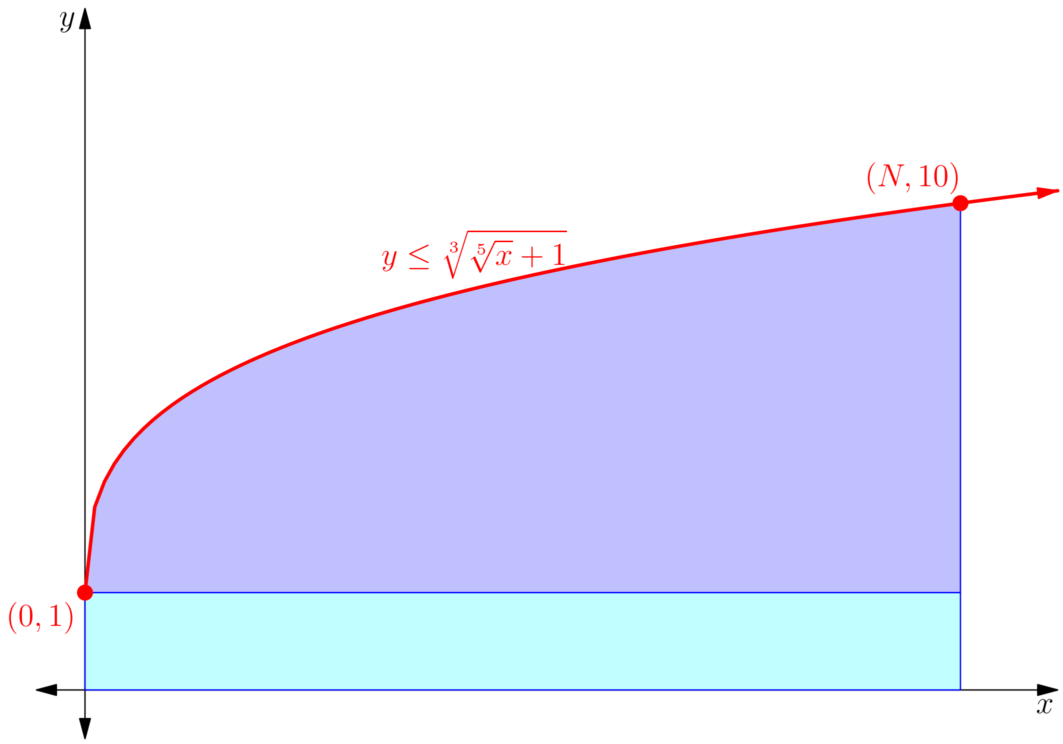

Figure 115: The region $0 \le y \le \sqrt[5]{\sqrt[3]{x}+1}$ for $0 \le x \le N$. Not at all to scale.

Let's first convert the region into inequality format: we have $0 \le x \le N$, $y \ge 0$ and

$$y \le \sqrt[3]{\sqrt[5]{x}+1} \iff x \ge \left(y^3-1\right)^5.$$

Hence, the area under the curve can be split into two parts. In the range $0 \le y \le 1$ we get the light blue rectangle shown above (bottom half of figure), which goes from $0 \le x \le N$ to $0 \le y \le 1$, and has area N. Then from $1 \le y \le 10$ the bounds on x are instead given by

$$\left(y^3-1\right)^5 \le x \le N.$$

This is the dark blue region (top half of figure) and it has area

$$\int_{y=1}^{2} \int_{x=(y^3-1)^5}^{N} 1 \, \mathrm{d}x \, \mathrm{d}y = \int_{y=1}^{N} \left(N - \left(y^3-1\right)^5\right) \mathrm{d}y.$$

The total area is thus

$$N + \int_{y=1}^{10} \left(N - (y^3 - 1)^5\right) \,\mathrm{d}y = N + N(N-1) - \int_{y=1}^{10} (y^3 - 1)^5 \,\mathrm{d}y.$$

This is a easily seen to be a rational number.

Remark

Using a calculator, one could explicitly compute

$$\int_{y=1}^{10} (y^3 - 1)^5 \,\mathrm{d}y = \int_{y=1}^{10} (y^{15} - 5y^{12} + 10y^9 - 10y^6 + 5y^3 - 1) \,\mathrm{d}y = \frac{904414539218186169}{1456}$$

if one is so inclined.

Digression

It is possible to evaluate the integral using 18.01 methods by making the u-substitution $u = \sqrt{\sqrt{x} + 1}$, but this is extremely tedious.

§48.5 Solution to Exercise 23.1 (integral over triangle)

Exercise 23.1. Let $\mathcal{R}$ be all the points on or inside the triangle with vertices $(0,0)$, $(1,2)$ and $(2,1)$. Compute

$$\iint_{\mathcal{R}} \frac{(x+y)^2}{xy} \,\mathrm{d}x \,\mathrm{d}y.$$

(Recommended approach: use change of variables with $u = x + y$ and $v = \frac{x}{y}$.)

We use the transformation:

$$u = x + y, \quad v = \frac{x}{y}.$$

The region under (u, v) coordinates can be expressed as

$$0 \le u \le 3, \quad \frac{1}{2} \le v \le 2.$$

This is drawn in Figure 116.

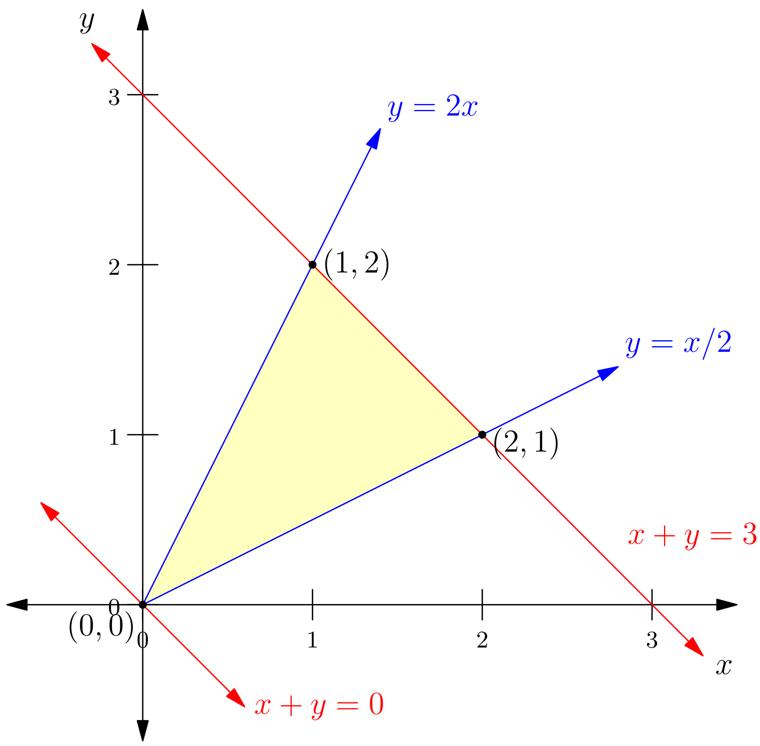

Figure 116: The triangle with vertices $(0, 0)$, $(1, 2)$, and $(2, 1)$, with a change of variables suggested using $u = x + y$ and $v = \frac{x}{y}$

This is a case where we want to use the inverse Jacobian

$$\det(J_{\mathbf{T}^{-1}}) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ \frac{1}{y} & -\frac{x}{y^2} \end{vmatrix} = -\frac{x+y}{y^2}.$$

So

$$\frac{1}{|\det(J_{\mathbf{T}^{-1}})|} = \frac{y^2}{x+y}.$$

Hence, the transformed integral becomes

$$\begin{aligned} \int_{u=0}^{3} \int_{v=\frac{1}{2}}^{2} \frac{(x+y)^2}{xy} \cdot \frac{y^2}{x+y} \, \mathrm{d}v \, \mathrm{d}u &= \int_{u=0}^{3} \int_{v=\frac{1}{2}}^{2} \frac{y}{x} \cdot (x+y) \, \mathrm{d}v \, \mathrm{d}u \\ &= \int_{u=0}^{3} \int_{v=\frac{1}{2}}^{2} \frac{1}{v} \cdot u \, \mathrm{d}v \, \mathrm{d}u \\ &= \left(\int_{u=0}^{3} u \, \mathrm{d}u \right) \left(\int_{v=\frac{1}{2}}^{2} \frac{1}{v} \, \mathrm{d}v \right) \\ &= \left[\frac{u^2}{2} \right]_{u=0}^{3} \cdot [\log v]_{v=\frac{1}{2}}^{2} \\ &= \frac{9}{2} \cdot \left(\log 2 - \log\left(\frac{1}{2}\right) \right) = \boxed{9 \log 2}. \end{aligned}$$

§48.6 Solution to Exercise 24.1 (polar integral 1)

Exercise 24.1. Compute

$$\int_{x=0}^{1} \int_{y=0}^{\sqrt{1-x^2}} xy \, dy \, dx.$$

The limits of integration describe the region bounded by:

- $0 \le x \le 1$,
- $0 \le y \le \sqrt{1-x^2}$.

This corresponds to the quarter-circle in the first quadrant of the unit disk, given by $x^2 + y^2 \le 1$ with $x \ge 0$.

Using the polar coordinate transformations:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad dx\,dy = r\,dr\,d\theta.$$

The given region corresponds to:

- $0 \le r \le 1$,
- $0 \le \theta \le \frac{\pi}{2}$.

Rewriting the integrand:

$$xy = (r\cos\theta)(r\sin\theta) = r^2\cos\theta\sin\theta.$$

Hence the integral transforms into:

$$\begin{aligned} I &= \int_{\theta=0}^{\pi/2} \int_{r=0}^{1} r^2 \cos\theta \sin\theta \cdot r \, dr \, d\theta \\ &= \left(\int_{r=0}^{1} r^3 \, dr \right) \left(\int_{\theta=0}^{\pi/2} \frac{\sin(2\theta)}{2} \, d\theta \right) \\ &= \left[\frac{r^4}{4} \right]_{r=0}^{1} \left[-\frac{\cos(2\theta)}{4} \right]_{\theta=0}^{\pi/2} \\ &= \frac{1}{4} \cdot \frac{1}{2} = \boxed{\frac{1}{8}}. \end{aligned}$$

§48.7 Solution to Exercise 24.2 (polar integral 2)

Exercise 24.2. Compute

$$\iint_{(x-1)^2+y^2 \le 1} \frac{1}{\sqrt{x^2+y^2}} \, dx \, dy.$$

In Section 24.3, we have already established that the given region in polar coordinates is described by:

$$0 \le r \le 2\cos\theta, \quad -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}.$$

Using the standard polar transformations:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad dx\,dy = r\,dr\,d\theta,$$

the given integrand is just r.

$$\frac{1}{\sqrt{x^2+y^2}} = \frac{1}{\sqrt{r^2}} = \frac{1}{r}.$$

Thus, the integral becomes:

$$\begin{aligned} I &= \int_{\theta=-\pi/2}^{\pi/2} \int_{r=0}^{2\cos\theta} \frac{1}{r} \cdot r \, \mathrm{d}r \, \mathrm{d}\theta \\ &= \int_{\theta=-\pi/2}^{\pi/2} \int_{r=0}^{2\cos\theta} \mathrm{d}r \, \mathrm{d}\theta \\ &= \int_{\theta=-\pi/2}^{\pi/2} 2\cos\theta \, \mathrm{d}\theta \\ &= [2\sin\theta]_{\theta=-\pi/2}^{\pi/2} \\ &= \boxed{4}. \end{aligned}$$

§48.8 Solution to Exercise 24.3 (polar integral 3)

Exercise 24.3 (*). Compute

$$\iint_{x^2+y^2\le 1} \sqrt{\left(x+\frac{3}{5}\right)^2 + \left(y+\frac{4}{5}\right)^2} \, \mathrm{d}x \, \mathrm{d}y.$$

This is actually a disguised version of the example in Section 24.3! That is, the answer is also $\frac{32}{9}$.

To repeat, in Section 24.3 the example can be thought of as showing

$$I_1 = \iint_{(x-1)^2+y^2\le 1} \sqrt{x^2+y^2} \, \mathrm{d}x \, \mathrm{d}y = \frac{32}{9}.$$

Our goal is to argue that

$$I_2 = \iint_{x^2+y^2\le 1} \sqrt{\left(x+\frac{3}{5}\right)^2 + \left(y+\frac{4}{5}\right)^2} \, \mathrm{d}x \, \mathrm{d}y = \boxed{\frac{32}{9}}.$$

Note that:

- The first integral I_1 is taken over the disk centered at $(1,0)$ with radius 1. We call this disk $\mathcal{R}_1$.
- The second integral I_2 is taken over the unit disk centered at $(0,0)$. We call this disk $\mathcal{R}_2$.

Observe that the disks $\mathcal{R}_1$ and $\mathcal{R}_2$ are congruent. Moreover,

- The integrand in I_1 is $\sqrt{x^2+y^2}$. This measures the distance of each point in $\mathcal{R}_1$ from the origin $O=(0,0)$. Note that O is a point on the boundary of $\mathcal{R}_1$.
- The integrand in I_2 is $\sqrt{\left(x+\frac{3}{5}\right)^2 + \left(y+\frac{4}{5}\right)^2}$. This measures the distance of each point in $\mathcal{R}_2$ from the origin $P = \left(-\frac{3}{5}, -\frac{4}{5}\right)$. Note that P is a point on the boundary of $\mathcal{R}_2$.

See the illustration in Figure 117.

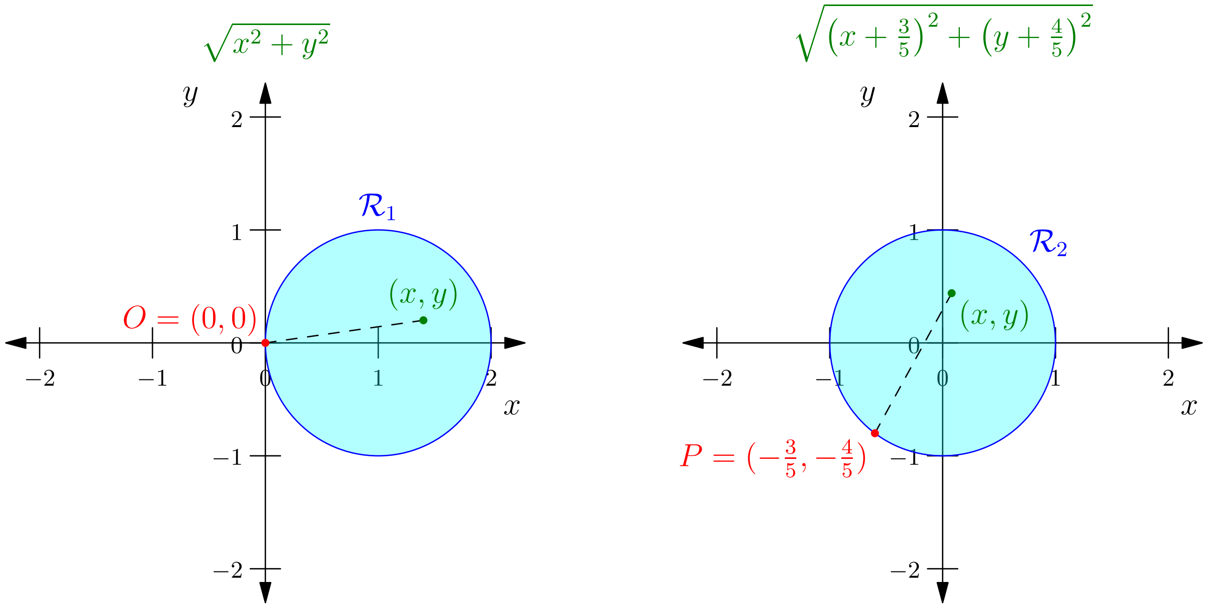

Figure 117: Illustration of the integrals I_1 and I_2 of from Section 24.3 and Exercise 24.3, showing they're computing the same thing.

Since both integrals effectively compute the same function over congruent regions, the results must be equal.

Chapter 49. Solutions to Part Hotel

§49.1 Solution to Exercise 26.1 (the napkin-ring problem)

Exercise 26.1 (Napkin-ring problem). Let $R > a > 0$ be given real numbers, and let $h := 2\sqrt{R^2 - a^2}$. A cylindrical hole of radius a is drilled through the center of a wooden ball of radius R to get a bead of height h, as shown in Figure 59. Compute the volume of the resulting bead as a function of h.

This is a famous exercise with its own page on Wikipedia at `https://w.wiki/CarU`. We give one solution using spherical coordinates below, but this is far from the only possible solution.

Using cylindrical coordinates the equation of the sphere is:

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

Since the hole is along the z-axis with radius a, we impose

$$a \le r \le R$$

that is, the bead consists of those points whose distance from the z-axis is at least a.

As for z, we require $r^2 + z^2 \le R^2$. That is, for each given r, the possible values of z are those with

$$-\sqrt{R^2 - r^2} \le z \le \sqrt{R^2 - r^2}.$$

The volume element in cylindrical coordinates is:

$$\mathrm{d}V = r \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}z.$$

Putting this all together, the volume integral is:

$$V = \int_{\theta=0}^{2\pi} \int_{r=a}^{R} \int_{z=-\sqrt{R^2-r^2}}^{\sqrt{R^2-r^2}} r \, \mathrm{d}z \, \mathrm{d}r \, \mathrm{d}\theta.$$

Evaluating the inner integral gives:

$$\int_{z=-\sqrt{R^2-r^2}}^{\sqrt{R^2-r^2}} \mathrm{d}z = 2\sqrt{R^2 - r^2}.$$

Thus, the volume integral simplifies to:

$$\begin{aligned} V &= \int_{\theta=0}^{2\pi} \int_{r=a}^{R} 2r\sqrt{R^2 - r^2} \, \mathrm{d}r \, \mathrm{d}\theta \\ &= \left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta \right) \left(\int_{r=a}^{R} 2r\sqrt{R^2 - r^2} \, \mathrm{d}r \right) \\ &= 2\pi \left(\int_{r=a}^{R} 2r\sqrt{R^2 - r^2} \, \mathrm{d}r \right). \end{aligned}$$

To evaluate the integral, we substitute $u = R^2 - r^2$, so that $\mathrm{d}u = -2r \, \mathrm{d}r$, we rewrite:

$$\begin{aligned}
\int_{r=a}^{R} 2r\sqrt{R^2-r^2}\,\mathrm{d}r &= \int_{u=R^2-R^2}^{R^2-a^2} \sqrt{u}(-\,\mathrm{d}u) \\
&= \int_{u=0}^{R^2-a^2} \sqrt{u}\,\mathrm{d}u \\
&= \left[\frac{2}{3}u^{3/2}\right]_{u=0}^{R^2-a^2} \\
&= \frac{2}{3}(R^2-a^2)^{\frac{3}{2}}.
\end{aligned}$$

Multiplying by 2π, we obtain:

$$V = 2\pi \cdot \frac{2}{3}(R^2-a^2)^{3/2} = \frac{4\pi}{3}(R^2-a^2)^{3/2}.$$

Since $h = 2\sqrt{R^2-a^2}$, we have: $R^2 - a^2 = \frac{h^2}{4}$, so

$$V = \boxed{\frac{\pi}{6}h^3}.$$

§49.2 Solution to Exercise 27.1 (average distance of sphere to line)

Exercise 27.1. Consider a solid ball of radius 1 and a line ℓ through its center. Across all points P inside the ball, compute the average value of the distance from P to ℓ. (The average is defined as $\frac{1}{\operatorname{Vol}(\mathcal{T})}\iiint_{\mathcal{T}} d(P)\,\mathrm{d}V$, where $d(P)$ is the distance from P to ℓ.)

We choose the z-axis to be aligned with the line ℓ. Hence the distance from a point $P = (r, \theta, \varphi)$ to ℓ is simply the perpendicular distance from P to the z-axis, which is:

$$d(P) = r\sin\varphi.$$

We integrate over the entire sphere with coordinates

$$0 \le r \le 1, \quad 0 \le \theta \le 2\pi, \quad 0 \le \varphi \le \pi.$$

The total volume of the ball is:

$$\operatorname{Vol}(\mathcal{T}) = \frac{4}{3}\pi(1)^3 = \frac{4\pi}{3}.$$

We now set up the integral $\iiint_{\mathcal{T}} d(P)\,\mathrm{d}V$. The volume element is:

$$\mathrm{d}V = r^2 \sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta.$$

Hence

$$\begin{aligned}
\iiint_{\mathcal{T}} d(P)\,\mathrm{d}V &= \iiint_{\mathcal{T}} r\sin\varphi \cdot r^2\sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi}\mathrm{d}\theta \int_{\varphi=0}^{\pi}\sin^2\varphi\,\mathrm{d}\varphi \int_{r=0}^{1} r^3\,\mathrm{d}r.
\end{aligned}$$

The center integral needs the following trig identity:

$$\begin{aligned}\sin^2\varphi &= \frac{1-\cos(2\varphi)}{2} \\ \implies \int \sin^2\varphi \, \mathrm{d}\varphi &= \frac{\varphi}{2} - \frac{\sin(2\varphi)}{4}.\end{aligned}$$

Hence,

$$\begin{aligned}\iiint_{\mathcal{T}} d(P)\,\mathrm{d}V &= \int_{\theta=0}^{2\pi} \mathrm{d}\theta \int_{\varphi=0}^{\pi} \sin^2\varphi\,\mathrm{d}\varphi \int_{r=0}^{1} r^3\,\mathrm{d}r = (2\pi)\cdot\left(\frac{\pi}{2}\right)\cdot\left(\frac{1}{4}\right) \\ &= \frac{\pi^2}{4}.\end{aligned}$$

So the final answer is

$$\frac{1}{\mathrm{Vol}(\mathcal{T})}\iiint_{\mathcal{T}} d(P)\,\mathrm{d}V = \frac{\pi^2/4}{4\pi/3} = \boxed{\frac{3}{16}\pi}.$$

§49.3 Solution to Exercise 27.2 (gravity on hemisphere)

Exercise 27.2. Suppose $\mathcal{T}$ is a solid metal hemisphere of radius 1 of constant unit density, and P is a point of mass m at the center of the base of the hemisphere. Calculate the magnitude of the force of gravity exerted on the point P.

We orient the hemisphere so it rests on the xy-plane with the point P at $(0,0,0)$ (so the hemisphere is $x^2+y^2+z^2 \le 1$ and $z \ge 0$). Then this is basically the same as the example in Section 27.5, except the bounds of integration change.

To be precise, we have $G_1 = G_2 = 0$, and Equation 17 and again

$$G_3 = Gm \iiint_{\mathcal{T}} \sin\varphi\cos\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta$$

after setting the density to 1. The only change is the bounds of integration: for the hemisphere we should have

$$0 \le \rho \le 1, \quad 0 \le \theta \le 2\pi, \quad 0 \le \phi \le \frac{\pi}{2}.$$

So, when we integrate to compute G_3 we have

$$\begin{aligned}G_3 &= Gm \int_{\theta=0}^{2\pi}\int_{\varphi=0}^{\frac{\pi}{2}}\int_{\rho=0}^{1} \sin\varphi\cos\varphi\,\mathrm{d}\rho\,\mathrm{d}\varphi\,\mathrm{d}\theta \\ &= Gm\left(\int_{\theta=0}^{2\pi}\mathrm{d}\theta\right)\left(\int_{\varphi=0}^{\frac{\pi}{2}}\sin\varphi\cos\varphi\,\mathrm{d}\varphi\right)\left(\int_{\rho=0}^{1}\mathrm{d}\rho\right) \\ &= Gm(2\pi)\left(\int_{\varphi=0}^{\frac{\pi}{2}}\sin\varphi\cos\varphi\,\mathrm{d}\varphi\right)(1).\end{aligned}$$

To evaluate the integral with φ, write

$$\int_{\varphi=0}^{\pi/2}\sin\varphi\cos\varphi\,\mathrm{d}\varphi = \frac{1}{2}\int_{0}^{\pi/2}\sin 2\varphi\,\mathrm{d}\varphi = \frac{1}{2}\left[-\frac{1}{2}\cos 2\varphi\right]_{\varphi=0}^{\pi/2} = \frac{1}{2}.$$

Hence,

$$G_3 = (Gm) \cdot (2\pi) \cdot \frac{1}{2} \cdot 1 = Gm\pi.$$

In other words, in the coordinate system we chose, gravity is given by

$$\mathbf{G} = \langle 0, 0, Gm\pi \rangle.$$

The magnitude is $|\mathbf{G}| = \boxed{Gm\pi}$.

§49.4 Solution to Exercise 28.1 (find point on parametrized surface)

Exercise 28.1. Consider a surface $\mathcal{S}$ given by the parametrization $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\mathbf{r}(u, v) = \langle u - v, u + v, uv \rangle$$

for all (u, v) in $\mathbb{R}^2$. Compute the real number k for which the point $P = (2, 8, k)$ lies on $\mathcal{S}$.

For $P = (2, 8, k)$ to lie on $\mathcal{S}$, there must exist u, and v such that:

$$u - v = 2, \quad u + v = 8, \quad uv = k.$$

Adding the first two equations:

$$(u - v) + (u + v) = 2 + 8 \implies 2u = 10 \implies u = 5.$$

Subtracting the first equation from the second:

$$(u + v) - (u - v) = 8 - 2 \implies 2v = 6 \implies v = 3.$$

Hence $k = uv = \boxed{15}$.

§49.5 Solution to Exercise 29.1 (find tangent plane to parametrized surface)

Exercise 29.1. Consider a surface $\mathcal{S}$ given by the parametrization $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\mathbf{r}(u, v) = \langle u - v, u + v, uv \rangle$$

for all (u, v) in $\mathbb{R}^2$. Compute the tangent plane to $\mathcal{S}$ at the point $(3, 7, 10)$.

The first step is to solve for (u, v), much like in the preceding Exercise 28.1. The point $(3, 7, 10)$ must satisfy the parametrization equations:

$$u - v = 3, \quad u + v = 7, \quad uv = 10.$$

Solving gives $(u, v) = (5, 2)$.

Now, from the discussion in Section 29.2, the idea is that the normal vector for our tangent plane ought to be given by

$$\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$$

at every point of the surface.

So, we first compute the derivatives:

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial u} &= \left(\frac{\partial}{\partial u}(u-v), \frac{\partial}{\partial u}(u+v), \frac{\partial}{\partial u}(uv) \right) \\
&= \langle 1, 1, v \rangle \\
\frac{\partial \mathbf{r}}{\partial v} &= \left(\frac{\partial}{\partial v}(u-v), \frac{\partial}{\partial v}(u+v), \frac{\partial}{\partial v}(uv) \right) \\
&= \langle -1, 1, u \rangle.
\end{aligned}$$

At $(u, v) = (5, 2)$ we get

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial u}(5,2) &= \langle 1, 1, 2 \rangle \\
\frac{\partial \mathbf{r}}{\partial v}(5,2) &= \langle -1, 1, 5 \rangle.
\end{aligned}$$

The cross product is then given by

$$\frac{\partial \mathbf{r}}{\partial u}(5,2) \times \frac{\partial \mathbf{r}}{\partial v}(5,2) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 1 & 2 \\ -1 & 1 & 5 \end{vmatrix} = 3\mathbf{e}_1 - 7\mathbf{e}_2 + 2\mathbf{e}_3.$$

Thus, the normal vector is $\langle 3, -7, 2 \rangle$. So the equation of the tangent plane should be

$$3x - 7y + 2z = k$$

for some number k. To pass through $(3, 7, 10)$, we take $k = 3 \cdot 3 - 7 \cdot 7 + 2 \cdot 10 = -20$. Hence the final answer is

$$\boxed{3x - 7y + 2z = -20}.$$

§49.6 Solution to Exercise 30.1 (surface area of paraboloid)

Exercise 30.1. Compute the surface area of the surface defined by $z = x^2 + y^2 \le 1$.

The surface $\mathcal{S}$ is parametrized by

$$\mathbf{r}(x, y) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix}$$

where (x, y) lies within the disk $x^2 + y^2 \le 1$. Compute the partial derivatives of $\mathbf{r}$ with respect to x and y:

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial x} &= \begin{pmatrix} \frac{\partial x}{\partial x} \\ \frac{\partial y}{\partial x} \\ \frac{\partial z}{\partial x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix} \\
\frac{\partial \mathbf{r}}{\partial y} &= \begin{pmatrix} \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial y} \\ \frac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2y \end{pmatrix}.
\end{aligned}$$

Hence, the cross product is given by $\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y}$:

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} &= (0 \cdot 2y - 1 \cdot 2x)\mathbf{e}_1 - (1 \cdot 2y - 0 \cdot 2x)\mathbf{e}_2 + (1 \cdot 1 - 0 \cdot 0)\mathbf{e}_3 \\ &= \begin{pmatrix} -2x \\ 2y \\ 1 \end{pmatrix}.\end{aligned}$$

Hence the magnitude of this cross product is:

$$\left| \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} \right| = \sqrt{(-2x)^2 + (-2y)^2 + (1)^2} = \sqrt{4x^2 + 4y^2 + 1}.$$

Hence, the surface area of the surface in question is given by

$$\begin{aligned}\operatorname{SurfArea}(\mathcal{S}) &= \iint_{x^2+y^2\le 1} \left| \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} \right| \mathrm{d}x\,\mathrm{d}y \\ &= \iint_{x^2+y^2\le 1} \sqrt{4x^2+4y^2+1}\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Due to the circular symmetry, it is convenient to switch to polar coordinates; we write

$$\begin{aligned}\operatorname{SurfArea}(\mathcal{S}) &= \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \sqrt{4r^2+1} \cdot r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \left(\int_{\theta=0}^{2\pi} \mathrm{d}\theta \right) \left(\int_{r=0}^{1} \sqrt{4r^2+1} \cdot r\,\mathrm{d}r \right).\end{aligned}$$

The integral over r can be evaluated by using u-substitute according to

$$u := 4r^2 + 1 \implies \mathrm{d}u = 8r\,\mathrm{d}r \implies r\,\mathrm{d}r = \frac{\mathrm{d}u}{8}$$

so

$$\begin{aligned}\int_{r=0}^{1} \sqrt{4r^2+1} \cdot r\,\mathrm{d}r &= \int_{u=1}^{5} \left(\frac{\sqrt{u}}{8} \right) \mathrm{d}u \\ &= \frac{1}{8} \int_1^5 \sqrt{u}\,\mathrm{d}u = \left[\frac{1}{8} \cdot \frac{2}{3} u^{3/2} \right]_{u=1}^{5} = \frac{1}{12}\left(5^{3/2} - 1^{3/2}\right) \\ &= \frac{1}{12}\left(5\sqrt{5} - 1\right).\end{aligned}$$

And of course $\int_{\theta=0}^{2\pi} \mathrm{d}\theta = 2\pi$. Hence the answer

$$\boxed{\frac{5\sqrt{5}-1}{6}\pi}.$$

§49.7 Solution to Exercise 30.2 (Archimedes hat-box theorem)

The area of the cylinder part is straightforward and does not need calculus: it's a cylinder whose base has circumference 2π and which has height $b - a$, so the surface area is

$$2\pi(b - a).$$

So the main part of the problem is to show that the blue part of the sphere in Figure 69 has the same surface area.

It's sufficeint to solve the problem in the case $0 \le a \le b \le 1$. (If a and b are both negative, then you can do a reflection argument. And if $a < 0 < b$, then one should split the surface area at the equator of the sphere (along $z = 0$) into two parts; then add them together.)

We'll adopt the calculation in Section 30.3 for our purposes. In that section, we were able to calculate the surface of the unit hemisphere by viewing the sphere as the level surface of $g(x, y, z) = x^2 + y^2 + z^2 = 1$, and using the formula

$$\frac{\nabla g}{\partial g/\partial z} = \frac{\langle 2x, 2y, 2z\rangle}{2z} = \left\langle \frac{x}{z}, \frac{y}{z}, 1\right\rangle.$$

to derive that

$$\operatorname{SurfArea}(\text{hemisphere}) = \iint_{x^2+y^2\le 1} \frac{1}{\sqrt{1-(x^2+y^2)}}\,\mathrm{d}x\,\mathrm{d}y.$$

For Archimedes hat-box theorem, the integral itself stays the same; the change is that rather than integrating along the entire $x^2 + y^2 \le 1$ (which would give the surface area of the hemisphere), we instead integrate along the annulus

$$1 - b^2 \le x^2 + y^2 \le 1 - a^2.$$

Indeed, this is the shadow of the surface area in Figure 69 onto the xy-plane. (Indeed, the bottom disk has radius $\sqrt{1-a^2}$ and the top disk has radius $\sqrt{1-b^2}$.)

Getting back to integration, the surface area we seek for the sphere is thus

$$S = \iint_{1-b^2\le x^2+y^2\le 1-a^2} \frac{1}{\sqrt{1-(x^2+y^2)}}\,\mathrm{d}x\,\mathrm{d}y.$$

Now we just have to redo the calculation in Section 29.2 with only slight modifications. We use polar coordinates to change this to

$$\begin{aligned} S &= \int_{\theta=0}^{2\pi}\int_{r=\sqrt{1-b^2}}^{\sqrt{1-a^2}} \frac{1}{\sqrt{1-r^2}}(r\,\mathrm{d}r\,\mathrm{d}\theta) \\ &= \left(\int_{\theta=0}^{2\pi}\mathrm{d}\theta\right)\left(\int_{r=\sqrt{1-b^2}}^{\sqrt{1-a^2}} \frac{r}{\sqrt{1-r^2}}\,\mathrm{d}r\right). \end{aligned}$$

The left integral is 2π. For the inner integral, use the u-substitution $u = 1 - r^2 \implies \frac{\mathrm{d}u}{\mathrm{d}r} = -2r$ to get

$$\int_{r=\sqrt{1-b^2}}^{\sqrt{1-a^2}} \frac{r}{\sqrt{1-r^2}}\,\mathrm{d}r = \int_{u=b^2}^{a^2} -\frac{1}{2}u^{-\frac{1}{2}}\,\mathrm{d}u = \int_{u=a^2}^{b^2} \frac{1}{2}u^{-\frac{1}{2}}\,\mathrm{d}u = \left[u^{\frac{1}{2}}\right]_{u=a^2}^{b^2} = b - a.$$

Thus we get

$$S = 2\pi(b-a)$$

as we needed.

Chapter 50. Solutions to Part India

§50.1 Solution to Exercise 31.1 (touch grass)

Exercise 31.1. Take a few deep breaths, touch some grass, and have a nice drink of water, so that you can look at Figure 75 without feeling fear.

Here's a picture of moonlight sailing after having to grade the third midterm.

Figure 118: Moonlight sailing after the third midterm.

§50.2 Solution to Exercise 31.2 (print a poster)

Exercise 31.2. Print out a copy of the high-resolution version of Figure 75 (which can be downloaded at `https://web.evanchen.cc/textbooks/poster-stokes.pdf`) and hang it in your room.

Here's a picture of a leftover poster taped up in my room.

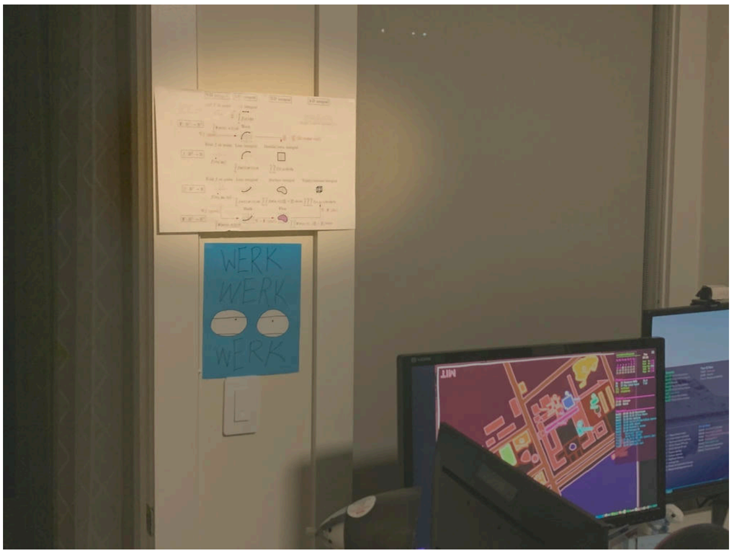

Figure 119: Hanging up a poster in my room.

§50.3 Solution to Exercise 32.1 (divergence of gravity)

Exercise 32.1. Consider the force of gravity $\mathbf{G}$ exerted by a point mass of mass m at a point O. Show that

$$\nabla \cdot \mathbf{G} = 0$$

at every point *except* O.

Recall that the gravity vector field is given by

$$\mathbf{G} = \left\langle \frac{-Gmx}{(x^2+y^2+z^2)^{\frac{3}{2}}}, \frac{-Gmy}{(x^2+y^2+z^2)^{\frac{3}{2}}}, \frac{-Gmz}{(x^2+y^2+z^2)^{\frac{3}{2}}} \right\rangle$$

For brevity, we let $\rho = \sqrt{x^2+y^2+z^2}$,

$$\mathbf{G} = \left\langle \frac{-Gmx}{\rho^3}, \frac{-Gmy}{\rho^3}, \frac{-Gmz}{\rho^3} \right\rangle.$$

Ignoring the constant factor Gm, we start by calculating the derivative of the first component with respect to x, that is:

$$\frac{\partial}{\partial x}\frac{x}{\rho^3}.$$

Using the quotient rule

$$\frac{\partial}{\partial x}\frac{x}{\rho^3} = \frac{(\rho^3)\frac{\partial}{\partial x}(x) - x\frac{\partial}{\partial x}(\rho^3)}{(\rho^3)^2}.$$

Since $\rho^3 = (x^2+y^2+z^2)^{3/2}$, the chain rule gives

$$\frac{\partial}{\partial x}(\rho^3) = \frac{3}{2}(x^2+y^2+z^2)^{1/2} \cdot 2x = 3x\rho.$$

Hence, the quotient rule gives

$$\frac{\partial}{\partial x}\frac{x}{\rho^3} = \frac{\rho^3 \cdot 1 - x \cdot 3x\rho}{r^6} = \frac{\rho^2 - 3x^2}{\rho^5}.$$

Now the divergence is given by

$$\begin{aligned}
\nabla \cdot \mathbf{G} &= -Gm\left(\frac{\partial}{\partial x}\frac{x}{\rho^3} + \frac{\partial}{\partial x}\frac{y}{\rho^3} + \frac{\partial}{\partial x}\frac{z}{\rho^3}\right) \\
&= -Gm\left(\frac{\rho^2 - 3x^2}{\rho^5} + \frac{\rho^2 - 3y^2}{\rho^5} + \frac{\rho^2 - 3z^2}{\rho^5}\right) \\
&= -Gm\frac{3\rho^2 - 3(x^2+y^2+z^2)}{\rho^5} \\
&= 0
\end{aligned}$$

as claimed.

§50.4 Solution to Exercise 33.1 (parabola arc v1)

Exercise 33.1 (Suggested by Ting-Wei Chao)**.** Let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Let

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$ using direct parametrization.

We split $\mathcal{C}$ into two parts $\mathcal{C}_1$ and $\mathcal{C}_2$ corresponding to the parabola and the segment, respectively. These are colored red and blue in Figure 120, respectively.

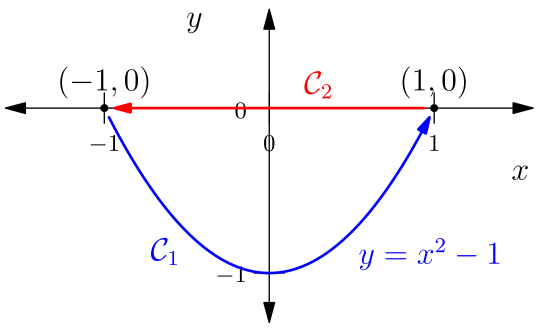

Figure 120: The curve $\mathcal{C}$ (and the enclosed region).

For $\mathcal{C}_1$, we choose the parametrization

$$\mathbf{r}(t) = (t, t^2 - 1)$$

for $-1 \le t \le 1$, so

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(t, t^2 - 1) = \begin{pmatrix} t^4 \\ t^4 \end{pmatrix} \quad \mathbf{r}'(t) = \begin{pmatrix} 1 \\ 2t \end{pmatrix}.$$

Then we get

$$\begin{aligned}\int_{\mathcal{C}_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{t=-1}^{1} \begin{pmatrix} t^4 \\ t^4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2t \end{pmatrix} \mathrm{d}t \\ &= \int_{t=-1}^{1} (t^4 + t^4 \cdot 2t)\, \mathrm{d}t \\ &= \left[\frac{t^5}{5} + \frac{1}{3} t^6 \right]_{t=-1}^{1} = \frac{2}{5}.\end{aligned}$$

Next, we parameterize $\mathcal{C}_2$ by $(1-t, 0)$ for $0 \le t \le 2$. Then

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(1-t, 0) = \begin{pmatrix} (1-t)^2 \\ 1 \end{pmatrix} \quad \mathbf{r}'(t) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Then

$$\begin{aligned}\int_{\mathcal{C}_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{t=-1}^{1} \begin{pmatrix} (1-t)^2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mathrm{d}t \\ &= \int_{t=0}^{2} -(1-t)^2\, \mathrm{d}t \\ &= -\frac{2}{3}.\end{aligned}$$

Putting this all together we get

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{\mathcal{C}_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} + \int_{\mathcal{C}_2} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \frac{2}{5} - \frac{2}{3} = \boxed{-\frac{4}{15}}.$$

§50.5 Solution to Exercise 33.2 (work from $45°$ angle)

Exercise 33.2. Let $\mathcal{C}$ be a curve in $\mathbb{R}^2$ from $(0,0)$ to $(2,3)$ whose arc length is 7. Let $\mathbf{F}$ be a vector field with the property that for any point P on the curve,

- $\mathbf{F}(P)$ has magnitude 5;
- $\mathbf{F}(P)$ makes a $45°$ angle with the tangent vector to $\mathcal{C}$ at P (the tangent vector points along the direction of $\mathcal{C}$).

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$.

Take any parametrization of the curve $\mathcal{C}$, say from $t = 0$ to 1. The work we seek is then

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \int_{t=0}^{1} \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\, \mathrm{d}t.$$

The point of the problem is that $\mathbf{F}(\mathbf{r}(t))$ is supposed to have magnitude 5 and form a $45°$ angle with $\mathbf{r}'(t)$. Hence, if we use the *geometric* definition of the dot product, we get

$$\begin{aligned}\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \int_{t=0}^{1} |\mathbf{F}(\mathbf{r}(t))| \, |\mathbf{r}'(t)| \cos(45^\circ) \, \mathrm{d}t \\ &= \int_{t=0}^{1} 5 \, |\mathbf{r}'(t)| \cos(45^\circ) \, \mathrm{d}t \\ &= 5\cos(45^\circ) \int_{t=0}^{1} |\mathbf{r}'(t)| \, \mathrm{d}t.\end{aligned}$$

But that integral is the arc length of $\mathcal{C}$. So

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = 5\cos(45^\circ) \cdot 7 = 35\cos(45^\circ) = \boxed{\frac{35\sqrt{2}}{2}}.$$

§50.6 Solution to Exercise 34.1 (checking conservativeness)

Exercise 34.1. Is the vector field

$$\mathbf{F}(x, y) = \begin{pmatrix} \sin(e^x) \\ \arctan(y^\pi + \pi^y) \end{pmatrix}$$

conservative?

Yes, because

$$\begin{aligned}\frac{\partial}{\partial y} \sin(e^x) &= 0 \\ \frac{\partial}{\partial x} \arctan(y^\pi + \pi^y) &= 0.\end{aligned}$$

In general, the point of this exercise is that every vector field of the form $\mathbf{F}(x, y) = \langle \text{stuff only involving } x, \text{stuff only involving } y \rangle$ is always conservative, because the relevant partial derivatives are both 0.

§50.7 Solution to Exercise 34.2 (work exercise)

Exercise 34.2. Calculate the line integral

$$\oint_{\mathcal{C}} (x^2 - y) \, \mathrm{d}x + (y^2 - x) \, \mathrm{d}y$$

where $\mathcal{C}$ is the boundary of the region enclosed by the circle $x^2 + y^2 = 4$, oriented counterclockwise.

The vector field is conservative, because

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

So the answer is 0. (If you didn't notice this at first and tried to use Green's theorem, you should notice at the moment where the curl turns out to 0.)

§50.8 Solution to Exercise 34.3 (parabola arc v2)

Exercise 34.3 (Suggested by Ting-Wei Chao)**.** As in Exercise 33.1, let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Again let

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r}$ this time using Green's Theorem.

Let $P = x^2(y+1)$ and $Q = (y+1)^2$. Thus

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0 - x^2 = -x^2.$$

Let $\mathcal{R}$ be the region between $y = x^2 - 1$ and $y = 0$, enclosed by $\mathcal{C}$. Then $\mathcal{R}$ can be described by the inequalities

$$\begin{aligned} -1 \le x \le 1 \\ x^2 - 1 \le y \le 0. \end{aligned}$$

See Figure 121.

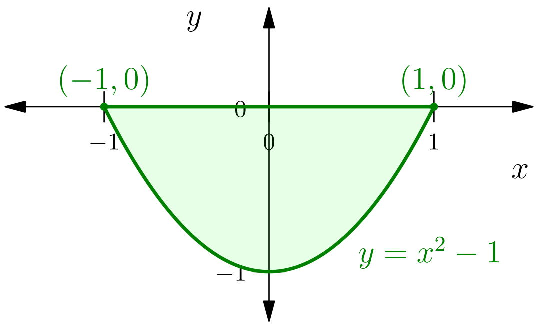

Figure 121: The curve $\mathcal{C}$ encloses a region $\mathcal{R}$.

Hence from Green's theorem, we have

$$\begin{aligned} \int_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= \iint_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A \\ &= \int_{x=-1}^{1} \int_{y=x^2-1}^{0} (-x^2)\, \mathrm{d}y\, \mathrm{d}x \\ &= \int_{x=-1}^{1} x^2(x^2-1)\, \mathrm{d}x \\ &= \left[\frac{x^5}{5} - \frac{x^3}{3} \right]_{x=-1}^{1} = \frac{2}{5} - \frac{2}{3} = \boxed{-\frac{4}{15}}. \end{aligned}$$

§50.9 Solution to Exercise 34.4 (the shoelace formula)

Exercise 34.4 (*) (Shoelace formula)**.** Let $n \ge 3$ be an integer and suppose $\mathcal{P} = P_1 P_2 \ldots P_n$ is a convex n-gon in $\mathbb{R}^2$, where the vertices $P_i = (x_i, y_i)$ are labeled counterclockwise. Use Green's theorem to prove the following formula for the area of $\mathcal{P}$:

$$\operatorname{Area}(\mathcal{P}) = \frac{1}{2}\sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i).$$

Here $x_0 = x_n$ and $y_0 = y_n$ by convention, so the $i = 0$ summand is $x_n y_1 - x_1 y_n$.

Green's theorem states that for a simple closed curve $\mathcal{C}$ enclosing a region $\mathcal{R}$, we have

$$\iint_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}A = \oint_{\mathcal{C}} (P\,\mathrm{d}x + Q\,\mathrm{d}y).$$

We take $\mathcal{C}$ as the polygon, oriented counterclockwise, and $\mathcal{R}$ as its interior. We will chose the vector field

$$\mathbf{F} = \begin{pmatrix} 0 \\ x \end{pmatrix}$$

so that the 2D scalar curl equals

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial}{\partial x}(x) - \frac{\partial}{\partial y}(0) = 1.$$

Thus, Green's theorem gives:

$$\operatorname{Area}(\mathcal{P}) = \oint_{\mathcal{C}} x\,\mathrm{d}y.$$

Now we evaluate the line integral "manually" across the n sides of the polygon. As the boundary $\mathcal{C}$ consists of the edges $P_i P_{i+1}$, we sum over each edge:

$$\oint_{\mathcal{C}} x\,\mathrm{d}y = \sum_{i=0}^{n-1} \int_{\text{From } P_i \text{ to } P_{i+1}} x\,\mathrm{d}y.$$

Let's parametrize the segment joining (x_i, y_i) to (x_{i+1}, y_{i+1}) by just the constant speed parametrization taking unit time:

$$\mathbf{r}(t) = \langle (1-t)x_i + t x_{i+1}, (1-t)y_i + t y_{i+1} \rangle \quad 0 \le t \le 1.$$

Thus,

$$\mathbf{F}(\mathbf{r})(t) = \begin{pmatrix} 0 \\ (1-t)x_i + t x_{i+1} \end{pmatrix}$$
$$\mathbf{r}'(t) = \begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix}$$

so we are integrating the dot product

$$\begin{aligned}\int_{\text{From } P_i \text{ to } P_{i+1}} x\,\mathrm{d}y &= \int_{t=0}^{1} \begin{pmatrix} 0 \\ (1-t)x_i + tx_{i+1} \end{pmatrix} \cdot \begin{pmatrix} x_{i+1} - x_i \\ y_{i+1} - y_i \end{pmatrix} \mathrm{d}t \\ &= \int_{t=0}^{1} ((1-t)x_i + tx_{i+1})(y_{i+1} - y_i)\,\mathrm{d}t \\ &= (y_{i+1} - y_i)\left(x_i \int_{t=0}^{1} (1-t)\,\mathrm{d}t + x_{i+1} \int_{t=0}^{1} t\,\mathrm{d}t \right) \\ &= (y_{i+1} - y_i)\left(x_i \cdot \frac{1}{2} + x_{i+1} \cdot \frac{1}{2} \right) \\ &= \frac{1}{2}(y_{i+1} - y_i)(x_i + x_{i+1}) \\ &= \frac{x_i y_{i+1} - x_{i+1} y_i}{2} + \frac{x_{i+1} y_{i+1} - x_i y_i}{2}.\end{aligned}$$

Summing over all edges, the first term is the desired right-hand side, while the second term cancels since

$$\frac{x_1 y_1 - x_n y_n}{2} + \frac{x_2 y_2 - x_1 y_1}{2} + \frac{x_3 y_3 - x_2 y_2}{2} + \dots + \frac{x_n y_n - x_{n-1} y_{n-1}}{2} = 0.$$

In other words,

$$\begin{aligned}\sum_{i=0}^{n-1} \int_{\text{From } P_i \text{ to } P_{i+1}} x\,\mathrm{d}y &= \sum_{i=0}^{n-1} \left(\frac{x_i y_{i+1} - x_{i+1} y_i}{2} + \frac{x_{i+1} y_{i+1} - x_i y_i}{2} \right) \\ &= \sum_{i=0}^{n-1} \left(\frac{x_i y_{i+1} - x_{i+1} y_i}{2} \right)\end{aligned}$$

and the proof is complete.

§50.10 Solution to Exercise 35.1 (parabola arc v3)

Exercise 35.1 (Suggested by Ting-Wei Chao)**.** As in Exercise 33.1 and Exercise 34.3, let $\mathcal{C}$ be the oriented closed curve formed by the arc of the parabola $y = x^2 - 1$ running from $(-1, 0)$ to $(1, 0)$, followed by a line segment from $(1, 0)$ back to $(-1, 0)$. Again let

$$\mathbf{F}(x, y) = \begin{pmatrix} x^2(y+1) \\ (y+1)^2 \end{pmatrix}.$$

Compute $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}s$ using direct parametrization and by using Green's Theorem for flux.

We can pretty much copy Exercise 33.1 and Exercise 34.3, and just make slight modifications to get the 2D flux instead of the work. See the following figure, which is just a copy of Figure 120 and Figure 121.

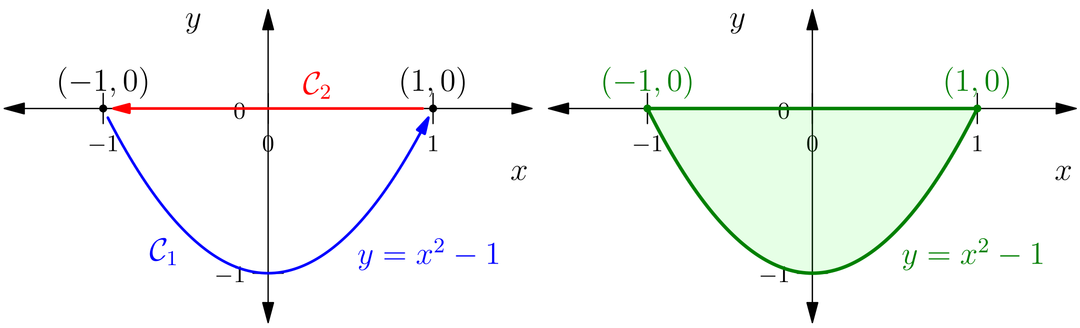

Figure 122: The parabola arc again.

Using direct parametrization

Again we split $\mathcal{C}$ into two parts $\mathcal{C}_1$ and $\mathcal{C}_2$ corresponding to the parabola and the segment, respectively.

For $\mathcal{C}_1$, we choose the parametrization

$$\mathbf{r}(t) = (t, t^2 - 1)$$

for $-1 \le t \le 1$. As we saw in Exercise 33.1, we have

$$\mathbf{r}(t) = (t, t^2 - 1)$$

for $-1 \le t \le 1$, so

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(t, t^2 - 1) = \begin{pmatrix} t^4 \\ t^4 \end{pmatrix} \quad \mathbf{r}'(t) = \begin{pmatrix} 1 \\ 2t \end{pmatrix}.$$

So rotating $\mathbf{F}$ as in Section 35.5, we have

$$\begin{aligned} \int_{\mathcal{C}_1} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s &= \int_{t=-1}^{1} \begin{pmatrix} -t^4 \\ t^4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2t \end{pmatrix} \mathrm{d}t \\ &= \int_{t=-1}^{1} (-t^4 + t^4 \cdot 2t) \, \mathrm{d}t \\ &= \left[-\frac{t^5}{5} + \frac{1}{3} t^6 \right]_{t=-1}^{1} = -\frac{2}{5}. \end{aligned}$$

Next, we parameterize $\mathcal{C}_2$ by $(1 - t, 0)$ for $0 \le t \le 2$. Again, repeating from Exercise 33.1, we have

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(1 - t, 0) = \begin{pmatrix} (1 - t)^2 \\ 1 \end{pmatrix} \quad \mathbf{r}'(t) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Rotating $\mathbf{F}$ as in Section 35.5, we have

$$\begin{aligned} \int_{\mathcal{C}_2} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s &= \int_{t=-1}^{1} \begin{pmatrix} -1 \\ (1 - t)^2 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} \mathrm{d}t \\ &= \int_{t=0}^{2} 1 \, \mathrm{d}t \\ &= 2. \end{aligned}$$

Putting this all together we get

$$\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = \int_{\mathcal{C}_1} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s + \int_{\mathcal{C}_2} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = -\frac{2}{5} + 2 = \boxed{\frac{8}{5}}.$$

Using Green's Theorem

This is similar to Exercise 34.3, and the only change we make is the integrand. Letting $P = x^2(y+1)$ and $Q = (y+1)^2$, we consider

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = 2x(y+1) + 2(y+1) = 2(x+1)(y+1).$$

Again let $\mathcal{R}$ be the region between $y = x^2 - 1$ and $y = 0$, enclosed by $\mathcal{C}$. The region $\mathcal{R}$ hasn't changed and is given by

$$\begin{aligned} -1 &\le x \le 1 \\ x^2 - 1 &\le y \le 0. \end{aligned}$$

Hence from Green's theorem, we have

$$\begin{aligned} \int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s &= \iint_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A \\ &= \int_{x=-1}^{1} \int_{y=x^2-1}^{0} 2(x+1)(y+1) \, \mathrm{d}y \, \mathrm{d}x \\ &= \int_{x=-1}^{1} (x+1) \int_{y=x^2-1}^{0} 2(y+1) \, \mathrm{d}y \, \mathrm{d}x \\ &= \int_{x=-1}^{1} (x+1)(1-x^4) \, \mathrm{d}x \\ &= \int_{x=-1}^{1} (-x^5 - x^4 + x + 1) \, \mathrm{d}x \\ &= \left[-\frac{x^6}{6} - \frac{x^5}{5} + x^2 + x \right]_{x=-1}^{1} = \boxed{\frac{8}{5}}. \end{aligned}$$

§50.11 Solution to Exercise 35.2 (flux across a triangle)

Exercise 35.2. Triangle ABC has vertices $A = (-5, 0)$, $B = (9, 0)$, and C on the positive y-axis. The flux of the vector field

$$\mathbf{F}(x, y) = \begin{pmatrix} x + 7y^2 \\ x^2 + 7y \end{pmatrix}$$

across the perimeter of ABC, oriented counterclockwise, is 672. Compute the length of the perimeter of ABC.

We are given a triangle ABC with vertices:

$$A = (-5, 0), \quad B = (9, 0), \quad C = (0, h),$$

where C is on the positive y-axis. We'll find h, after which we can get the length of the perimeter easily.

The vector field is:

$$\mathbf{F}(x, y) = (P, Q) = (x + 7y^2, x^2 + 7y).$$

Let $\mathcal{C}$ be the boundary of that triangle. Green's theorem gives

$$672 = \oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = \iint_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A.$$

Computing partial derivatives:

$$\begin{aligned} \frac{\partial P}{\partial x} &= \frac{\partial}{\partial x}(x + 7y^2) = 1, \\ \frac{\partial Q}{\partial y} &= \frac{\partial}{\partial y}(x^2 + 7y) = 7. \end{aligned}$$

Thus, the divergence is:

$$\nabla \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = 1 + 7 = 8.$$

Hence, we get

$$672 = \iint_{\mathcal{R}} 8 \, \mathrm{d}A = 8 \operatorname{Area}(\mathcal{R}).$$

But the area of a triangle is:

$$\operatorname{Area}(\triangle ABC) = \frac{1}{2} \cdot \text{base} \cdot \text{height} = \frac{1}{2} \cdot 14 \cdot h = 7h.$$

Substituting this into the integral equation:

$$8 \cdot 7h = 672.$$

Solving for h:

$$56h = 672 \implies h = \frac{672}{56} = 12.$$

Hence, we know $C = (0, 12)$.

We can then get the perimeter by computing all the side lengths:

$$\begin{aligned} AB &= |9 - (-5)| = 14 \\ BC &= \sqrt{(9-0)^2 + (0-12)^2} = \sqrt{81 + 144} = \sqrt{225} = 15 \\ CA &= \sqrt{(0 - (-5))^2 + (12 - 0)^2} = \sqrt{25 + 144} = \sqrt{169} = 13. \end{aligned}$$

Thus, the perimeter is: $14 + 15 + 13 = \boxed{42}$.

§50.12 Solution to Exercise 37.1 (return of the butterfly)

Exercise 37.1. Another butterfly is fluttering in the xy plane with position $\mathbf{r}(t) = \langle \sin(t), \sin(t) \rangle$. Let $\mathcal{C}$ denote its trajectory between $0 \le t \le 2\pi$. Compute $\int_{\mathcal{C}} (x \, \mathrm{d}x)$ and $\int_{\mathcal{C}} (y \, \mathrm{d}x)$.

Note the butterfly starts at $\mathbf{r}(0) = (0, 0)$ and ends at $\mathbf{r}(2\pi) = (0, 0)$ as well.

The short solution to both parts

The answer is $\boxed{0}$ regardless of the vector field $\mathbf{F}$! Two ways to see this.

- This follows by Green's theorem, because the trajectory $\mathbf{r}$ cuts a degenerate parallelogram of area zero.
- The butterfly is tracing its own path in reverse, so the part from $0 \le t \le \frac{\pi}{2}$ cancels $\frac{\pi}{2} \le t \le \pi$ while the part from $\pi \le t \le \frac{3}{2}\pi$ cancels $\frac{3}{2}\pi \le t \le 2\pi$.

Another short way for the first vector field

The first integral is of the conservative vector field $\mathbf{F} = \begin{pmatrix} x \\ 0 \end{pmatrix}$, because its 2D scalar curl is $0 - 0 = 0$. So the fundamental theorem of calculus also implies the answer is 0, because the path is a loop. (If f is a potential function, then the answer should be $f(0,0) - f(0,0) = 0$. You could compute the potential function $f(x,y) = \frac{x^2}{2} + C$ if you want, but it's not needed.)

The long way for the second non-conservative field

In the second part, $\mathbf{F} = \begin{pmatrix} y \\ 0 \end{pmatrix}$ is not conservative. Let's say you didn't come up with the idea in the slick solution. Then you could still compute the integral manually by taking

$$\mathbf{r}'(t) = \langle \cos(t), \cos(t) \rangle$$

so the line integral is given by

$$\int_{t=0}^{2\pi} \mathbf{F}(\sin(t), \sin(t)) \cdot \mathbf{r}'(t)\, dt = \int_{t=0}^{2\pi} \sin(t)\cos(t)\, dt.$$

This integral is 0; here are many ways to evaluate it.

1. Notice this is actually the same trig integral you got if you evaluated for the first line integral manually as well, so the answer should be the same, namely 0.
2. Notice the contribution from t and $t + \pi$ cancel for $0 \le t \le \pi$.
3. Another way to evaluate the integral is via the u-substitution $u = \sin(t)$, where $du = \cos(t)\, dt$:
$$\int_{t=0}^{2\pi} \sin(t)\cos(t)\, dt = \int_{u=\sin(0)}^{u=\sin(2\pi)} u\, du = \int_0^0 u\, du = 0.$$
4. Another way is to use the trig substitution
$$\int_{t=0}^{2\pi} \sin(t)\cos(t)\, dt = \int_{t=0}^{2\pi} \frac{\sin(2t)}{2}\, dt = \left[-\frac{\cos(2t)}{4} \right]_{t=0}^{2\pi} = 0.$$
5. If you are allergic to trig functions, a fifth approach is to remember that line integrals don't depend on the exact parametrization. So rather than using $\mathbf{r}(t) = \langle \sin(t), \sin(t) \rangle$, you could imagine cutting the butterfly's motion into three constant-velocity trajectories:
   - $\mathbf{r}_1(t) = (t, t)$ for $0 \le t \le 1$
   - $\mathbf{r}_2(t) = (1 - t, 1 - t)$ for $0 \le t \le 2$
   - $\mathbf{r}_3(t) = (t - 1, t - 1)$ for $0 \le t \le 1$.

   If you compute the three line integrals, the sum will also be zero.

§50.13 Solution to Exercise 37.2 (standard work and flux)

Exercise 37.2. Let $\mathcal{C}$ denote the unit circle $x^2 + y^2 = 1$ oriented counterclockwise, and consider the vector field $\mathbf{F}(x,y) = \langle x + 2y, 4x + 8y \rangle$. Compute $\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r}$ and $\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, ds$.

This is a cookie-cutter application of Green's theorem (both forms).

For the line integral, use Green's theorem with 2D scalar curl:

$$\begin{aligned}
\int_{\mathcal{C}} \mathbf{F} \cdot d\mathbf{r} &= \iint_{x^2+y^2\le 1} \left(\frac{\partial}{\partial x}(4x+8y) - \frac{\partial}{\partial y}(x+2y) \right) dA \\
&= \iint_{x^2+y^2\le 1} (4-2)\, dA \\
&= 2 \iint_{x^2+y^2\le 1} dA \\
&= \boxed{2\pi}.
\end{aligned}$$

For the flux, use Green's theorem with divergence:

$$\begin{aligned}
\int_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, ds &= \iint_{x^2+y^2\le 1} \nabla \cdot \mathbf{F}\, dA \\
&= \iint_{x^2+y^2\le 1} \left(\frac{\partial}{\partial x}(x+2y) + \frac{\partial}{\partial y}(4x+8y) \right) dA \\
&= \iint_{x^2+y^2\le 1} (1+8)\, dA \\
&= 9 \iint_{x^2+y^2\le 1} dA \\
&= \boxed{9\pi}.
\end{aligned}$$

§50.14 Solution to Exercise 37.3 (region with area π)

Exercise 37.3. Compute all real numbers k for which the following region has area π:

$$(kx+y)^2 + (x+ky)^2 \le \frac{1}{4}.$$

Let $\mathcal{R}$ denote the region in the problem. This is a change of variables problem where

$$\begin{aligned}
u &= x + ky \\
v &= kx + y
\end{aligned}$$

changes $\mathcal{R}$ into the disk $u^2 + v^2 \le \frac{1}{4}$ of radius $\frac{1}{2}$.

Let $\mathbf{T}$ denote the corresponding map $(u,v) \mapsto (x,y)$. Compute the inverse of the Jacobian

$$J_{\mathbf{T}^{-1}} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & k \\ k & 1 \end{pmatrix}.$$

So

$$\det J_{\mathbf{T}} = \frac{1}{\begin{vmatrix} 1 & k \\ k & 1 \end{vmatrix}} = \frac{1}{1-k^2}.$$

Now, the problem condition tells us

$$\pi = \iint_{\mathcal{R}} \mathrm{d}x\,\mathrm{d}y = \iint_{u^2+v^2\le\frac{1}{4}} \left|\frac{1}{1-k^2}\right| \mathrm{d}u\,\mathrm{d}v = \left|\frac{1}{1-k^2}\right| \iint_{u^2+v^2\le\frac{1}{4}} \mathrm{d}u\,\mathrm{d}v = \left|\frac{1}{1-k^2}\right| \cdot \frac{\pi}{4}.$$

So the equation we are trying to solve is

$$\pi = \left|\frac{1}{1-k^2}\right| \cdot \frac{\pi}{4} \iff k^2 - 1 = \pm\frac{1}{4} \iff k^2 = \frac{3}{4} \quad \text{or} \quad k^2 = \frac{5}{4}.$$

Hence the answers are

$$k = \boxed{\pm\frac{\sqrt{3}}{2} \ \text{ or } \ \pm\frac{\sqrt{5}}{2}}.$$

§50.15 Solution to Exercise 37.4 (center of mass of wedge)

Exercise 37.4. Compute the center of mass of the region where $y \ge 0$ and $3x^2 \le y^2 \le 9 - x^2$, assuming constant density.

First, we sketch the region. The condition $x^2 + y^2 \le 9$ represents a circle of radius 3 centered at the origin. The inequality $y \ge 0$ and $y \ge \sqrt{3}|x|$ cuts out a wedge covering the top half of the circle within the angle range $\frac{\pi}{3} \le \theta \le \frac{2\pi}{3}$ in polar coordinates. See Figure 123.

Switching to polar coordinates, the region in polar coordinates is bounded by:

- $0 \le r \le 3$ (radius of the circle),
- $\frac{\pi}{3} \le \theta \le \frac{2\pi}{3}$ (angular bounds determined by $y = \sqrt{3}x$ and $y = -\sqrt{3}x$).

The coordinates of the center of mass are then given by:

$$\bar{x} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} x\,\mathrm{d}A, \quad \bar{y} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} y\,\mathrm{d}A.$$

- The area can computed by noticing the region is one-sixth of the area of the full circle:

$$\operatorname{Area}(\mathcal{R}) = \frac{1}{6} \cdot (3^2 \cdot \pi) = \frac{3\pi}{2}.$$

- We have $\bar{x} = 0$ by symmetry around the y-axis.
- We need to compute $\bar{y}$. Use $y = r\sin\theta$:

$$\bar{y} = \iint_{\mathcal{R}} y\,\mathrm{d}A = \int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \int_{r=0}^{3} r\sin\theta \cdot r\,\mathrm{d}r\,\mathrm{d}\theta = \int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \sin\theta \int_{r=0}^{3} r^2\,\mathrm{d}r\,\mathrm{d}\theta.$$

First, we compute

$$\int_{r=0}^{3} r^2\,\mathrm{d}r = \left[\frac{r^3}{3}\right]_{r=0}^{3} = \frac{27}{3} = 9.$$

Hence $\bar{y} = \iint_{\mathcal{R}} y\,\mathrm{d}A = 9\int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \sin\theta\,\mathrm{d}\theta$. Integrate $\sin\theta$ with respect to θ:

$$\int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \sin\theta\,\mathrm{d}\theta = -[\cos\theta]_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} = -\cos\left(\frac{2\pi}{3}\right) + \cos\left(\frac{\pi}{3}\right) = -\left(-\frac{1}{2}\right) + \frac{1}{2} = 1.$$

Thus, $\iint_{\mathcal{R}} y\,\mathrm{d}A = 9 \cdot 1 = 9$, and so

$$\bar{y} = \frac{1}{\operatorname{Area}(\mathcal{R})} \iint_{\mathcal{R}} y \, \mathrm{d}A = \frac{9}{\frac{3\pi}{2}} = \frac{6}{\pi}.$$

In conclusion, the center of mass is given by

$$(\bar{x}, \bar{y}) = \boxed{\left(0, \frac{6}{\pi}\right)}.$$

Digression: the long way for area

If you don't want to do geometry, you can manually compute $\operatorname{Area}(\mathcal{R})$ by the definition $\operatorname{Area}(\mathcal{R}) = \int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \int_{r=0}^{3} r \, \mathrm{d}r \, \mathrm{d}\theta$. First, integrate with respect to r:

$$\int_{r=0}^{3} r \, \mathrm{d}r = \left[\frac{r^2}{2}\right]_{r=0}^{3} = \frac{9}{2}.$$

Then, integrate with respect to θ:

$$A = \int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \frac{9}{2} \, \mathrm{d}\theta = \frac{9}{2}\left(\frac{2\pi}{3} - \frac{\pi}{3}\right) = \frac{9}{2} \cdot \frac{\pi}{3} = \frac{3\pi}{2}.$$

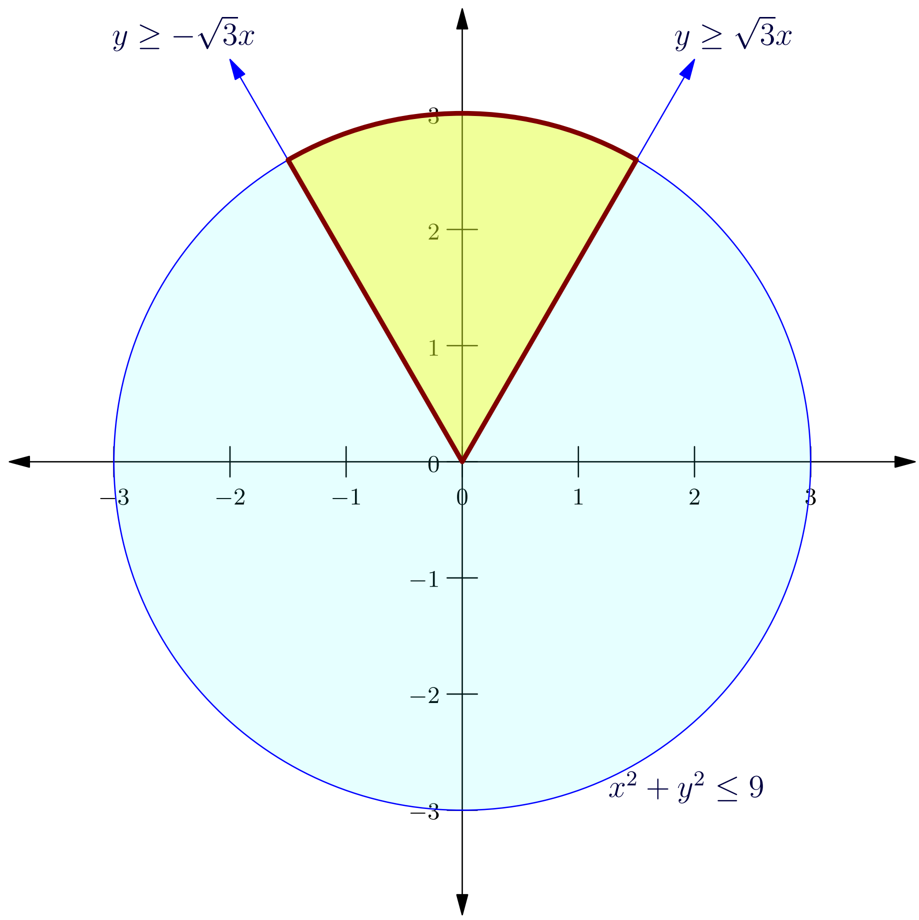

Figure 123: It's a one-sixth slice of a pizza or something.

Digression: the long way for $\overline{x}$

For comparison, we show what happens if you didn't notice the symmetry and proceed to integrate. In polar coordinates, $x = r\cos\theta$ and $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$. So:

$$\iint_{\mathcal{R}} x\,\mathrm{d}A = \int_{r=0}^{3}\int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} r\cos\theta\cdot r\,\mathrm{d}r\,\mathrm{d}\theta = \int_{r=0}^{3} r^2 \int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \cos\theta\,\mathrm{d}\theta\,\mathrm{d}r.$$

However, the inner integral is

$$\begin{aligned}\int_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \cos\theta\,\mathrm{d}\theta &= [\sin\theta]_{\theta=\frac{\pi}{3}}^{\frac{2\pi}{3}} \\ &= \sin\left(\frac{2\pi}{3}\right) - \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = 0.\end{aligned}$$

So the whole thing is 0.

§50.16 Solution to Exercise 37.5 (recovering h)

Exercise 37.5. Let $\mathcal{C}$ denote any path from $(0,0)$ to (π,π). Determine the unique function $h(x)$ for which $\mathbf{F}(x,y) = \langle xy + \cos(x), h(x) + \cos(y)\rangle$ is conservative, and moreover $\int_{\mathcal{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} = 0$.

Because $\mathbf{F}$ is conservative, we know the following two partial derivatives must be equal:

$$\begin{aligned}\frac{\partial}{\partial y}(xy+\cos(x)) &= x \\ \frac{\partial}{\partial x}(h(x)+\cos(y)) &= h'(x).\end{aligned}$$

From $h'(x) = x$ we deduce

$$h(x) = \frac{x^2}{2} + C$$

for some constant C.

So we almost know h, except we need to use the last piece of information to find C. First, recover a potential function for $\mathbf{F}$ in terms of C:

$$f(x,y) = \frac{1}{2}x^2y + \sin(x) + \sin(y) + Cy + C'$$

for some constant C' (which is irrelevant). Then use the fundamental theorem calculus for line integrals:

$$\begin{aligned}\int_{\mathcal{C}} \mathbf{F}\cdot\mathrm{d}\mathbf{r} &= f(\pi,\pi) - f(0,0) \\ \implies 0 &= \left(\frac{\pi^3}{2} + C\pi + C'\right) - C' \implies C = -\frac{\pi^2}{2}.\end{aligned}$$

Thus we've completely recovered the function h:

$$h(x) = \boxed{\frac{x^2}{2} - \frac{\pi^2}{2}}.$$

§50.17 Solution to Exercise 37.6 (integrating $\log(\sqrt{x}+1)$)

Exercise 37.6. Assume log is base $e \approx 2.718$. Use any method you want to compute

$$\int_{x=0}^{(e-1)^2} \log\left(\sqrt{x}+1\right) \mathrm{d}x.$$

Recommended approach: view the integral as the area under a curve, then switch from vertical to horizontal slicing.

At face value, this looks like an 18.01 integral, but we know from 18.01 that this integral is actually measuring the area under some curve; we denote that region by $\mathcal{R}$, shaded in blue below. The idea is that, to avoid having to deal with log and square root, we are going to use horizontal slicing for the region under the curve shown in the figure.

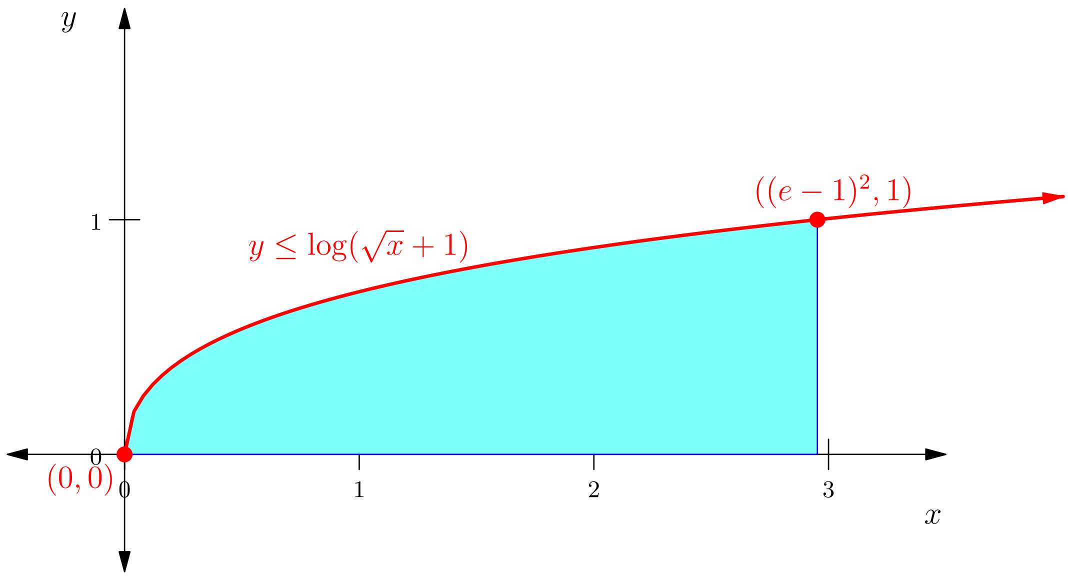

Figure 124: The region $y \le \log(\sqrt{x}+1)$ for $0 \le x \le (e-1)^2$.

The point is to now rewrite

$$y \le \log\left(\sqrt{x}+1\right) \iff e^y - 1 \le \sqrt{x} \iff x \ge \left(e^y - 1\right)^2.$$

Hence, we could equally well rewrite the shaded blue region $\mathcal{R}$ as:

$$0 \le y \le 1 \quad \text{and} \quad x \ge \left(e^y - 1\right)^2.$$

Writing this as a double integral gives

$$\begin{aligned}\operatorname{Area}(\mathcal{R}) &= \int_{y=0}^{1} \int_{x=(e^y-1)^2}^{(e-1)^2} \mathrm{d}x \, \mathrm{d}y = \int_{y=0}^{1} \left((e-1)^2 - (e^y-1)^2 \right) \mathrm{d}y \\ &= e^2 - 2e - \int_{y=0}^{1} (e^{2y} - 2e^y) \, \mathrm{d}y = e^2 - 2e - \left[\frac{e^{2y}}{2} - 2e^y \right]_{y=0}^{1} \\ &= e^2 - 2e - \left(\frac{e^2}{2} - 2e \right) - \left(\frac{1}{2} - 2 \right) = \boxed{\frac{e^2-3}{2}}.\end{aligned}$$

Remark

It is also possible to calculate an antiderivative of $\log(\sqrt{x}+1)$ directly by using integration by parts and u-substitution, but this process is time-consuming. The anti-derivative turns out to equal $-\frac{x}{2} + \sqrt{x} + (x-1)\log(\sqrt{x}+1) + C$.

Chapter 51. Solutions to Part Juliett

§51.1 Solution to Exercise 38.1 (flux through a surface)

Exercise 38.1. Calculate the flux of the vector field

$$\mathbf{F}(x,y,z) = \left\langle \frac{x}{3}, \frac{y}{4}, \frac{1}{5} \right\rangle$$

across the portion of the surface defined by

$$x^3 + y^4 = e^z, \quad 0 \le x \le 5, \quad 0 \le y \le 5$$

where the normal vector is oriented upwards.

We parametrize the surface by $\mathbf{r}(x,y) = \langle x, y, \log(x^3+y^4) \rangle$. Rather than deal with log, we use the second row of Table 23 and define the function

$$g(x,y,z) = x^3 + y^4 - e^z.$$

The surface is given implicitly by $g(x,y,z) = 0$, so we compute its gradient:

$$\nabla g = \left\langle \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z} \right\rangle = (3x^2, 4y^3, -e^z).$$

Hence,

$$\frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = \frac{\nabla g}{\frac{\partial g}{\partial z}} = \frac{\langle 3x^2, 4y^3, -e^z \rangle}{-e^z} = \begin{pmatrix} -3x^2 e^{-z} \\ -4y^3 e^{-z} \\ 1 \end{pmatrix}.$$

This normal vector is oriented upwards because its z-component is positive, so we take this as our $\mathbf{n}\,\mathrm{d}S$.

Hence, the flux of $\mathbf{F}$ through S is given by

$$\begin{aligned}
\iint_S \mathbf{F} \cdot \mathbf{n}\,\mathrm{d}S &= \int_{x=0}^{5} \int_{y=0}^{5} \begin{pmatrix} \frac{x}{3} \\ \frac{y}{4} \\ \frac{1}{5} \end{pmatrix} \cdot \begin{pmatrix} -3x^2 e^{-z} \\ -4y^3 e^{-z} \\ 1 \end{pmatrix} \mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{5} \int_{y=0}^{5} \left(\frac{x^3+y^4}{e^z} + \frac{1}{5} \right) \mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{5} \int_{y=0}^{5} \left(\frac{-e^z}{e^z} + \frac{1}{5} \right) \mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{5} \int_{y=0}^{5} -\frac{4}{5}\,\mathrm{d}y\,\mathrm{d}x \\
&= 25 \cdot -\frac{4}{5} = \boxed{-20}.
\end{aligned}$$

§51.2 Solution to Exercise 39.1 (sealing a surface)

Exercise 39.1. Let $\mathcal{S}$ be the part of the surface $z = e^{x^2+y^2}$ where $z \le e$, with normal vector oriented downwards. Let $\mathbf{F}(x, y, z) = \langle \cos(z^2) - x, \sin(z^2) - y, 2z \rangle$. Compute the flux of $\mathbf{F}$ through $\mathcal{S}$. (Recommended approach: sealing.)

The divergence of $\mathbf{F}$ is

$$\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(\cos(z^2) - x) + \frac{\partial}{\partial y}(\sin(z^2) - y) + \frac{\partial}{\partial z}(2z) = (-1) + (-1) + 2 = 0.$$

We seal the region $\mathcal{S}$ by adding $\mathcal{S}_{\text{lid}}$, the surface of points with $z = e$ and $x^2 + y^2 \le 1$. This encloses a closed volume $\mathcal{T}$. We orient the normal vector pointing upward (away from $\mathcal{T}$). Note $\mathcal{S}$ also has normal vector pointing away from $\mathcal{T}$.

The divergence theorem on $\mathcal{S}$ and $\mathcal{S}_{\text{lid}}$, enclosing $\mathcal{T}$, now gives

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} + \iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \,\mathrm{d}V = 0.$$

Thus, we compute the flux through the disk and use it to determine the flux through $\mathcal{S}$.

We can take $\mathbf{n}\,\mathrm{d}S = \langle 0, 0, 1 \rangle$ for the flat surface $\mathcal{S}_{\text{lid}}$, so

$$\begin{aligned} \iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \langle 0, 0, 1 \rangle \,\mathrm{d}A &= \iint_{\mathcal{S}_{\text{lid}}} 2z \,\mathrm{d}A \\ &= \iint_{x^2+y^2 \le 1} 2e \,\mathrm{d}A \\ &= 2e \operatorname{Area}(x^2 + y^2 \le 1) = 2e\pi. \end{aligned}$$

Hence the answer:

$$\iint_{\mathcal{S}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = -\iint_{\mathcal{S}_{\text{lid}}} \mathbf{F} \cdot \mathrm{d}\mathbf{S} = \boxed{-2\pi e}.$$

§51.3 Solution to Exercise 39.2 (gravity)

Exercise 39.2. Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ are two closed surfaces that don't intersect and such that $\mathcal{S}_2$ is contained inside $\mathcal{S}_1$. Orient both surfaces outwards. Let O be a point contained inside $\mathcal{S}_2$. Consider the force of gravity $\mathbf{G}$ exerted by a point mass of mass m at O. Show that

$$\oiint_{\mathcal{S}_1} \mathbf{G} \cdot \mathbf{n} \,\mathrm{d}S = \oiint_{\mathcal{S}_2} \mathbf{G} \cdot \mathbf{n} \,\mathrm{d}S.$$

This basically follows from Exercise 32.1 which told us $\mathbf{G}$ has divergence 0. See the cartoon in Figure 125.

We consider the solid volume $\mathcal{T}$ contained between $\mathcal{S}_1$ and $\mathcal{S}_2$. Since $\mathcal{S}_2$ has normal vector oriented toward $\mathcal{T}$ while $\mathcal{S}_1$ has normal vector oriented away from $\mathcal{T}$, the divergence theorem says that

$$\iint_{\mathcal{S}_2} \mathbf{G} \cdot \mathbf{n} \,\mathrm{d}S - \iint_{\mathcal{S}_1} \mathbf{G} \cdot \mathbf{n} \,\mathrm{d}S = \iiint_{\mathcal{T}} \nabla \cdot \mathbf{G} \,\mathrm{d}V.$$

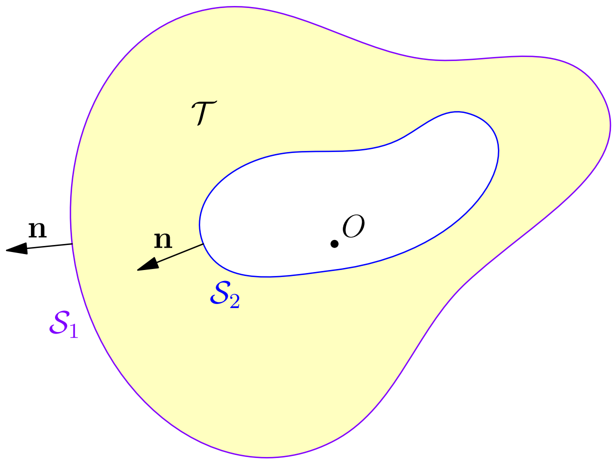

Figure 125: Cartoon of Exercise 39.2.

However, $\mathcal{T}$ does not contain the point O. Therefore, applying Exercise 32.1, we have

$$\iiint_{\mathcal{T}} \nabla \cdot \mathbf{G} \, \mathrm{d}V = \iiint_{\mathcal{T}} 0 \, \mathrm{d}V = 0.$$

The proof is complete.

§51.4 Solution to Exercise 39.3 (divergence to Green for flux)

Exercise 39.3 (*). Prove Green's theorem for flux by quoting the divergence theorem.

That is, suppose $\mathbf{F} = \begin{pmatrix} p \\ q \end{pmatrix}$ is a vector field in $\mathbb{R}^2$ and $\mathcal{C}$ is a closed loop enclosing a region $\mathcal{R}$ counterclockwise. Find a way to use the divergence theorem to prove

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}s = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A.$$

This is covered in Section 20.2 of Poonen's notes. We give the same solution here.

Instead of working in two dimensions, we extend $\mathcal{R}$ into a three-dimensional region $\mathcal{T}$ that consists of a slab of height 1:

$$\mathcal{T} = \{(x, y, z) \mid (x, y) \text{ in } \mathcal{R}, 0 \le z \le 1\}.$$

To put this in words, imagine if we had one of those cool printers that schools use that can make a stack of hundreds of copies of the same image in a moment. Then we print a bunch of copies of $\mathcal{R}$ until the stack of photocopies has height 1; then $\mathcal{T}$ denotes the part of the papers that has ink within it. See Figure 126.

Now the boundary of $\mathcal{T}$ consists of three parts in its surface:

- The top lid $\mathcal{S}_{\text{top}}$, which looks like a copy of $\mathcal{R}$ at height $z = 1$.
- The bottom lid $\mathcal{S}_{\text{bottom}}$, which looks like a copy of $\mathcal{R}$ at height $z = 0$;
- The curved part $\mathcal{S}$ which look like walls of $\mathcal{T}$.

We orient all three away from $\mathcal{T}$.

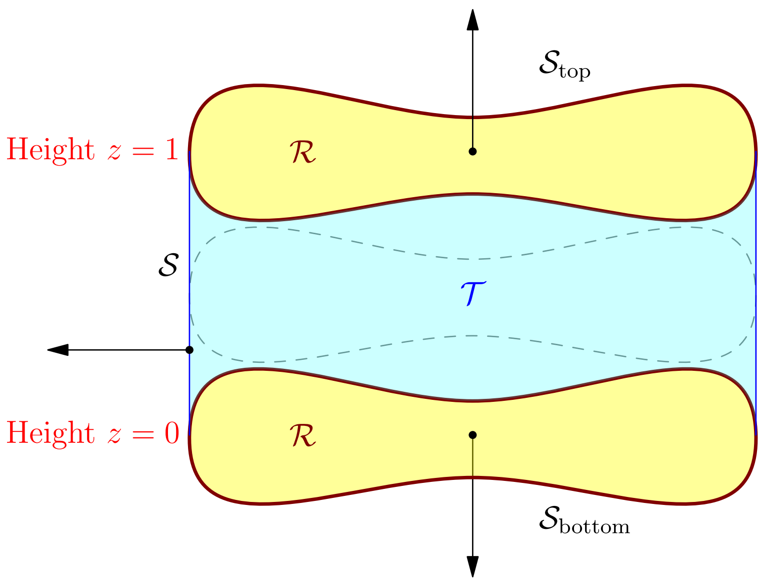

Figure 126: We make a stack of copies of $\mathcal{R}$ of height 1. This produces a solid volume $\mathcal{T}$.

Next, we define a new three-dimensional vector field based on our 2D field $\mathbf{F}$:

$$\mathbf{F}^* = \langle p, q, 0 \rangle.$$

The divergence of $\mathbf{F}^*$ is:

$$\nabla \cdot \mathbf{F}^* = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}.$$

Applying the divergence theorem to $\mathcal{T}$, we obtain:

$$\iint_{\mathcal{S}_{\text{top}}} \mathbf{F}^* \cdot \mathbf{n} \, \mathrm{d}S + \iint_{\mathcal{S}_{\text{bottom}}} \mathbf{F}^* \cdot \mathbf{n} \, \mathrm{d}S + \iint_{\mathcal{S}} \mathbf{F}^* \cdot \mathbf{n} \, \mathrm{d}S = \iiint_{\mathcal{T}} \nabla \cdot \mathbf{F}^* \, \mathrm{d}V.$$

Since $\mathcal{T}$ has height 1, the volume integral simplifies to:

$$\begin{aligned} \iiint_{\mathcal{T}} \nabla \cdot \mathbf{F}^* \, \mathrm{d}V &= \int_{z=0}^{1} \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A \, \mathrm{d}z \\ &= \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A. \end{aligned}$$

Now let's look at the three parts of the boundary of $\mathcal{T}$:

- On the top face at $z = 1$, the outward normal is $\mathbf{n} = \langle 0, 0, 1 \rangle$, and since $\mathbf{F}^* = \langle p, q, 0 \rangle$, we have:

  $$\mathbf{F}^* \cdot \langle 0, 0, 1 \rangle = 0.$$

  Thus, there is no contribution from the top face.
- Similarly the bottom face $z = 0$ gives no contribution.
- The vertical sidewalls project exactly onto $\mathcal{C}$, the boundary of $\mathcal{R}$. The normal to these walls is $\mathbf{n}$ in the xy-plane, so the surface element is $\mathbf{n} \, \mathrm{d}S = \mathbf{n} \, \mathrm{d}s \, \mathrm{d}z$. The flux contribution from these sidewalls is:

  $$\oint_{\mathcal{C}} \int_{z=0}^{1} \mathbf{F}^* \cdot \mathbf{n} \, \mathrm{d}z \, \mathrm{d}s.$$

Since $\mathbf{F}^*$ does not depend on z, this simplifies to:

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}s.$$

Putting this all together we get

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{n}\, \mathrm{d}s = \iint_{\mathcal{R}} \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \right) \mathrm{d}A.$$

This completes the proof of Green's theorem for flux using the divergence theorem.

§51.5 Solution to Exercise 40.1 (Stokes to Green for work)

Exercise 40.1. Prove Green's theorem for work by quoting classical Stokes' theorem.

That is, suppose $\mathbf{F} = \binom{p}{q}$ is a vector field in $\mathbb{R}^2$ and $\mathcal{C}$ is a closed loop enclosing a region $\mathcal{R}$ counterclockwise. Find a way to use classical Stokes' theorem to prove

$$\oint_{\mathcal{C}} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \iint_{\mathcal{R}} \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) \mathrm{d}A.$$

This is covered in Section 20.1 of Poonen's notes. It's actually easier than Exercise 39.3, because one doesn't need to print a whole bunch of copies of $\mathcal{R}$ — just one.

Specifically, we'll just take a single copy of $\mathcal{C}$ on the plane $z = 0$. That's all — it's completely flat. Then we take our surface $\mathcal{S}$ to be just the flat $\mathcal{R}$, again, still contained in $z = 0$. And we take the force field to be

$$\mathbf{F}^* = \langle p, q, 0 \rangle$$

as before.

The classical Stokes' theorem now states that for a surface $\mathcal{S}$ with boundary $\mathcal{C}$,

$$\oint_{\mathcal{C}} \mathbf{F}^* \cdot \mathrm{d}\mathbf{r} = \iint_{\mathcal{S}} \nabla \times \mathbf{F}^* \cdot \mathbf{n}\, \mathrm{d}S.$$

The left-hand side is the same as $\oint_{\mathcal{C}} \mathbf{F} \cdot \mathbf{r}$ in 2D: the work doesn't change if we add an extra dimension. So we just evaluate the curl on the right-hand side:

$$\nabla \times \mathbf{F}^* = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ p & q & 0 \end{vmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \end{pmatrix}.$$

As the surface $\mathcal{S}$ lies in the plane $z = 0$, so the unit normal to $\mathcal{S}$ is $\mathbf{n} = \langle 0, 0, 1 \rangle$. So, the dot product is

$$(\nabla \times \mathbf{F}^*) \cdot \mathbf{n} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}.$$

Plugging this into the right-hand side of our Stokes' theorem application ends the proof.

Chapter 52. Solutions to Part Kilo

Statements of the exercises are in Chapter 41. Right now we only have solutions to the "mock half-final" here. If you want to submit a pull request for the remaining problems, you'd be welcome to.

§52.1 Solution to Exercise 41.1 (cube roots)

We'll just find all the answers to $z^3 = -1000i$ as in Section 10.5 and then identify the relevant one. In polar form, write

$$z^3 = -1000i = 1000 \cdot (\cos 270^\circ + i \sin 270^\circ).$$

Take the cube roots with the standard recipe: the magnitudes should be $\sqrt[3]{1000} = 10$ and the arguments should start from $\frac{270^\circ}{3} = 90^\circ$ and be spaced 120° apart. That is, the three cube roots should be

$$\begin{aligned} z_1 &= 10(\cos 90^\circ + i \sin 90^\circ) \\ z_2 &= 10(\cos 210^\circ + i \sin 210^\circ) \\ z_3 &= 10(\cos 330^\circ + i \sin 330^\circ). \end{aligned}$$

Of these three answers, we want the one whose real and imaginary part are both negative. Only z_2 works; in rectangular form it is

$$z_2 = 10\left(-\frac{\sqrt{3}}{2} - i\frac{1}{2}\right) = \boxed{-5\sqrt{3} - 5i}$$

(and this is the only possible example).

§52.2 Solution to Exercise 41.2 (one of two eigenvalues)

Solution with bare-hands

Given that 2 is an eigenvalue of M, we should have $\det(M - 2I) = 0$. Write

$$0 = \det(M - 2I) = \begin{vmatrix} 1-2 & 1 \\ a & 6-2 \end{vmatrix} = \begin{vmatrix} -1 & 1 \\ a & 4 \end{vmatrix} = -4 - a \implies \boxed{a = -4}.$$

Now let's go back to the characteristic polynomial again. The eigenvalues are those λ such that

$$0 = \begin{vmatrix} 1-\lambda & 1 \\ -4 & 6-\lambda \end{vmatrix} = (1-\lambda)(6-\lambda) + 4 = \lambda^2 - 7\lambda + 10.$$

We know that $\lambda = 2$ is one root of the quadratic; the other one is $\boxed{\lambda = 5}$.

To get the eigenvector, write $\binom{x}{y}$ so that we need

$$M\begin{pmatrix} x \\ y \end{pmatrix} = 5\begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{cases} (1-5)x + y = 0 \\ -4 + (6-5)y = 0 \end{cases} \implies y = 4x.$$

So an eigenvector for 5 is $\boxed{\begin{pmatrix} 1 \\ 4 \end{pmatrix}}$ (or any nonzero multiple of it).

Solution using trace and determinant shortcut

Let λ_2 be the other eigenvalue. If you happen to remember that the trace is the sum of the eigenvalues while the determinant was the product of the eigenvalues (Section 9.7), then this question can be done even more quickly:

$$\lambda_2 + 2 = \operatorname{Trace} M = 1 + 6$$
$$\lambda_2 \cdot 2 = \det M = \begin{vmatrix} 1 & 1 \\ a & 6 \end{vmatrix} = 6 - a.$$

The first equation implies $\lambda_2 = 5$; then the second implies $a = -4$. The eigenvector is then recovered in the same way as the first solution.

§52.3 Solution to Exercise 41.3 (plane)

We start by determining the equation of the plane through $P_1 = (b, 0, 0)$, $P_2 = (0, b, 0)$ and $P_3 = (0, 0, b)$. You might be able to guess the equation just by looking, but if you didn't see it, you could also use the cross product

$$(P_2 - P_1) \times (P_3 - P_1) = \begin{pmatrix} -b \\ b \\ 0 \end{pmatrix} \times \begin{pmatrix} -b \\ 0 \\ b \end{pmatrix} = \begin{pmatrix} b^2 \\ b^2 \\ b^2 \end{pmatrix} = b^2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Therefore, $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is a normal vector to the plane, so the plane's equation should be $x + y + z = \text{const}$. The plane should pass through $(b, 0, 0)$ and so on; hence the plane's equation is

$$x + y + z = b.$$

In order for this to also pass through $(2, 3, 6)$, we need $b = 2 + 3 + 6 = \boxed{11}$.

It remains to calculate the distance from $(0, 0, 0)$ to the plane $x + y + z = 11$. If you use the point-to-plane formula (Section 5.6) for this, you get

$$\frac{|1 + 2 + 3 - 11|}{\sqrt{1^2 + 1^2 + 1^2}} = \boxed{\frac{5}{\sqrt{3}}}.$$

If you didn't remember this formula, you should instead compute the length of the projection of the vector $\mathbf{v} = (1, 2, 3) - (2, 3, 6) = \langle -1, -1, -3 \rangle$ (you can replace $(2, 3, 6)$ with any other point on the plane, like $(0, 0, 11)$ or similar) along the direction of $\mathbf{n} = \langle 1, 1, 1 \rangle$. Doing this by hand gives

$$\frac{\mathbf{v} \cdot \mathbf{n}}{|\mathbf{n}|} = \frac{\langle -1, -1, -3 \rangle \cdot \langle 1, 1, 1 \rangle}{\sqrt{3}} = -\frac{5}{\sqrt{3}}$$

like before, although as I've described before, we're really just repeating the proof of the point-to-plane formula.

§52.4 Solution to Exercise 41.4 (level curves through critical points)

Let $f(x, y) = \cos(x) + \sin(y)$. The gradient is given by

$$\nabla f = \langle -\sin(x), \cos(y) \rangle.$$

So a critical point occurs at any point for which $\sin(x) = \cos(y) = 0$. (These are the points where $\cos(x) = \pm 1$ and $\sin(y) = \pm 1$.)

Saddle point

To identify a saddle point, we compute the double derivatives:

$$\begin{aligned} f_{xx} &= -\cos(x) \\ f_{xy} &= 0 \\ f_{yy} &= -\sin(y). \end{aligned}$$

It's enough to pick any (x, y) for which f_{xx} and f_{yy} have opposite sign. One example would be $(x, y) = \boxed{\left(0, \frac{3\pi}{2}\right)}$, among many others. At this value we get $f\left(0, \frac{3\pi}{2}\right) = 0$.

In fact, the complete list of saddle points is given as follows: whenever m and n are integers where $m + n$ is odd, the point

$$(x, y) = \left(m\pi, \left(n + \frac{1}{2}\right)\pi\right)$$

is a saddle point, and these are all saddle points. The previous example was the special case $m = 0$ and $n = 1$.

The level curve of f is the set of points (x, y) with $\cos(x) + \sin(y) = 0$, so in fact every saddle point lies on this level curve. In Figure 127, we draw the level curve below in blue, and the saddle points in red. Since $\cos(x) = \sin(y)$ whenever $x \pm y + \frac{\pi}{2}$ is a multiple of 2π, the level curves are a mesh of lines running through the plane at diagonals.

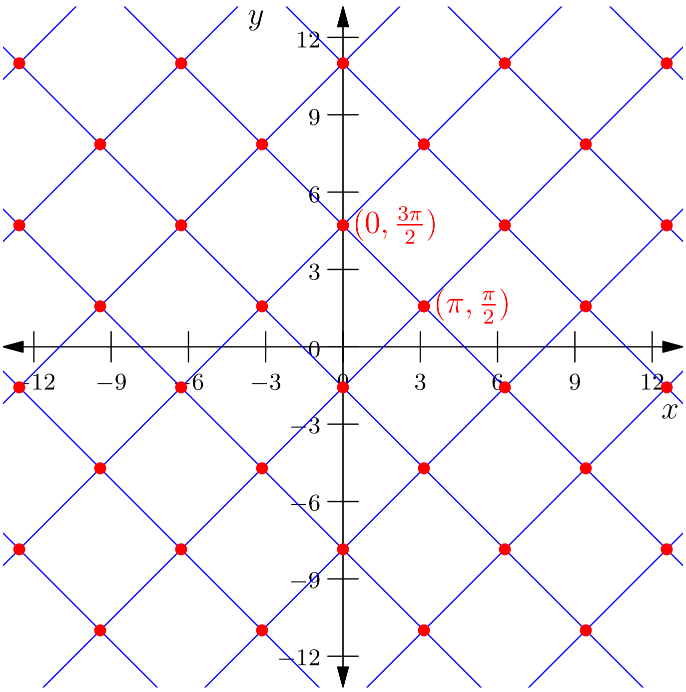

Figure 127: The level curve of $f = \cos(x) + \sin(y) = 0$. Saddle points are marked in red, and these are all the saddle points.

Local maximum

An example of a local maximum would be $\boxed{0, \frac{\pi}{2}}$, at which $f\left(0, \frac{\pi}{2}\right) = \cos 0 + \sin\left(\frac{\pi}{2}\right) = 2$. In fact, the level curve of $f(x, y) = 2$ passes through all the local maximums, which occur only when $\cos x = \sin y = 1$, meaning x and $y - \frac{\pi}{2}$ are integer multiples of 2π. So the level curve of f for the value 2 contains *only* a disjointed set of points, as shown in Figure 128.

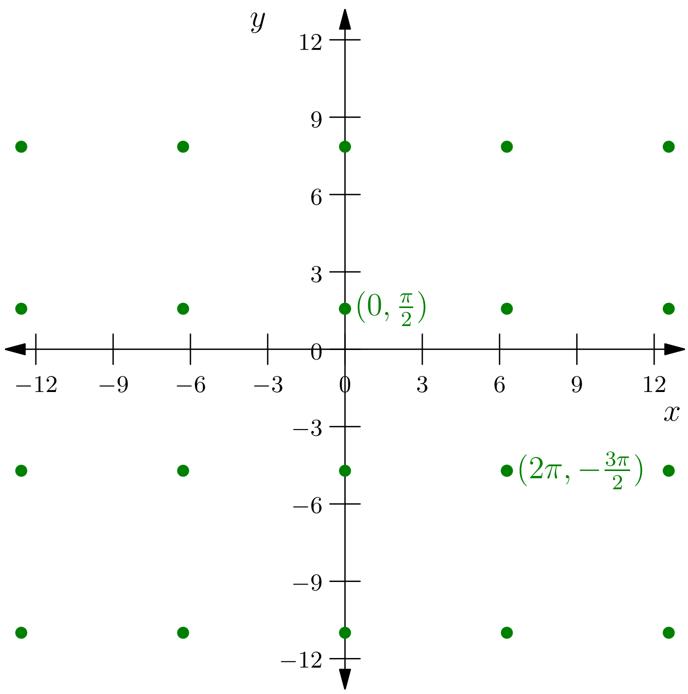

Figure 128: The level curve of $f = \cos(x) + \sin(y) = 2$, in green. Each green point is a local maximum of f (and these are the only local maximums).

§52.5 Solution to Exercise 41.5 (optimization)

Let $f(x, y) = x^2 + 2y^2 + 4x$. Let $\mathcal{R}$ denote the region $x^2 + y^2 \le 9$, which is 2D, has no limit cases, and boundary $x^2 + y^2 = 9$. We carry out our optimization recipe from Chapter 19.

0. There are no limit cases, but a boundary $x^2 + y^2 = 9$, a circle of radius 3.
1. First let's find the critical points of $f(x, y, z) = x^2 + 2y^2 + 4x$. Write

$$\nabla f = \begin{pmatrix} 2x + 4 \\ 4y \end{pmatrix}.$$

   The only point at which $\nabla f = 0$ is $(-2, 0)$, at which

$$f(-2, 0) = -4.$$

2. The boundary of $\mathcal{R}$ is $x^2 + y^2 = 9$. We use Lagrange multipliers on the boundary, which we denote $\mathcal{S}$, with constraint $g(x, y) = x^2 + y^2 = 9$.
   0. The new region $\mathcal{S}$ has no boundary and no limit cases.
   1. Let's find the LM-critical points for f on $\mathcal{S}$. Take the gradient of g to get

$$\nabla g = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

      The only point at which $\nabla g = \mathbf{0}$ is $x = y = 0$ which isn't on $\mathcal{S}$, so we don't have to worry about $\nabla g = \mathbf{0}$ the case. Now we instead solve

$$\begin{pmatrix} 2x + 4 \\ 4y \end{pmatrix} = \lambda \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

      The second equation says

$$4y = \lambda 2y \implies \lambda = 2 \quad \text{or} \quad y = 0.$$

If $y = 0$, we get the points $(3,0)$ and $(-3,0)$ which we need to check. We have

$$f(3,0) = 21$$
$$f(-3,0) = -3.$$

Now suppose instead $\lambda = 2$. Then $2x + 4 = 4x \implies x = 2$, and hence $y = \pm\sqrt{5}$. We check those points

$$f\left(2, \sqrt{5}\right) = 22$$
$$f\left(2, -\sqrt{5}\right) = 22.$$

2. $\mathcal{S}$ has no boundary to consider.
3. $\mathcal{S}$ has no limit cases to consider.

3. $\mathcal{R}$ has no limit cases to consider.

Of the five points we've checked, $\boxed{f(-2,0) = -4}$ and $\boxed{f\left(2, \pm\sqrt{5}\right) = 22}$ give the optimal values.

§52.6 Solution to Exercise 41.6 (triple integral)

The region being integrated over can be succinctly described as

$$\mathcal{R} = \{0 \le x \le y \le z \le 1\}.$$

Swap the order of integration so that z is outermost:

$$\begin{aligned}
\int_{x=0}^{1} \int_{y=x}^{1} \int_{z=y}^{1} e^{z^3} \, dz \, dy \, dx &= \int_{z=0}^{1} \int_{y=0}^{z} \int_{x=0}^{y} e^{z^3} \, dx \, dy \, dz \\
&= \int_{z=0}^{1} e^{z^3} \int_{y=0}^{z} \int_{x=0}^{y} 1 \, dx \, dy \, dz \\
&= \int_{z=0}^{1} e^{z^3} \int_{y=0}^{z} y \, dy \, dz \\
&= \int_{z=0}^{1} e^{z^3} \frac{z^2}{2} \, dy \, dz \\
&= \frac{1}{6} \int_{z=0}^{1} e^{z^3} 3z^2 \, dy \, dz \\
&= \frac{1}{6} \left[e^{z^3}\right]_{z=0}^{1} = \boxed{\frac{e-1}{6}}.
\end{aligned}$$

§52.7 Solution to Exercise 41.7 (curl)

The curl of $\mathbf{F}$ can be computed as

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 7\cos(x) & \cos(y)\cos(2z) & c\sin(y)\sin(2z) \end{vmatrix} = \begin{pmatrix} c\cos(y)\sin(2z) - \cos(y) \cdot (-2\sin(2z)) \\ 0 \\ 0 \end{pmatrix}$$

which is identically zero only for $\boxed{c=-2}$. For that value of c, we can recover a potential function f by writing

$$\begin{aligned}
\frac{\partial f}{\partial x} &= 7\cos(x) \implies f = 7\sin(x) + C_1(y,z) \\
\frac{\partial f}{\partial y} &= \cos(y)\cos(2z) \implies f = \sin(y)\cos(2z) + C_2(z,x) \\
\frac{\partial f}{\partial z} &= -2\sin(y)\sin(2z) \implies f = \sin(y)\cos(2z) + C_3(x,y).
\end{aligned}$$

Hence, the potential function can be extracted:

$$f(x,y,z) = 7\sin(x) + \sin(y)\cos(2z).$$

For a curve $\mathcal{C}$ starting at P and ending at Q, we have

$$\int_{\mathcal{C}} f \cdot \mathrm{d}\mathbf{r} = f(Q) - f(P).$$

However, since both the trig functions sin and cos take values in $[-1,1]$, it's easy to see that $\max f = 8$ (for example $f(\frac{\pi}{2}, \frac{\pi}{2}, 0) = 8$) while $\min f = -8$ (for example $f(-\frac{\pi}{2}, -\frac{\pi}{2}, 0) = -8$). Hence the largest possible value of the line integral is $8-(-8) = \boxed{16}$.

§52.8 Solution to Exercise 41.8 (flux)

The divergence is $\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(x + e^y + z^3) + \frac{\partial}{\partial y}(e^x + y + z^3) + \frac{\partial}{\partial z} z = 1 + 1 + 1 = 3$.

Given $\mathcal{S}$, we add two lids, $\mathcal{S}_{\text{top}}$ and $\mathcal{S}_{\text{bottom}}$. The top lid is the flat surface given by $z = 9$ and $x^2 + y^2 \le 100$, with normal vector oriented upwards. The bottom lid is the flat surface given by $z = 7$ and $x^2 + y^2 \le 100$, with normal vector oriented outwards. Finally, let $\mathcal{T}$ denote the cylinder $1 \le z \le 2$ and $x^2 + y^2 \le 100$, which is enclosed by $\mathcal{S}$, $\mathcal{S}_{\text{top}}$, $\mathcal{S}_{\text{bottom}}$. Then the divergence theorem states that

$$\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V = \iint_{\mathcal{S}_{\text{top}}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S + \iint_{\mathcal{S}_{\text{bottom}}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S + \iint_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S.$$

The fourth quantity is the flux we want, so our strategy is to calculate the first three quantities.

The divergence is straightforward because its constant:

$$\iiint_{\mathcal{T}} \nabla \cdot \mathbf{F} \, \mathrm{d}V = \iiint_{\mathcal{T}} 3 \, \mathrm{d}V = 3 \operatorname{Vol}(\mathcal{T}) = 3 \cdot 100\pi \cdot 2 = 600\pi$$

(the volume of a cylinder with height 2 and base of area 100π).

For the top lid, we recall that for a flat surface parallel to the xy-plane, we have $\mathbf{n} \, \mathrm{d}S = \pm\langle 0,0,1\rangle \, \mathrm{d}x \, \mathrm{d}y$. For the top lid, we thus have

$$\begin{aligned}
\iint_{\mathcal{S}_{\text{top}}} \mathbf{F} \cdot \mathbf{n} \, \mathrm{d}S &= \iint_{\mathcal{S}_{\text{top}}} \langle x + e^y + 729, e^x + y + 729, 9\rangle \cdot \langle 0,0,1\rangle \, \mathrm{d}x \, \mathrm{d}y \\
&= \iint_{\mathcal{S}_{\text{top}}} 9 \, \mathrm{d}x \, \mathrm{d}y \\
&= 9 \operatorname{Area}(\mathcal{S}_{\text{top}}) = 900\pi.
\end{aligned}$$

For the bottom lid, we instead have

$$\begin{aligned}\iint_{\mathcal{S}_{\text{bottom}}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S &= \iint_{\mathcal{S}_{\text{bottom}}} \langle x+e^y+343, e^x+y+343, 7\rangle \cdot \langle 0,0,-1\rangle \,\mathrm{d}x\,\mathrm{d}y \\ &= \iint_{\mathcal{S}_{\text{bottom}}} (-7)\,\mathrm{d}x\,\mathrm{d}y \\ &= -7\operatorname{Area}(\mathcal{S}_{\text{bottom}}) = -700\pi.\end{aligned}$$

Hence, the quantities in the divergence theorem become

$$600\pi = 900\pi - 700\pi + \iint_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S$$

so $\iint_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}S = \boxed{400\pi}$.

Part Mike: Appendix

Nothing past this point is for exam, obviously.

Chapter 53. If you are thinking of majoring in math...

During the course, one of the students asked me about academic advice saying they wanted to become a math major at MIT. If that also describes you, here's what I told them. The course numbers here are with respect to MIT, but this advice should hold equally well at other universities.

§53.1 The two starter topics are algebra and analysis, not calculus

It may come as a surprise to you that 18.02 isn't a prerequisite, even indirectly, for most upper-division math classes (18.xyz for $x \geq 1$). The two most important areas to take in pure math are **18.100** (real analysis) and **18.701–18.702** (algebra); these are sort of the barrier between the world of pre-university math and serious math. Once you clear these two classes, the floodgates open and the world of modern math is yours to explore (see the dependency chart in the Napkin for more on this).

For example, if you take 18.701, the instructor will literally *throw away* the "definitions" of linear transformations (and others) you learned in 18.02 and replace them with the "correct" ones. You've seen me do this already. Similarly, you will have new rigorous definitions of derivatives and integrals. In some sense, 18.100 is really *redoing* all of 18.01 and 18.02 with actual proofs.

§53.2 Proof-writing

A prerequisite to both 18.100 (real analysis) and 18.701−18.702 (algebra) isn't any particular theory, but **proof experience**, and that's the biggest priority if you don't have that yet. (And I don't mean two-column proofs in 9th grade geometry. Two-column proofs were something made up for K-12 education and never used again.)

At MIT, I've been told in recent years there's a class called 18.090 for this. This class is new enough I don't even have any secondhand accounts, but if Poonen is on the list of instructors who developed the course, I trust him. If you're at a different school, my suggestion would be to ask any of the math professors a question along the lines of "I'd like to major in math, but I don't have proof experience yet. Which class in your department corresponds to learning proof arguments?". They should know exactly what you're talking about.

Alternatively, if you are willing to study proof-writing independently, the FAQ `https://web.evanchen.cc/faq-contest.html#C-5` on my website has some suggestions. In particular, if you're a textbook kind of person, the book I used growing up was Rotman's *Journey into Math: An Introduction to Proofs*, available at `https://store.doverpublications.com/products/9780486453064` it worked well for me. I'm sure there are other suitable books as well.

§53.3 The three phases of math education (from Tao's blog)

Let me put proof-writing into the bigger framework. Terence Tao, on his blog, describes a division of mathematical education into three stages. The descriptions that follows are copied verbatim from that link:

1. The "pre-rigorous" stage, in which mathematics is taught in an informal, intuitive manner, based on examples, fuzzy notions, and hand-waving. (For instance, calculus is usually first introduced in terms of slopes, areas, rates of change, and so forth.) The emphasis is more on computation than on theory.
2. The "rigorous" stage, in which one is now taught that in order to do maths "properly", one needs to work and think in a much more precise and formal manner (e.g. re-doing calculus by using

epsilons and deltas all over the place). The emphasis is now primarily on theory; and one is expected to be able to comfortably manipulate abstract mathematical objects without focusing too much on what such objects actually "mean".

3. The "post-rigorous" stage, in which one has grown comfortable with all the rigorous foundations of one's chosen field, and is now ready to revisit and refine one's pre-rigorous intuition on the subject, but this time with the intuition solidly buttressed by rigorous theory. (For instance, in this stage one would be able to quickly and accurately perform computations in vector calculus by using analogies with scalar calculus, or informal and semi-rigorous use of infinitesimals, big-O notation, and so forth, and be able to convert all such calculations into a rigorous argument whenever required.) The emphasis is now on applications, intuition, and the "big picture".

These notes are still in the first stage. The introduction-to-proofs class at your school will essentially be the beginning of the second stage.

Chapter 54. Proofs of the dot product property

§54.1 Deriving the geometric definition of dot product from the algebraic one

This proof is short, but harder to come up with.

We have two definitions in play and we want to show they coincide, which makes notation awkward. So in what follows, our notation $\mathbf{u} \cdot \mathbf{v}$ will always refer to the *algebraic* definition; and we will *prove* that $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}|\ |\mathbf{v}| \cos\theta$.

The algebraic definition is already enough to tell us that

$$\mathbf{w} \cdot \mathbf{w} = |\mathbf{w}|^2 \tag{18}$$

by the Pythagorean theorem: if $\mathbf{w} = \langle a_1, ..., a_n \rangle$ then both sides equal $a_1^2 + ... + a_n^2$.

Let C denote the origin, and let A and B denote the endpoints of $\mathbf{u}$ and $\mathbf{v}$ when we draw them as arrows emanating from the origin. Hence $\mathbf{v} - \mathbf{u}$ is a vector pointing from A to B.

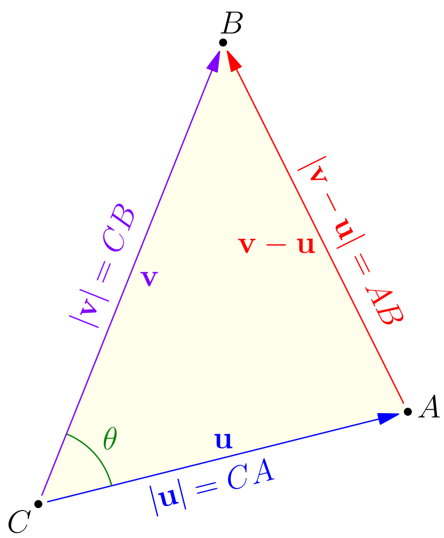

Figure 129: We use the law of cosines on triangle ABC together with three applications of Equation 18 to show the geometric definition of dot product.

We now use Equation 18 three times as follows:

$$\begin{aligned} AB^2 &= (\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u}) \\ &= \mathbf{v} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{u} - 2\mathbf{u} \cdot \mathbf{v} \\ &= CA^2 + CB^2 - 2\mathbf{u} \cdot \mathbf{v}. \end{aligned}$$

However, the law of cosines on triangle ABC also tells us that

$$AB^2 = CA^2 + CB^2 - 2(CA)(CB)\cos\theta$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Setting the two equations for AB^2 equal gives

$$\mathbf{u} \cdot \mathbf{v} = (CA)(CB)\cos\theta.$$

As $|\mathbf{u}| = CA$ and $|\mathbf{v}| = CB$, the proof is complete.

§54.2 Deriving the algebraic definition of dot product from the geometric one

The proof in Section 54.1 might seem magical. Indeed, it's so short because it's cheating in some way: it starts with the algebraic definition. But if you've never seen the dot product before, that algebraic definition is unnatural; you wouldn't have any idea to write the expression $a_1 b_1 + \ldots + a_n b_n$. So in this section we give a proof that *starts* from the geometric formula and shows how you would come up with $a_1 b_1 + \ldots + a_n b_n$.

So this time our convention is flipped from Section 54.1: in what follows, our notation $\mathbf{u} \cdot \mathbf{v}$ will always refer to the *geometric* definition; that is $\mathbf{u} \cdot \mathbf{v} := |\mathbf{u}|\ |\mathbf{v}| \cos\theta$. And our goal is to show that it matches the algebraic definition.

We will assume that $|\mathbf{u}| = 1$ (i.e. $\mathbf{u}$ is a unit vector) so that $\mathbf{u} \cdot \mathbf{v}$ is the length of the projection of $\mathbf{v}$ onto $\mathbf{u}$. This is OK to assume because in the general case one just scales everything by $|\mathbf{u}|$.

Easy special case

As a warmup, try to show that if $\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ is any vector, then $\mathbf{u} \cdot \mathbf{e}_1 = a$. (This is easy. The projection of $\mathbf{u}$ onto $\mathbf{e}_1$ is literally a.)

Main proof

For concreteness, specialize to $\mathbb{R}^2$ and consider $\mathbf{u} \cdot \mathbf{v}$ where $\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ is a unit vector (i.e. $\mathbf{u} = 1$), and $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ is any vector in $\mathbb{R}^2$. Then we want to show that the projection of $\mathbf{v}$ onto $\mathbf{u}$ has length $xa + yb$. See Figure 130.

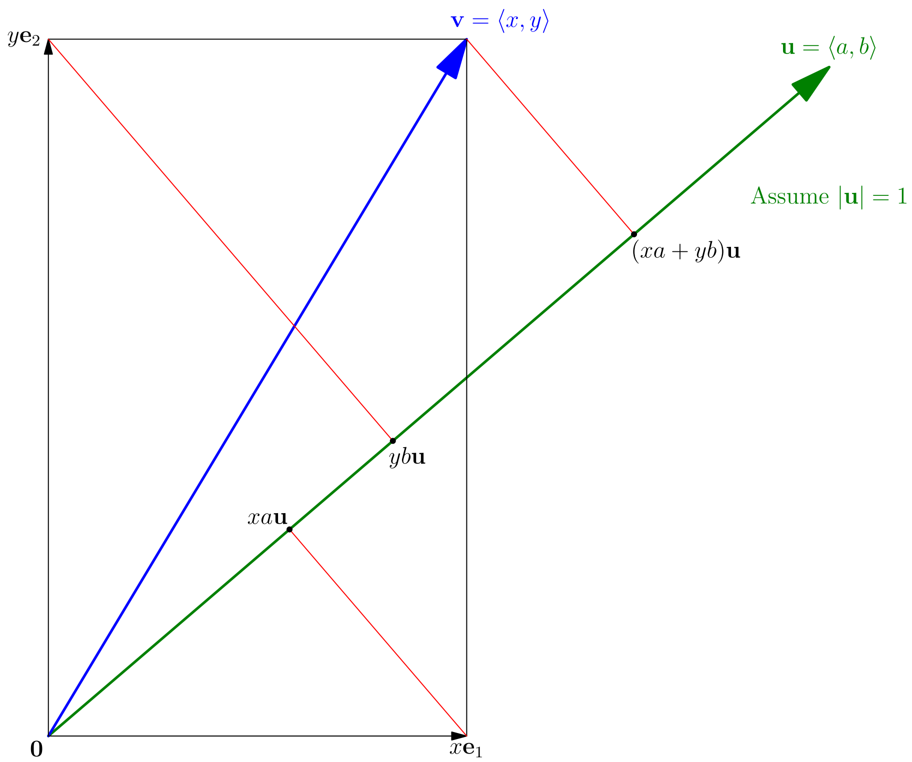

Figure 130: Proof that the dot product is given by the projection

The basic idea is to decompose $\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2$. The length of projection of $\mathbf{v}$ onto $\mathbf{u}$ can be decomposed then into the lengths of projections of $x\mathbf{e}_1$ and $y\mathbf{e}_2$. (To see this, tilt your head so the green line is horizontal; recall that the black quadrilateral is a rectangle, hence also a parallelogram). In other words,

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot (x\mathbf{e}_1 + y\mathbf{e}_2) = x(\mathbf{u} \cdot \mathbf{e}_1) + y(\mathbf{u} \cdot \mathbf{e}_2).$$

But we already did the special cases before:

$$\begin{aligned}\mathbf{u} \cdot \mathbf{e}_1 &= a \\ \mathbf{u} \cdot \mathbf{e}_2 &= b.\end{aligned}$$

Hence, we get the right-hand side is

$$\mathbf{u} \cdot \mathbf{v} = xa + yb,$$

as advertised. In summary, by using the black parallelogram, we were able to split $\mathbf{u} \cdot \mathbf{v}$ into two easy cases we already know how to do.

The same idea will work in $\mathbb{R}^3$ if you use $\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$ instead, and replace the parallelogram with a parallelepiped, in which case one now has 3 easy cases. And so on in n dimensions.

Chapter 55. What does i^i mean?

When learning mathematics, I believe definitions are actually more important than theorems. A lot of confusion comes from not having been given careful definitions of the objects. (See `https://web.evanchen.cc/handouts/NaturalProof/NaturalProof.pdf` for more on that.)

So in general any time you are confused about whether an operation is "legal" — and this is true in all of math, not just 18.02 — **the first thing to really check whether you have been given a precise definition**. The endless Internet debates on whether 0 is even or whether $0.999... = 1$ or whether $\frac{1}{x}$ is a continuous function (hint: yes) are all examples of people who don't know the definitions of objects they're discussing.

§55.1 Real exponents, real base

With that in mind, let's fix $a > 0$ a positive real number and think about what a^r should mean.

Definition 55.1 (18.100 definition)**.**

- When $n > 0$ is an integer, then $a^n := a \cdot ... \cdot a$, where a is repeated n times.
- Then we let $a^{-n} := \frac{1}{a^n}$ for each integer $n > 0$.
- When $\frac{m}{n}$ is a rational number, $a^{\frac{m}{n}}$ means the unique $b > 0$ such that $a^m = b^n$. (In 18.100, one proves this b is unique and does exist.)
- It's less clear what a^x means when $x \in \mathbb{R}$, like $x = \sqrt{2}$ or $x = \pi$. I think usually one takes a limit of rational numbers q close to x and lets $a^x := \lim_{q \to x} a^q$. (In 18.100, one proves this limit does in fact exist.)

§55.2 Complex exponents, real base

But when $z \in \mathbb{C}$, what does a^z mean? There's no good way to do this.

You likely don't find an answer until 18.112, but I'll tell you now. In 18.100 you will also prove that the Taylor series

$$e^x = \sum_{k \geq 0} \frac{r^k}{k!}$$

is correct, where $e := \sum_{k \geq 0} \frac{1}{k!}$ is Euler's constant.

So then when you start 18.112, we will flip the definition on its head:

Definition 55.2 (18.112 definition)**.** If $z \in \mathbb{C}$, we *define*

$$e^z := \sum_{k \geq 0} \frac{z^k}{k!}.$$

Then for $a > 0$, we let $a^z = e^{z \log a}$.

To summarize: in 18.100, we defined exponents in the way you learned in grade school and then proved there was a Taylor series. But in 18.112, you *start* with the Taylor series and *then* prove that the rules in grade school you learned still applied.

And checking this consistency requires work. Because we threw away Definition 55.1, identities like $e^{z_1 + z_2} = e^{z_1} e^{z_2}$ and $(e^{z_1})^{z_2} = e^{z_1 z_2}$ are no longer "free": they have to be proved rigorously too. (To be fair, they need to be proved in 18.100 too, but there it's comparatively easier.) I think you shouldn't be *surprised* they're true; we know it's true for $\mathbb{R}$, so it's one heck of a good guess. But you shouldn't take these on faith. At least get your professor to acknowledge they *require* a (non-obvious) proof, even if you aren't experienced enough to follow the proof yourself yet.

Anyway, if we accept this definition, then Euler's formula makes more sense:

Theorem 55.3 (Euler). *We have*

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

The point is that cosine and sine also have a Taylor series that is compatible with definition:

$$\begin{aligned} \cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \\ \sin(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots. \end{aligned} \tag{19}$$

And if you put these together, you can verify Theorem 55.3, up to some technical issues with infinite sums. I think the professor even showed this in class:

$$\begin{aligned} \cos(\theta) + i\sin(\theta) &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \ldots\right) + \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots\right)i \\ &= 1 + (\theta i) + \frac{(\theta i)^2}{2!} + \frac{(\theta i)^3}{3!} + \frac{\theta i^4}{4!} + \frac{(\theta i)^5}{5!} \\ &= e^{i\theta}. \end{aligned}$$

§55.3 Complex exponents, complex base

But what about i^i? Our Definition 55.2 above only worked for positive real numbers $a > 0$. Here, it turns out you're out of luck. There isn't any way to define i^i in a way that makes internal sense. The problem is that there's no way to take a single log of a complex number, so the analogy with $\log a$ breaks down.

Put another way: there's no good way to assign a value to $\log(i)$, because $e^{i\pi/2} = e^{5i\pi/2} = \ldots$ are all equal to i. You might hear this phrased "complex-valued logarithms are multivalued". You can have some fun with this paradox:

$$\begin{aligned} i = e^{i\pi/2} &\implies i^i = e^{-\pi/2} \\ i = e^{5i\pi/2} &\implies i^i = e^{-5\pi/2}. \end{aligned}$$

Yeah, trouble.

§55.4 Trig functions with complex arguments

On the other hand, $\cos(i)$ can be defined: use the Taylor series Equation 19, like we did for e^z. To spell it out:

Definition 55.4 (18.112 trig definitions). If z is a complex number, we define

$$\begin{aligned} \cos(z) &:= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \ldots \\ \sin(z) &:= z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \ldots. \end{aligned}$$

If you do this, then Definition 55.2 implies the following identities are kosher:

Proposition 55.5. *Under Definition 55.4, we have the identities*

$$\cos(z) := \frac{e^{iz} + e^{-iz}}{2}$$
$$\sin(z) := \frac{e^{iz} - e^{-iz}}{2i}.$$

Proof. If you write out $e^{iz} = \sum \frac{(iz)^k}{k!}$ and $e^{-iz} = \sum \frac{(-iz)^k}{k!}$ and add them, the odd k's cancel and the even k's don't, which gives you

$$e^{iz} + e^{-iz} = 2 - 2 \cdot \frac{z^2}{2!} + 2 \cdot \frac{z^4}{4!} - 2 \cdot \frac{z^6}{6!} + \dots.$$

So dividing by 2, we see $\cos(z)$ on the right-hand side, as needed. The argument with sin is similar, but this time the even k's cancel and you divide by $2i$ instead. □

So for example, from Proposition 55.5, we conclude for example that

$$\cos(i) = \frac{e + \frac{1}{e}}{2}.$$

Strange but true.

§55.5 The future: what are 18.100 and 18.112 anyway?

First I need to tell you what analysis is. When students in USA ask me what analysis is, I sometimes say "calculus but you actually prove things". But that's actually a bit backwards; it turns out that in many parts of the world, there is no topic called "calculus".[29] It would be more accurate to say calculus is analysis with proofs, theorems, and coherent theorem statements deleted, and it only exists in some parts of the world (which is why mathematicians will tend to look down on it).

With that out of the way,

- 18.100 is real analysis, i.e. analysis of functions over $\mathbb{R}$
- 18.112 is complex analysis, i.e. analysis of functions over $\mathbb{C}$.

If you ever take either class, I think the thing to know about them is:

> *Complex analysis is the good twin and real analysis is the evil one: beautiful formulas and elegant theorems seem to blossom spontaneously in the complex domain, while toil and pathology rule the reals.*
>
> — *Charles Pugh, in* Real Mathematical Analysis

[29]See https://web.evanchen.cc/faq-school.html#S-10.

Chapter 56. Saddle point simulation code for Section 17.3

```
import random

random.seed("18.02 Fall 2024")


def classify_critical_points(a3, a2, a1, b3, b2, b1):
    # f = a3 * x**3 + a2 * x**2 + a1 * x + b3 * y**3 + b2 * y**2 + b1 * y
    # the constant term has no effect on the critical points, so we ignore it
    assert a3 != 0 and b3 != 0

    # fx = 3 a3 x^2  + 2 a2 x + 1; fy = 3 b3 y^2  + 2 b2 y + 1
    # If either of these have negative discriminant, rage-quit
    if 4 * a2 * a2 - 12 * a3 * a1 < 0 or 4 * b2 * b2 - 12 * b3 * b1 < 0:
        return (0, 0, 0)

    # Otherwise, let's get the two critical values
    x1 = (-2 * a2 + (4 * a2 * a2 - 12 * a3 * a1) ** 0.5) / (6 * a3)
    x2 = (-2 * a2 - (4 * a2 * a2 - 12 * a3 * a1) ** 0.5) / (6 * a3)
    y1 = (-2 * b2 + (4 * b2 * b2 - 12 * b3 * b1) ** 0.5) / (6 * b3)
    y2 = (-2 * b2 - (4 * b2 * b2 - 12 * b3 * b1) ** 0.5) / (6 * b3)

    local_minima, local_maxima, saddle_points = 0, 0, 0

    for x0 in (x1, x2):
        for y0 in (y1, y2):
            fxx = 6 * a3 * x0 + 2 * a2
            fyy = 6 * b3 * x0 + 2 * b2
            assert fxx != 0 and fyy != 0  # give up lol
            if fxx > 0 and fyy > 0:
                local_minima += 1
            elif fxx < 0 and fyy < 0:
                local_maxima += 1
            else:
                saddle_points += 1
    return (local_minima, local_maxima, saddle_points)


local_minima = 0
local_maxima = 0
saddle_points = 0

N = 10**6
for _ in range(10000):
    a1 = random.randint(-N, N + 1)
    a2 = random.randint(-N, N + 1)
    a3 = random.randint(-N, N + 1)
    b1 = random.randint(-N, N + 1)
    b2 = random.randint(-N, N + 1)
    b3 = random.randint(-N, N + 1)
    u, v, w = classify_critical_points(a3, a2, a1, b3, b2, b1)
    local_minima += u
    local_maxima += v
    saddle_points += w
total = local_minima + local_maxima + saddle_points
print(local_minima / total, local_maxima / total, saddle_points / total, total)
```