Course Notes for Calculus III

Remkes Kooistra
The King's University

August 4, 2019

Contents

Chapter 1

Introduction

Calculus I and II are the study of functions of a single real variable, using the techniques of limits, derivatives, intergrals and series. In Calculus I and II, we built a set of very powerful tools to understand functions, their increase and descrease, their asymptotic growth, their behaviour near undefined points, their maxima and minima, their slopes, and the areas under their graphs. Qualitatively, we can describe and visualize single-variable functions very precisely using the tools of calculus.

However, all the functions studied so far have been functions of *one* real variable with *one* real output. This course starts with the obvious observation that a function can depend on numerous variables and have outputs in several variables. Vector calculus asks this question: What can the techniques of calculus tell us about these multi-variable functions? How do the major ideas of derivatives, integrals and series extend to situations with more than one variable?

The short answer is that they all extend; moreover, they extend in deeply fascinating, mysterious and intricate ways.

Chapter 2

Infinite Series

The first two chapters of these notes, covering Infinite Series and Taylor Series, cover the same as the final two chapters of the Calculus II notes. Much of what we cover is review and most of the these notes are identical to the Calculus II notes, but we will find a new focus on Taylor series and new material on approximation and error analysis.

There are three classical branches of calculus. The first two, derivatives and integrals, command the vast majority of the time and energy in most first year calculus classes. In many universities, these two topics are the entire course. However, there is a third branch of the calculus which deserves equal attention: infinite series.

In some ways, the problem of infinite series is older than the problems motivating derivatives and integrals. Issues of infinite series go back to at least early Greek mathematics, where thinkers struggled with the puzzle known as Zeno's Paradox.

There are many forms of Zeno's Paradox; I will present one relatively common version. If you wish to travel from point a to point b, then first you must travel half-way. Having gone halfway to b, you must again cover half the remaining distance. Having gone 3/4 of the way to b, there is still a distance remaining, and you still must first cover half that distance. Repeating this process gives an infinite series of halves, all of which must be traversed to travel from a to b. Since doing an infinite number of things is not humanly possible, you will never be able to reach b. Finally, since this holds for any two points a and b, movement is impossible.

Obviously, Zeno's paradox doesn't hold, since we are able to move from one place to another. But Zeno's paradox has commanded the attention and imagination of philosophers and mathematicians for over 2000 years, as they struggled to deal with the infinity implicit in even the smallest movement. Infinite series is one way (though, some would argue, an incomplete way) of dealing with Zeno's paradox.

2.1 Sequences

Before we jump into series themselves, we need to start with infinite sequences.

2.1.1 Definition

Definition 2.1.1. An *infinite sequence* of real numbers is a set of real numbers indexed by $\mathbb{N}$. These are the common notations for an infinite sequence:

$$\{a_n\}_{n\in\mathbb{N}} \qquad \{a_n\}_{n=0}^{\infty} \qquad \{a_n\} \qquad \{a_1, a_2, a_3, a_4, \ldots\}$$

Example 2.1.2. Sequences can be entirely random, or patterned by some formula or recursion. Here are some familiar examples. I show the first few terms as well as a method of generating higher terms (either direct or recursive).

The sequence of natural numbers:
$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}$ $\quad a_n = n$
The sequence of even numbers:
$\{2, 4, 6, 8, 10, \ldots\}$ $\quad a_n = 2n$
The harmonic sequence:
$\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5} \ldots\}$ $\quad a_n = \frac{1}{n}$
The alternating harmonic sequence:
$\{1, \frac{-1}{2}, \frac{1}{3}, \frac{-1}{4}, \frac{1}{5} \ldots\}$ $\quad a_n = \frac{(-1)^n}{n}$
The geometric sequence with common ratio $\frac{-1}{2}$:
$\{1, \frac{-1}{2}, \frac{1}{4}, \frac{-1}{8}, \frac{1}{16} \ldots\}$ $\quad a_n = \left(\frac{-1}{2}\right)^n$ $\quad a_n = \left(\frac{-1}{2}\right) a_{n-1}$
The arithmetic sequence with common difference 6:
$\{1, 7, 13, 19, 25, \ldots\}$ $\quad a_n = 1 + 6n$ $\quad a_n = a_{n-1} + 6$
The Fibonacci sequence:
$\{1, 1, 2, 3, 5, 8, 13, 21, \ldots\}$ $\quad a_1 = a_2 = 1$ $\quad a_n = a_{n-1} + a_{n-2}$
The sequence of ratios of Fibonacci terms:
$\{1, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \frac{8}{5} \ldots\}$ $\quad a_1 = 1$ $\quad a_n = 1 + \frac{1}{a_{n-1}}$

Definition 2.1.3. There is another definition of sequences which is quite useful. Instead of thinking of $\mathbb{N}$ as an index, we can think of a sequence $\{a_n\}_{n=1}^{\infty}$ as a function:

$$f : \mathbb{N} \to \mathbb{R} \qquad f(n) = a_n$$

If we think of sequences as functions on $\mathbb{N}$, then we can use all of the language of functions. In this way, sequences can be increasing, decreasing, monotonic, bounded above, bounded below and bounded. However, since the domain $\mathbb{N}$ is seperated into discrete numbers, this function f has no continuity properties.

Even though we stated the definition for indices in $\mathbb{N}$, we can choose another starting point: $\{a_n\}_{n=3}^{\infty}$ is a sequence which starts with a_3, and $\{a_n\}_{n=-2}^{\infty}$ is a sequences which starst with a_{-2}. We still always count up from the starting point.

There are many, many sequences studied in mathematics. The Online Encyclopedia of Integer Sequences (OEIS) is a repository for interesting sequences with integer values. As of August 20, 2018, there were 313927 sequences in the OEIS.

2.1.2 Limits of Sequences

As functions $\mathbb{N} \to \mathbb{R}$, sequences are not continuous, so we can't ask for limits at finite values. However, since the index $n \to \infty$, we can ask for the long term behaviour of the sequence.

Definition 2.1.4. The statement

$$\lim_{n\to\infty} a_n = L$$

means that as n gets larger and larger without bound, a_n gets closer and closer to L. Similarly, the statement

$$\lim_{n\to\infty} a_n = \infty$$

means that as n gets larger and larger without bound a_n also gets larger and larger without bound. Sequences with finite limits are *convergent* sequences, and all others (where the limit is either infinite or non-existant) are *divergent* sequences.

As we did for limits of real valued functions, we could restate these limits with ϵ and δ definitions.

Definition 2.1.5. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence. $L \in \mathbb{R}$ is the *limit of the sequence* if

$$\forall \epsilon > 0 \;\; \exists N \in \mathbb{N} \text{ such that } \forall n > N \;\; |a_n - L| < \epsilon$$

We say that the sequence has an infinite limit if

$$\forall M \in \mathbb{N} \;\; \exists N \in \mathbb{N} \text{ such that } \forall n > N \;\; a_n > M$$

To understand limits, we can make great use of the perspective of sequences as functions. We know that limits of functions have many useful properties; all those properties transfer to sequences.

Proposition 2.1.6. *If $\{a_n\}$ and $\{b_n\}$ are convergent sequences, then the following properties hold.*

$$\begin{aligned}
\lim_{n\to\infty} (a_n + b_n) &= \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \\
\lim_{n\to\infty} (a_n - b_n) &= \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n \\
\lim_{n\to\infty} c a_n(x) &= c \lim_{n\to\infty} a_n(x) \\
\lim_{n\to\infty} a_n b_n &= \left(\lim_{n\to\infty} a_n\right)\left(\lim_{n\to\infty} b_n\right) \\
\lim_{n\to\infty} \frac{a_n}{b_n} &= \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n} \\
\lim_{n\to\infty} (a_n(x))^n &= (\lim_{n\to\infty} a_n(x))^n \\
\lim_{n\to\infty} \sqrt[n]{a_n} &= \sqrt[n]{\lim_{n\to\infty} a_n} \qquad \text{if} \qquad a_n \geq 0
\end{aligned}$$

Example 2.1.7. Limits of sequences are limits of functions as the input goes to ∞, so asymptotic analysis applies. We can use asymptotic analysis to easily calculate these examples.

$$\lim_{n\to\infty} n^2 = \infty$$
$$\lim_{n\to\infty} \frac{1}{n} = 0$$
$$\lim_{n\to\infty} \frac{n+1}{n^2} = 0$$

Example 2.1.8. Asymptotic anlysis doesn't solve everything; some limits are still difficult to determine. One such limit is the limit definition of the number e.

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$$

Example 2.1.9. One of our example sequences was the ratio of the Fibonacci terms. How do we calculate this limit of Fibonacci terms?

$$\lim_{n\to\infty} \frac{f_{n+1}}{f_n} = \ ?$$

Let $a_n = \frac{f_n+1}{f_n}$. We can look at the recursive definition of the sequence of Fibonacci terms: $f_{n+1} = f_n + f_{n-1}$. We divide this equation by f_n and manipulate to identify terms of the sequence a_n.

$$\frac{f_{n+1}}{f_n} = \frac{f_n}{f_n} + \frac{f_{n-1}}{f_n} \implies a_n = 1 + \frac{1}{\frac{f_n}{f_{n-1}}} \implies a_n = 1 + \frac{1}{a_{n-1}}$$

We are going to use this expression to find the limit of the sequence. Write ϕ for the value of the limit we wish to calculate (this is a traditional notational choice for this limit). Then we can apply the limit to the above formula and solve for ϕ. (Since all our terms are positive, we only take the positive root in the final step).

$$a_n = 1 + \frac{1}{a_{n-1}}$$
$$\lim_{n\to\infty} a_n = 1 + \frac{1}{\lim_{n\to\infty} a_{n-1}}$$
$$\phi = 1 + \frac{1}{\phi} \implies \phi^2 = \phi + 1 \implies \phi^2 - \phi - 1 = 0$$
$$\phi = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2} = \frac{1+\sqrt{5}}{2}$$

This ϕ is the celebrated Golden Ratio.

2.2 Definition of Infinite Series

Definition 2.2.1. If $\{a_n\}$ is a sequence, then the sum of all infinitely many terms a_n is called an *infinite series*. We write infinite series with sigma notation.

$$\sum_{n=1}^{\infty} a_n$$

The number n is called the *index* and the numbers a_n are called the *terms*. If we want to forget the sum, we can talk about the *sequence of terms* $\{a_n\}_{n=1}^{\infty}$. Though we started with $n = 1$ in this definition, we could start with any integer.

2.2.1 Partial Sums and Convergence

Unlike finite sums, we have no guarantee that this expression evaluates to anything. The problem of infinite series is precisely this: how do we add up infinitely many things? This isn't a problem that algebra can solve, but calculus, with the use of limits, can give a reasonable answer. We need to set up an approximation process and take the limit, just as we did for derivatives and integrals. The approximation process is called partial sums. Instead of taking the entire sum to infinity, let's just take a piece of finite length.

Definition 2.2.2. The nth *partial sum* of an infinite series is the sum of the first n terms.

$$s_n := \sum_{k=1}^{n} a_k$$

Since these are finite sums, we can actually calculate them. They serve as approximations to the total infinite sum. Moreover, these partial sums $\{s_n\}_{n=1}^{\infty}$ define a sequence. We can take the limit of the sequcne of partial sums. This is the limit of the approximation process, so it should calculate the value of the series.

Definition 2.2.3. The value of an infinite series is the limit of the sequence of partial sums, if the limit exists.

$$\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k$$

If this limit exists, we call the series *convergent*. Otherwise, we call the series *divergent*.

Example 2.2.4. The first and most classical example is simply Zeno's paradox. If we are trying to go from 0 to 1, first we travel $\frac{1}{2}$, then $\frac{1}{4}$, then $\frac{1}{8}$, and so on. We represent this paradox as an infinite sum.

$$\sum_{n=1}^{\infty} \frac{1}{2^n}$$

Let's look at the partial sums.

$$
\begin{aligned}
s_1 &= \frac{1}{2} \\
s_2 &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \\
s_3 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} \\
s_4 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16} \\
s_5 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} = \frac{31}{32} \\
s_6 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} = \frac{63}{64} \\
&\vdots \qquad \vdots
\end{aligned}
$$

We can generate a formula to describe the pattern.

$$s_n = \frac{2^n - 1}{2n}$$

Since we have a general expression for the partial sums, we can take the limit.

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{2^n - 1}{2n} = 1$$

Unsurprisingly, we get that the total distance travelled from 0 to 1 is simply 1 unit. This gives a justification for saying that we *can* travel an infinite number of smaller and smaller intervals, since all those infinitely many intervals add up to a finite distance. (Whether this actually soves Zeno's paradox is a question left for the philosophers.)

Example 2.2.5. Now consider the sum of the harmonic series. We are going to analyze the partial sums. We don't get a general formula, but we can define some lower bounds for these partial sums.

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

$$
\begin{aligned}
s_1 &= 1 \\
s_2 &= 1 + \frac{1}{2} = \frac{3}{2} \\
s_3 &= 1 + \frac{1}{2} + \frac{1}{3} \\
s_4 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 2
\end{aligned}
$$

The inequatity holds since $\frac{1}{3} > \frac{1}{4}$ and all other terms remain the same.

$$
\begin{aligned}
s_8 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \\
&> 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{5}{2}
\end{aligned}
$$

We replace all the fraction without powers of 2 in the demoninator with smaller terms to satify the inequality.

$$s_{16} > 3$$

We can generate a lower bound for s_{2^n} in this pattern.

$$\begin{aligned} s_{32} &> \frac{7}{2} \\ s_{64} &> 4 \\ s_{128} &> \frac{9}{2} \\ s_{256} &> 5 \end{aligned}$$

Taking every second power of two gives us partial sums larger than the sequence of positive numbers.

$$s_{2^{2k-2}} > k \qquad \forall k \geq 2$$

The lower bounds get larger and larger. The limit of the sequence of partial sums is larger than this limit of larger bounds.

$$\lim_{n\to\infty} s_n = \lim_{k\to\infty} s_{2^{2k-2}} \geq \lim_{k\to\infty} k = \infty$$

The harmonic series is divergent. This is something of a surprising result, since the harmoinc series looks similar to the series defining Zeno's paradox. However, the terms of the harmonic series are large enough to eventually add up to something larger than any finite number.

This example makes a very important point: it is possible to build a series where the terms are getting smaller and smaller and still end up with an infinite value. This gives some credit to the initial concern of Zeno's paradox, since all these smaller and smaller pieces may eventually add up to something infinite. One might have the intuition that if the terms are becoming very small (as with the harmonic series), the series should have a finite sum; the harmonic series is a counter-example to this intuition. However, the reverse intuition holds, as seen in the following result.

Proposition 2.2.6. *(The Test for Divergence) If we have an infinite series*

$$\sum_{n=1}^{\infty} a_n$$

such that

$$\lim_{n\to a_n} \neq 0$$

then the series must diverge.

Using the test for divergence and the harmonic series example, we can rephrase the important relation between convergence of a series and the limits of the terms. For an infinite series, the fact that the terms tend to zero is *necessary* for convergence but is *not sufficient*. It is a very common temptation to assume the fact that the terms tend to zer is sufficient; be careful not to fall into this trap.

Example 2.2.7. Another important example is an alternating series of positive and negative ones.

$$\sum_{n=1}^{\infty}(-1)^n$$

$$\begin{aligned} s_1 &= 1 \\ s_2 &= 1-1=0 \\ s_3 &= 1-1+1=1 \\ s_4 &= 1-1+1-1=0 \\ s_5 &= 1+1-1-1+1=1 \\ s_6 &= 1+1-1-1+1-1=0 \\ &\vdots \qquad \vdots \end{aligned}$$

We can determine a pattern for even and odd terms.

$$\begin{aligned} s_{2n} &= 0 \quad \forall n \in \mathbb{N} \\ s_{2n+1} &= 1 \quad \forall n \in \mathbb{N} \\ \lim_{n\to\infty} s_n &\qquad DNE \end{aligned}$$

This series does not converge, even though it doesn't grow to infinity. There is simply no way to settle on a value when the partial sums keep switching back and forth from 0 to 1.

Example 2.2.8. There are some nice examples where algebraic manipulation leads to reasonable partial sums. In this example (and similar series), the middle terms in each successive partial sum cancel; these series are called telescoping series.

$$\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$$

$$\begin{aligned} s_n &= \sum_{k=1}^{n}\frac{1}{k(k+1)} = \sum_{k=1}^{n}\frac{1}{k}-\frac{1}{k+1} \\ &= \frac{1}{1}-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\frac{1}{4}\cdots-\frac{1}{n+1} \end{aligned}$$

Almost all the terms conveniently cancel out, leaving only the first and the last.

$$\begin{aligned} &= 1-\frac{1}{n+1} \\ \sum_{n=1}^{\infty}\frac{1}{n(n+1)} &= \lim_{n\to\infty} 1-\frac{1}{n+1} = 1 \end{aligned}$$

Definition 2.2.9. The *factorial* of a natural number n is written $n!$. It is defined to be the product of all natural numbers up to and including n.

$$n! = (1)(2)(3)(4)(5)\ldots(n-2)(n-1)(n)$$

In addition, we define $0! = 1$. (Why? There are good reasons!)

The factorial grows very rapidly. Even by the time we get to $n = 40$, the factorial is already a ridiculously large number.

$$40! = 815915283247897734345611269596115894272000000000$$

Asymptotically, the factorial grows even faster than the exponential.

Example 2.2.10. Here's a series example using the factorial.

$$\sum_{n=0}^{\infty} \frac{1}{n!}$$

$$s_0 = 1$$
$$s_1 = 1 + 1 = 2$$
$$s_2 = 1 + 1 + \frac{1}{2} = \frac{5}{2}$$
$$s_3 = \frac{5}{2} + \frac{1}{6} = \frac{16}{6}$$
$$s_4 = \frac{16}{6} + \frac{1}{24} = \frac{61}{24}$$
$$s_5 = \frac{61}{24} + \frac{1}{120} = \frac{51}{20}$$
$$s_6 = \frac{51}{20} + \frac{1}{720} = \frac{1837}{720}$$

It looks like these terms are growing slowly and possibly leveling off at some value, perhaps less than 3. We can't prove it now, but the value of this series is surprising.

$$\sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n\to\infty} s_n = e$$

This is another definition for the number e. We'll prove that this definition is equivalent our existing definitions in Section 3.4.

Example 2.2.11. The study of values of particular infinite series is a major project in the history of

mathematics. There are many interesting results, some of which are listed here for your curiosity.

$$\pi = 4\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{4}{1} - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} + \frac{4}{9} - \dots$$

$$\pi = \sqrt{12}\sum_{n=-0}^{\infty} \frac{(-1)^n}{3^n(2n+1)} = \sqrt{12}\left(1 - \frac{1}{9} + \frac{1}{45} - \dots\right)$$

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801}\sum_{n=0}^{\infty} \frac{(4n)!(1103+26390n)}{(n!)^4(396)^{4n}}$$

$$\frac{1}{\pi} = \frac{1}{426880\sqrt{16005}}\sum_{n=0}^{\infty} \frac{(6n)!(13591409+545140134n)(-1)^n}{(3n)!(n!)^3(640320)^n}$$

$$\frac{\pi^4}{90} = \sum_{n=1}^{\infty} \frac{1}{n^4}$$

$$e = \sum_{n=0}^{\infty} \frac{(3n)^2+1}{(3n)!}$$

$$e = \sum_{n=0}^{\infty} \frac{n^7}{877n!}$$

2.2.2 Geometric and ζ Series

There are two important classes of convergent series which we will use throughout this chapter. The first is the geometric series.

Definition 2.2.12. For $|r| < 0$, the geometric series with common ratio r is this series.

$$\sum_{n=0}^{\infty} r^n$$

Proposition 2.2.13. *The geometric series with common ratio r converges to $\frac{1}{1-r}$ as long as $|r| < 1$.*

Proof. For $|r| < 1$, we look at the partial sums. We multiply by $\frac{1-r}{1-r}$. In the expansion in the denominator, most of the terms cancel and we are left with a simple expression.

$$s_k = 1 + r + r^2 + r^3 + \ldots + r^k = \frac{1-r}{1-r}\left(1 + r + r^3 + r^3 + \ldots + r^k\right) = \frac{(1-r^k)}{1-r}$$

The convergence of the series is shown by the limit of these partial sums.

$$\sum_{n=0}^{\infty} r^n = \lim_{k\to\infty} \frac{1-r^k}{1-r} = \frac{1}{1-r}$$

□

The second class of convergent series are the ζ (zeta) series. These are often called p-series in standard texts. We are calling them ζ series since this definition, in a broader context, gives the famous Riemann ζ function.

Definition 2.2.14. The ζ series is the infinite series with terms $\frac{1}{n^p}$.

$$\zeta(p) = \sum_{n=1}^{\infty} \frac{1}{n^p}$$

Proposition 2.2.15. *The ζ series converges when $p > 1$.*

We give this without proof for now. The convergence of the ζ series can be proved with the Integral Test in Section 2.6. Unlike the geometric series, where we can easily write the value, the actual value of $\zeta(p)$ is not easy to express in conventional algebraic terms.

2.3 Manipulation of Series

Once we have defined convergent series, we want to be able to work with them algebraically. There are several important manipulations and techniques.

First, series are linear as long as the indices match up. This means we can bring out constants and split series over sums.

$$c\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} ca_n$$
$$\sum_{n=0}^{\infty}(a_n \pm b_n) = \sum_{n=0}^{\infty} a_n \pm \sum_{n=0}^{\infty} b_n$$

Second, we can remove terms. Since a series is just notation for a sum, we can take out leading terms and write them in conventional notation.

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \sum_{n=3}^{\infty} a_n$$

Third, we can shift the indices. The key idea here is balance: whatever we do to the index in the bounds, we do the opposite to the index in the terms to balance it out.

$$\sum_{n=0}^{\infty} a_n = \sum_{n=1}^{\infty} a_{n-1} = \sum_{n=-1}^{\infty} a_{n+1}$$

Both techniques are very useful, particularly for combining series.

Example 2.3.1. In this example, we want to add two series which don't have matching indices. We shift the first series to make the indices match and allow the addition.

$$\sum_{n=2}^{\infty} \frac{3^n}{n!} + \sum_{n=0}^{\infty} \frac{1}{n(n+2)} = \sum_{n=0}^{\infty} \frac{3^{n+2}}{(n+2)!} + \sum_{n=0}^{\infty} \frac{1}{n(n+2)} = \sum_{n=0}^{\infty} \left(\frac{3^{n+2}}{(n+2)!} + \frac{1}{n(n+2} \right)$$
$$= \sum_{n=0}^{\infty} \left(\frac{3^{n+2} + (n-1)!(n+1)}{(n+2)!} \right)$$

2.4 Absolute and Conditional Convergence

Among convergent series, there is a distincting between a stronger and a weaker kind of convergence. This section explores that distinction via the important example of the alternating harmonic series.

2.4.1 Alternating Series Test

Definition 2.4.1. If we have a sequences of terms $\{a_n\}$ such that $a_n > 0 \ \forall n \in \mathbb{N}$, then the following expression is called an alternating series.

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n$$

In an alternating series, each term has a different sign from the previous term. Recall the test for divergence: for convergence, is it *necessary* but not *sufficient* for the terms to tend to zero. Intuitively, we would like to have sufficiency as well, but the harmonic series was the counter example. For alternating series, we get our wish.

Proposition 2.4.2. *(The Alternating Series Test) An alternating series converges if and only if the limit of the terms is zero.*

2.4.2 The Alternating Harmonic Series

Definition 2.4.3. The alternating harmonic series is the harmonic series with $(-1)^n$ in the numerator.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

This series converges by the alternating series test. It is difficult to prove, but the value is $\ln 2$.

We must be very careful here. Consider the following series, which is a pattern of two positive terms with odd denominators followed by one negative term with an even denominator.

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \dots$$

We write this series in a strange new way: as the difference of the alternating harmonic series and another series with only even denominators. You can see, if you add the two series, how some of the terms cancel and other add to the correct terms in the original.

$$\begin{aligned} &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots \\ &\quad + 0 + \frac{1}{2} + 0 - \frac{1}{4} + 0 + \frac{1}{6} + 0 - \frac{1}{8} + \dots \end{aligned}$$

The first series is the harmonic series. If we factor $\frac{1}{2}$ out of the second series, it is also the harmonic series. We use the known value of $\ln 2$ for the harmonic series to calculate the value of this series.

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} + \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2 + \frac{1}{2} \ln 2 = \frac{3}{2} \ln 2$$

This looks reasonable as well, but what are the terms of this series? If we group them by sign, the positive terms are $\{1, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \dots\}$ and the negative terms are $\{\frac{-1}{2}, \frac{-1}{4}, \frac{-1}{6}, \dots\}$. These are exactly the same terms at the alternating harmonic series, just in a different order. However, the alternating harmonic series summed to $\ln 2$, not $\frac{3}{2} \ln 2$.

It seems we can re-arrange the alternating harmonic series to sum to a different number. This is exceedingly odd: for finite sums, any re-arrangement was irrelevant to the value of the sum. It seems, for infinite sums, re-arrangement can actually change the value. There is an important result which is even stranger.

Proposition 2.4.4. *For any real number α, there is a re-arrangement of the alternating harmonic series that sums to α.*

Proof. This is a very strange result, but the proof has a remarkably simple argument.First, groups the terms as positive and negative. Each set of terms is asymptotically similar to the (non-alternating) harmonic series, so each set sums to $\pm\infty$.

Then choose a real number α. Start adding positive terms until we get past α. (This can always be done, since the possitive terms by themslves sum to ∞). Then, when we are past α, start adding negative terms until we're below α again. (Again, this can always be done, since the negative terms sum to $-\infty$). Then simply repeat this process, adding positives until we get above α and negatives until we get back below α. This process can be continued indefinitely, and since the terms get arbitrarily small, we will approach α in the limit. □

Example 2.4.5. There are some regular arrangements of the alternating harmonic which have specific values. Let $A(m,n)$ be the sum where we take m positive terns, then n negative, then back to m positive and so on. It can be proved that this converges to

$$A(m,n) = \ln 2 + \frac{1}{2}\ln\left(\frac{m}{n}\right).$$

In particular, the combination of one positive and four negative terms sums to zero.

$$A(1,4) = \ln 2 + \frac{1}{2}\ln\frac{1}{4} = \ln 2 + \ln\left(\frac{1}{4}\right)^{\frac{1}{2}} = \ln 2 + \ln\frac{1}{2} = \ln 2 - \ln 2 = 0$$

$$\begin{aligned} 0 = 1 &- \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \frac{1}{8} \\ &+ \frac{1}{3} - \frac{1}{10} - \frac{1}{12} - \frac{1}{14} - \frac{1}{16} \\ &+ \frac{1}{5} - \frac{1}{18} - \frac{1}{20} - \frac{1}{22} - \frac{1}{24} \\ &+ \frac{1}{7} - \frac{1}{26} - \frac{1}{28} - \frac{1}{30} - \frac{1}{32} \\ &+ \frac{1}{9} - \frac{1}{34} - \frac{1}{36} - \frac{1}{38} - \frac{1}{40} \cdots \end{aligned}$$

2.4.3 Conditional Convergence

This situation for the alternating harmonic series is not unique.

Definition 2.4.6. A convergent series $\sum a_n$ is called *absolutely convergent* if

$$\sum_{n=1}^{\infty} |a_n| < \infty.$$

Otherwise, if a series is convergent but not absolutely convergent, it is called *conditionally convergent.*

The alternating harmonic series was a conditionally convergent series, since the (non-alternating) harmonic series diverges. The behaviour that we saw for the alternating harmonic series is the same for *any* conditionally convergent series.

Proposition 2.4.7. *An absolutely convergent series converges to the same value regardless of reordering, but a conditionally convergent series can be rearranged to converge to any real number.*

2.5 Decimal Expansions for Real Numbers

A nice application of infinite series is a proper and complete account of decimal expansions for real numbers. Such expansions have infinite length, so some of the standard problems involving infinite process are involved; limits are required.

The starting question: what are decimal expansions and why do infinite strings of decimals actually represent numbers? Let's write a real number α as $a.d_1d_2d_3d_4\ldots$ where $a \in \mathbb{N}$ and the d_i are digits in $\{0, 1, \ldots, 9\}$. The meaning of this decimal expansion is an infinite series.

$$a + \frac{d_1}{10} + \frac{d_2}{100} + \frac{d_3}{1000} + \frac{d_4}{10000} + \frac{d_5}{100000} + \ldots \implies \alpha = a + \sum_{n=1}^{\infty} \frac{d_n}{10^n}$$

Asymptotically, since the numerators are bounded by 9, this series behaves like $\frac{1}{10^n}$. The terms $\frac{1}{10^n}$ are terms of a geometric series with common ratio $\frac{1}{10}$, which is convergent, so the decimal expansion is also convergent. Therefore, we can be confident that decimal expansions, though infinite, are a reasonable representation of real numbers.

Now think about the partial sums of a decimal expansion.

$$s_n = a + \sum_{k=1}^{\infty} \frac{d_k}{10^k}$$

These are all finite sums of fractions, hence they are rational numbers. Any real number α is the limit of sums of this type, since its decimal expansion is the limit of the partial series sums. This establishes an important and useful fact about real numbers (sometimes even taken as the definition of real numbers)

Proposition 2.5.1. *All real numbers are limits of sequences of rational numbers.*

2.6 Convergence Tests

When we are using a series, the most pressing problem is convergence. We want to work with actual values, but we are only guaranteed those values when the series converge. Therefore, mathematicians have devised many ways to test a series for convergence. We already have the test for divergence. In this section, we will present several more ways to test a series for convergence.

2.7 Comparison on Series

Proposition 2.7.1. *(Direct Comparison) Let $\{a_n\}$ and $\{b_n\}$ be the terms of two infinite series. Then an inequality of the terms implies an inequality of the series.*

$$a_n \leq b_n \quad \forall n \in \mathbb{N} \implies \sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n$$

In addition, let a_n and b_n be positive for all $n \in \mathbb{N}$.

- *If $\sum b_n$ is convergent, since the sum $\sum a_n$ is smaller, it must also be convergent.*
- *If $\sum a_n$ is divergent, since the sum $\sum b_n$ is larger, so it must also be divergent.*

Example 2.7.2. Here are some comparison examples.

$$\sum_{n=3}^{\infty} \frac{1}{n-2}$$

The terms $\frac{1}{n-2}$ are larger that $\frac{1}{n}$ and the harmonic series $\sum \frac{1}{n}$ is divergent, so this series is also divergent.

$$\sum_{n=1}^{\infty} \frac{1}{3^n + 4n + 1}$$

The terms $\frac{1}{3^n+4n+1}$ are smaller that $\frac{1}{3^n}$. These term $\frac{1}{3^n}$ are the terms of a geometric series with common ratio $\frac{1}{3}$, which converges. Therefore, this series converges.

$$\sum_{n=1}^{\infty} \frac{n+1}{n^2}$$

The terms $\frac{n+1}{n^2}$ are larger than $\frac{1}{n}$. The latter are the terms of the divergent harmonic series, so this series diverges.

$$\sum_{n=1}^{\infty} \frac{2^n}{n!}$$

For $n \geq 4$ we have

$$\frac{2^n}{n!} = \frac{2}{1}\frac{2}{2}\frac{2}{3}\cdots \leq \frac{2}{3}\left(\frac{1}{2}\right)^{n-4}$$

Therefore, we can compare to a geometric series with common ratio $\frac{1}{2}$, which converges. Therefore, this series also converges (and converges to e^2, as it happens).

As the last paranthetical comment hinted, comparison doesn't actually give the value of the series. These comparison arguments are very useful for determining convergence and divergence, but they don't calculate exact values. Also, in the last example the comparison only held for $n \geq 4$ instead of all $n \in \mathbb{N}$. This is typical and perfectly acceptable; for everything involving series other than calculating the exact value, we only need to consider the long term behaviour. For comparison, it is enough that $a_n < b_n$ for all n past some finite fixed value.

In addition to the exact comparisons listed above, we can also compare asymptotically. Asymptotic comparison is particularly useful, since we don't actually have to calculate the inequalities.

Proposition 2.7.3. *(Asymptotic Comparison) Let $a_n, b_n \geq 0$ be the terms of two series. If a_n and b_n have the same asymptotic order (in the variable n), then the two series*

$$\sum_{n=1}^{\infty} a_n \qquad \text{and} \qquad \sum_{n=1}^{\infty} b_n$$

have the same convergence behaviour: either they both converge or they both diverge.

In Example 2.7.2, we could have simply said that $\frac{1}{n-2}$ is asymptotically the same order as $\frac{1}{n}$, and likewsie for $\frac{n+1}{n^2}$. Asymptotic comparison is often easier since we don't need to explicitly construct the necessary inequality.

Example 2.7.4. As an example for both asymptotic comparison and conditional convergence, here are three alternating series. They are all convergent by the alternating series test. Comparison to geometric series or a ζ series is used to check their absolute convergence.

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^6} \qquad \text{absolutely convergent by asymptotic comparison to } \frac{1}{n^6}$$

$$\sum_{n=1}^{\infty} \frac{(-1)^n \arctan n}{n^2} \qquad \text{absolutely convergent by asymptotic comparison to } \frac{1}{n^2}$$

$$\sum_{n=2}^{\infty} \frac{(-1)^n}{\ln n} \qquad \text{conditionally convergent by asymptotic comparison to} \frac{1}{\ln n}$$

In the last compoarision, $\frac{1}{\ln n} > \frac{1}{n}$, so the asymptotic order of $\frac{(-1)^n}{\ln n}$ is that of a divergent series, growing faster than the harmonic series.

2.7.1 The Integral Test

Proposition 2.7.5. *(Integral Test) If a series has positive terms and $a_n = f(n)$ for f an integrable function, then the series is convergent if and only if the following improper integral is convergent:*

$$\int_1^{\infty} f(x)dx$$

Note that the integral and the resulting series will sum to different numbers: this test doesn't calculate the value of the sum. It just tells us whether the sum is convergent.

Example 2.7.6. As promised, the integral tests allows us to prove the that ζ series converges if and only if $p > 1$.

$$\int_1^{\infty} \frac{1}{x} dx = \infty \implies \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

$$\int_1^{\infty} \frac{1}{x^p} dx < \infty \implies \sum_{n=1}^{\infty} \frac{1}{n^p} < \infty \text{ for } p > 1$$

2.7.2 The Ratio and Root Tests

Here are two final tests.

Proposition 2.7.7. *(Ratio Test) If a_n are the terms of a series, consider the limit of the ratio of the terms.*

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

If this limit is infinite or finite and > 1, then the series diverges. If this limit is < 1, then the series converges. If the limit is 1, the test is inconclusive.

Proposition 2.7.8. *(Root Test) If a_n are the terms of a series, consider the limit of the roots of the terms.*

$$\lim_{n\to\infty} \sqrt[n]{|a_n|}$$

If this limit is infinite or finite and > 1, then the series diverges. If this limit is < 1, then the series converges. If the limit is 1, the test is inconclusive.

The ratio test is useful for powers and particularly for factorials. The root test is obviously useful for powers.

2.7.3 Testing Strategies

There are many approaches to testing the convergence of a series: looking at partial sums, testing for divergence, comparison, asymptotic comparison, the alternating series test, the integral test, the ratio test, and the root test. It is difficult to know where to start and which tests or techniques to use. Here are some pointers and strategies to help you.

- Looking at a series for asymptotic order is often the easiest first step. The main comparisons are with geometric series and ζ-series.
- Using the test for divergence is also often an easy first step. If the terms do not tend to 0, the series cannot converge.
- If the series is an alternating series, the alternating series test is likely the easiest approach.
- The integral test is often the best approach if the series involves complicated functions, such as exponentials, logarithms or trigonometric functions.
- The ratio test is often the best approach when the terms involve factorials. It is also very useful for terms which have the index in the exponent.
- The root test is rarely used. It also helps when the index is in the exponent, but most of those cases can also be done with the ratio test.

A final important observation is that convergence only cares about the long-term behaviour of the series. Any finite pieces at the start are negligible. This is a nice observation for many of the tests: comparisons only need to work eventually, integrals can be taken on $[a, \infty)$ for some $a > 0$, and a series which eventually becomes an alternating series can use the alternating series test.

Example 2.7.9. For an extreme example, consider this series:

$$\sum_{n=1}^{10^{300}} (n^2+n)^{75} + \sum_{n=10^{300}+1}^{\infty} \frac{1}{n^2}$$

The first 10^{300} terms of this series are enormous numbers and their sum is simply ridiculous. However, the series is eventually is a ζ-series with $p = 2$, which converges. Therefore, this sum is finite. The ridiculous number we get from the first 10^{300} terms is very, very large, but certainly finite. Any very, very large number is negligible when asking about infinity.

2.7.4 Testing Examples

Now that we have all the tools at our disposal, here are a bunch of examples.

Example 2.7.10.

$$\sum_{n=1}^{\infty} n^{-\frac{2}{3}}$$

The terms are $\frac{1}{n^{\frac{2}{3}}}$, so this is a ζ series. Since $\frac{2}{3} < 1$, this diverges.

Example 2.7.11.

$$\sum_{k=1}^{\infty} \frac{2^k}{e^k}$$

The terms are $\left(\frac{2}{e}\right)^k$, so this is a geometric series. $\frac{2}{e} < 1$ so converges.

Example 2.7.12.

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2-1}$$

This is an alternating series with decreasing terms, so it converges.

Example 2.7.13.

$$\sum_{n=1}^{\infty} \sin\left(\frac{n^2+1}{n}\right)$$

The terms do not tend to zero, so the series is divergent by the test for divergence.

Example 2.7.14.

$$\sum_{k=1}^{\infty} \frac{1}{k^2-1}$$

This is asymptotically $\frac{1}{k^2}$, which is a convergent ζ series.

Example 2.7.15.

$$\sum_{k=1}^{\infty} \frac{(-1)^k k}{e^k}$$

This is an alternating series with decreasing terms, so it converges.

Example 2.7.16.

$$\sum_{n=1}^{\infty} \frac{3}{2+e^n}$$

This is asymptotically $\frac{1}{e^n}$, which is a convergent geometric series.

Example 2.7.17.

$$\sum_{k=1}^{\infty} \frac{k\sqrt{k}}{k^3}$$

This is asymptitically $\frac{1}{k^{\frac{3}{2}}}$, which is a convergent ζ series.

Example 2.7.18.

$$\sum_{n=1}^{\infty} \frac{(-1)^n n}{n^3+4}$$

This is an alternating series. The terms are decreasing, so the alternating series test gives convergence. In addition, the absolute value of the terms is $\frac{n}{n^3+4}$ which is asymptotically $\frac{1}{n^2}$. That converges, so the series is absolutely convergent and can be rearranged without changing the value.

Example 2.7.19.

$$\sum_{k=1}^{\infty} \frac{2^k k!}{k^k}$$

The factorial suggests that the ratio test is the best approach.

$$\begin{aligned}\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| &= \lim_{k\to\infty}\frac{2^{k+1}(k+1)!k^k}{2^k k!(k+1)^{k+1}} = \lim_{k\to\infty} 2(k+1)\left(\frac{k}{k+1}\right)^k \frac{1}{k+1} \\ &= \lim_{k\to\infty} 2\left(\frac{k}{k+1}\right)^k = \frac{2}{e} < 1\end{aligned}$$

By the ratio test, this is convergent.

Example 2.7.20.

$$\sum_{k=1}^{\infty} k^5 e^{-k}$$

We use the ratio test:

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\frac{(k+1)^5 e^k}{k^5 e^{k+1}} = \lim_{k\to\infty}\frac{1}{e}\left(\frac{k+1}{k}\right)^5 = \lim_{k\to\infty}\frac{1}{e}\left(1+\frac{1}{k}\right)^5 = \frac{1}{e} < 1$$

By the ratio test, this is convergent.

Example 2.7.21.

$$\sum_{n=1}^{\infty}\frac{\ln n^2}{n^2}$$

We use the integral test, with $u = lnx$.

$$\int_1^{\infty}\frac{\ln x^2}{x^2}dx = 2\int_1^{\infty}\frac{\ln x}{x^2}dx = 2\int_0^{\infty} ue^{-u}du = -ue^{-u}\big|_1^{\infty} + 2\int_0^{\infty} e^{-u} = 0 - e^{-u}\big|_1^{\infty} = 1 \leq \infty$$

By the integral test, this is convergent.

Example 2.7.22.

$$\sum_{n=1}^{\infty}\frac{(n!)^2}{(2n+4)!}$$

The presence of a factorial means that ratio test is probably the best.

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{\frac{((n+1)!)^2}{(2(n+1)+4)!}}{\frac{(n!)^2}{(2(n)+4)!}} = \lim_{n\to\infty}\frac{(n+1)^2}{(2n+5)(2n+6} = \lim_{n\to\infty}\frac{n^2+2n+1}{4n^2+22n+30} = \frac{1}{4} < 1$$

The limit is less than 1, so the series is convergent.

Example 2.7.23.

$$\sum_{n=1}^{\infty}\frac{n!}{e^{n^2}}$$

We have factorials again, so ratio test is likely the best choice.

$$\begin{aligned}\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\frac{\frac{(n+1)!}{e^{(n+1)^2}}}{\frac{n!}{e^{n^2}}} = \lim_{n\to\infty}\frac{(n+1)e^{n^2}}{e^{n^2+2n+1}} \\ &= \lim_{n\to\infty}\frac{(n+1)e^{n^2}}{e^{n^2}e^{2n}e} = \lim_{n\to\infty}\frac{n+1}{e^{2n+1}} = 0\end{aligned}$$

The limit is less than 1, so the series is convergent.

Example 2.7.24.

$$\sum_{n=1}^{\infty} \frac{\ln n}{\sqrt[3]{n}}$$

The integral test if most appropriate here, even though the interal is difficult.

$$\begin{aligned}
\int_1^{\infty} \frac{\ln x}{\sqrt[3]{x}} &= \int_1^{\infty} \ln x x^{\frac{-1}{3}}\, dx \\
f = \ln x \implies f' &= \frac{1}{x} \\
g' = x^{\frac{-1}{3}} \implies g &= \frac{3x^{\frac{2}{3}}}{2} \\
&= \frac{3x^{\frac{2}{3}}}{2} \ln x \Big|_1^{\infty} - \int_1^{\infty} \frac{3x^{\frac{2}{3}}}{2} \frac{1}{x} dx = \lim_{a\to\infty} \left[\frac{3a^{\frac{2}{3}} \ln a}{2} - \frac{3}{2} 1^{\frac{3}{2}} \ln 1 - \frac{3}{2} \int_1^a x^{-\frac{1}{3}} dx \right] \\
&= \lim_{a\to\infty} \frac{3}{2} \left[a^{\frac{2}{3}} \ln a - \frac{3x^{\frac{2}{3}}}{2} \Big|_1^a \right] = \lim_{a\to\infty} \frac{3}{2} \left[a^{\frac{2}{3}} \ln a - \frac{3a^{\frac{2}{3}}}{2} + \frac{3}{2} \right] \\
&= \lim_{a\to\infty} \frac{3}{2} \left[a^{\frac{2}{3}} \left(\ln a - \frac{3}{2} \right) + \frac{3}{2} \right] = \infty
\end{aligned}$$

The integral diverges, so the series must as well.

Example 2.7.25.

$$\sum_{n=2}^{\infty} \frac{1}{n \ln n}$$

We use the integral text again. In the integral, we use the substitution $u = \ln x$.

$$\int_2^{\infty} \frac{1}{x \ln x} dx = \int_{\ln 2}^{\infty} \frac{1}{u} du = \ln u \Big|_{\ln 2}^{\infty} = \infty$$

The integral is divergent, so the sum is divergent as well. Also, note the following inequality.

$$\frac{1}{n} > \frac{1}{n \ln n} > \frac{1}{n^p}$$

It seems that comparison should be helpful with a series of this type. However, the inequality shows that this series is asymptotically between the harmonic series and the other convergent p series. In comparison, it's slightly larger than a convergent series and slightly smaller than a divergent series, which is entirely unhelpful.

Example 2.7.26.

$$\sum_{n=1}^{\infty} n e^{-n^2}$$

We use the integral test again. In the integral, we use the substitution $u = x^2$.

$$\int_1^\infty xe^{-x^2}dx = \frac{1}{2}\int_1^\infty e^{-u}du = \left.\frac{-1}{2}e^{-u}\right|_1^\infty = \frac{e}{2} < \infty$$

The integral converges, so the sum does as well. Note that the sum *does not* have the value $\frac{e}{2}$.

Example 2.7.27.

$$\sum_{n=1}^{\infty}(-1)^n 3^{\frac{1}{n}}$$

The root test is good for exponents.

$$\lim_{n\to\infty}\sqrt[n]{3^{\frac{1}{n}}} = \lim_{n\to\infty} 3^{\frac{1}{2n}} = 1$$

This limit is 1, so the test is inconclusive. Instead of using the root test, look at the limit of the terms. That limit is ± 1, which is not zero, so the series must diverge by the test for divergence.

Example 2.7.28.

$$\sum_{n=1}^{\infty}\left(\frac{n}{n+1}\right)^{n^2}$$

We use the root test again.

$$\begin{aligned}\lim_{n\to\infty}\sqrt[n]{|a_n|} = \lim_{n\to\infty}\sqrt[n]{\left(\frac{n}{n+1}\right)^{n^2}} &= \lim_{n\to\infty}\left(\frac{n}{n+1}\right)^n = \frac{1}{\lim_{n\to\infty}\left(\frac{n+1}{n}\right)^n} \\ &= \frac{1}{\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n} = \frac{1}{e} < 1\end{aligned}$$

The limit is less than 1, so the series converges.

Example 2.7.29.

$$\sum_{n=1}^{\infty}\tan\left(\frac{1}{n}\right)$$

There is an interesting comparison argument which we can use to tackle this difficult example. The derivative of tangent is $\sec^2 x$. Since $\sec x$ is always > 1 or < -1, we have $\sec^2 x > 1$. That is, the slope of tangent is always larger than 1. Since $\tan 0 = 0$, that means that near the origin, $\tan x > x$. Equivalently, for large n, $\tan\frac{1}{n} > \frac{1}{n}$. This allows us to compare our series to the harmonic series: our terms are larger than the harmonic series and the harmonic series diverges, so this series also diverges.

2.8 Values of Infinite Series

In all the examples of the previous section, we have been testing series convergence. One might complain that we are ignoring the more practical question of actually calculating the values. However, the problem of finding closed form expressions for the values of series is difficult. Consider the following series.

$$\sum_{n=1}^{\infty} \frac{n^2 + 4n}{(n!)2^n}$$

This is easy to check for convergence: therefore, this does converge to a number. Can we say what number? Most likely the number is some number we simply don't recognize, some irrational number which has no name. This is an inherent issue with series: often the series itself is the best representation of the new number, whatever it is.

However, we would often like to know the rough value of the number. We can use the series to approximate the number. In the previous example, we can approximate the sum by its fourth partial sum.

$$\frac{5}{2} + \frac{12}{8} + \frac{21}{48} + \frac{32}{384} = \frac{217}{48}$$

If we add more terms, we get a closer approximation. Other than this series of improving approximations, we really don't know what it is. The accuracy and precision of our approximation is an important question in the study of infinite series; we will return to it at the end of the next chapter.

Chapter 3

Taylor Series

3.1 Series as Functions

Our two main examples for comparison in the previous chapter were the geometric series and the ζ series. Both converged for certain values of the parameter; $|r| < 1$ for the geometric series and $p >$ for the ζ series. To start this section, I'd like to re-interpret these two series. Instead of thinking of a whole family of different series which depend on a parameter (r or p), we can think of each family of series as a *function* of the parameter. In this view, we have only one series, but the series produces a function instead of just a number.

For the geometric series, this new perspective defines a function $f(x)$.

$$f(x) = \sum_{n=0}^{\infty} x^n$$

The domain of this function is $|x| < 1$, since those are the values of the parameter (now the variable) where the geometric series converges. We even know what this function is:

$$f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad\qquad |x| < 1$$

In this way, the geometric series now defines the function $f(x) = \frac{1}{1-x}$. The domain restriction of the function is determined by the convergence of the series: a point x is in the domain of the function if the series converges for that choice of x.

We can do the same with the ζ series. The reason we called them ζ series is that the associated function is called the Riemann ζ-function.

$$\zeta(x) = \sum_{n=0}^{\infty} \frac{1}{n^x}$$

The domain of this function is $(1, \infty)$, since that is where the series converges. (In other branches of mathematics, the domain of ζ is extended in new and interesting ways. The vanishing of the ζ function is the subject of the famous Riemann Hypothesis, an important unsolved problem in modern mathematics.)

In general, an infinite series can represent a function $f(x)$ when the terms a_n of the series also depend on x.

$$f(x) = \sum_{n=1}^{\infty} a_n(x)$$

Notice that the variable of the function, x, is not the index of the sum n. These two numbers are different and must not be confused. The domain of this function is the set of values of x for which the series converges. Instead of previous domain restrictions, involving division by zero, square roots and other problems, domain restrictions are now all about convergence. For these series, convergence is no longer a yes/no question. Instead, it is a domain question: for which real numbers does the series converge?

3.2 Power Series

3.2.1 Definition

A polynomial $p(x)$ of degree d can be written as a finite sum in sigma notation.

$$p(x) = \sum_{n=0}^{d} c_n x^n$$

The terms involve powers of the variables (x^n) and coefficients of those powers (c_n). What if we let the degree become arbitrarily large, going to infinity?

Definition 3.2.1. A series of the form

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

is called a *power series*. The real numbers c_n are called the *coefficients* of the power series. The whole expression $c_n x^n$ is still the *term* of the power series.

The full definition is slightly more general. The previous series was a power series *centered at 0*. We can centre a power series at any $\alpha \in \mathbb{R}$.

Definition 3.2.2. A series of the form

$$f(x) = \sum_{n=0}^{\infty} c_n (x - \alpha)^n$$

is called a *power series centered at* α. The numbers c_n are still called the *coefficients* and the number α is called the *centre point*. The whole expression $c_n(x - \alpha)^n$ is still the *term*.

3.2.2 Radii of Convergence

Polynomials were defined for all real numbers; they had no domain restrictions. However, series do have domain restrictions. A power series may or may not converge for all real values of x. The first and most important issue when we start using series as functions is determining the domain of convergence. For power series, we will almost always use the ratio test. Recall that the ratio tests shows convergence when the limit of the ratio of the terms is < 1. We will use some examples to show the various types of behaviours.

Example 3.2.3.

$$\sum_{n=0}^{\infty} \frac{(x+2)^n}{n^2}$$

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \left|\frac{\frac{(x+2)^{n+1}}{(n+1)^2}}{\frac{(x+2)^n}{n^2}}\right| = \lim_{n\to\infty} \left|\frac{(x+2)n^2}{(n+1)^2}\right|$$

$$= |x+2| \lim_{n\to\infty} \frac{n^2}{n^2+2n+1} = |x+2| < 1$$

This series is centered at $\alpha = -2$, and the ratio test tells us that we have convergence on $|x+2| < 1$, which is the interval $(-3,-1)$. Outside the interval, the series diverges and doesn't represent a function. The convergence at the endpoints -3 and -1 is undetermined; we would need to check them individually using another type of test.

Example 3.2.4.

$$\sum_{n=0}^{\infty} nx^n$$

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \left|\frac{x^{n+1}(n+1)}{x^n n}\right| = |x| \lim_{n\to\infty} \frac{n+1}{n} = |x| < 1$$

The ratio test allows us to conclude that this converges on $(-1,1)$.

Example 3.2.5.

$$\sum_{n=0}^{\infty} n!x^n$$

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \left|\frac{x^{n+1}(n+1)!}{x^n n!}\right| = |x| \lim_{n\to\infty} \frac{n+1!}{n!} = \infty$$

This limit is never finite unless $x = 0$, so this converges almost nowhere. This is essentially useless as the definition of a function, since its only value is $f(0) = 0$.

Example 3.2.6.

$$\sum_{n=0}^{\infty} \frac{(-1)^n (x-7)^n}{2^n n!}$$

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \left| \frac{\frac{(x-7)^{n+1}}{2^{n+1}(n+1)!}}{\frac{(x-7)^n}{2^n n!}} \right| = |x-7| \lim_{n\to\infty} \frac{1}{2(n+1)} = 0 < 1$$

The limit here is 0 regardless of the value of x, so we have convergence for all real numbers.

The previous examples represent all of the possible types of convergence behaviour of power series. We summarize the situation in a proposition.

Proposition 3.2.7. *Consider a power series centered at $x = \alpha$.*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-\alpha)^n$$

Such a series will have precisely one of three convergence behaviours.

- *It will only converge for $x = \alpha$, where it has the value c_0.*
- *If will converge for all of $\mathbb{R}$.*
- *There will be a real number $R > 0$ such that the series converges on $(\alpha - R, \alpha + R)$. It will diverge outside this interval, and the behaviour at the end points is undetermined and has to be checked individually.*

Definition 3.2.8. The positive real number R in the third case is called the *radius of convergence* of a power series. We can use this terminology to cover the other two cases as well: in the first case, we say $R = 0$ and in the second case, we say $R = \infty$.

Example 3.2.9.

$$f(x) = \sum_{n=1}^{\infty} \sqrt{n} x^n \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left| \frac{\sqrt{n+1} x^{n+1}}{\sqrt{n} x^n} \right| = \lim_{n\to\infty} |x| \sqrt{\frac{n+1}{n}}$$

$$= |x| \lim_{n\to\infty} \sqrt{1 + \frac{1}{n}} = |x| < 1$$

The radius of convergence is $R = 1$, so this series converges on $(-1, 1)$.

Example 3.2.10.

$$f(x) = \sum_{n=1}^{\infty} \frac{n(x-6)^n}{4^{2n+2}} \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left| \frac{\frac{(n+1)(x-6)^{n+1}}{4^{2n+3}}}{\frac{n(x-6)^n}{4^{2n+2}}} \right|$$

$$= |x-6| \lim_{n\to\infty} \frac{n+1}{n} \frac{1}{4^2} = \frac{|x-6|}{16} < 1 \implies |x-6| < 16$$

The radius of convergence is $R = 16$, centered around $x = 6$. This series converges on $(-10, 22)$.

Example 3.2.11.

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{7^n} \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left| \frac{\frac{x^{n+1}}{7^{n+1}}}{\frac{x^n}{7^n}} \right|$$
$$= |x| \lim_{n\to\infty} \left| \frac{1}{7} \right| = \frac{|x|}{7} < 1 \implies |x| < 7$$

The radius of convergence is 7 and the series converges on $(-7, 7)$.

Example 3.2.12.

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{(1)(3)(5)\dots(2n+1)} \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left| \frac{\frac{x^{n+1}}{(1)(3)(5)\dots(2n+3)}}{\frac{x^n}{(1)(3)(5)\dots(2n+1)}} \right|$$
$$= |x| \lim_{n\to\infty} \frac{1}{2n+3} = 0$$

This convergence doesn't depend on x, since the limit is 0 in any case. Therefore, this has an infinite radius of convergence and is a function defined on all of $\mathbb{R}$.

Sometimes, we like to simply calculate the radius directly. Here are two formulae to do so.

Proposition 3.2.13. *If we have a power series where all the coefficients c_n are non-zero, then we can calculate the radius of convergence directly in either of two ways.*

$$R = \lim_{n\to\infty} \left| \frac{c_n}{c_{n+1}} \right|$$
$$R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}}$$

3.2.3 Properties of Power Series

Inside the radius of convergence, a power series has all the properties of a normal function. We can add and subtract two power series as long as we remain inside the radii of both series. We can multiply as well, though the calculations become difficult. The same is true for division: if a series is non-zero inside its radius of convergence, we can divide by the series (though the results of the calculation are difficult to use).

Other properties of series can be calculated with various ease or difficulty, depending on the series. We can investigate the growth of series, whether or not they are bound, symmetric or periodic, and whether or not they are invertible. The key idea to remember is that power series, inside their radii of convergence, are functions; anything that applies to functions can be applied to power series.

3.2.4 Calculus of Power Series

Since power series are functions, we can try to do calculus with them, investigating their limits, continuity, derivatives and integrals.

Proposition 3.2.14. *Assume we have a power series centered at α:*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-\alpha)^n$$

This f is a continuous function inside its radius of convergence. In addition, f is infinitely differentiable inside its radius of convergence.

There is a convenient notation for differentiability which we will use frequently.

Definition 3.2.15. If f is a function on a domain D and the n-th derivative of f is defined and continuous, we say that f is in class $C^n(D)$. If the domain is understood implicitly, we just say f is in class C^n. If f is infinitely differentiable, we say f is in class C^∞.

The proposition says that power series are in class C^∞, but how are these derivatives calculated? The answer is as nice as possible.

Proposition 3.2.16. *If f is a power series, then derivative of f is calculated term-wise, simply by differentiating every term in the series.*

$$f'(x) = \sum_{n=1}^{\infty} c_n n (x-\alpha)^{n-1}$$

Therefore, the derivative is a power series as well; moreover, it will have the same radius of convergence as the original.

Integration is just as pleasant for power series.

Proposition 3.2.17. *If f is a power series centered at α, then f is integrable and its indefinite integral is calculated termwise.*

$$\int f(x)dx = \sum_{n=0}^{\infty} c_n \frac{(x-\alpha)^{n+1}}{n+1} + C$$

The simplicity of integration is particularly helpful. As we saw in Calculus II, integration is difficult business. For functions which be expressed as series, integration is almost trivial. This makes power series a very useful and convenient class of functions.

3.3 Taylor Series

3.3.1 Analytic Functions

Once again, consider the geometric series:

$$f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

Unlike most of the power series we've seen so far, we actually know the values of the geometric series. This series, as a function, is the same as the function $\frac{1}{1-x}$ on the domain $(-1, 1)$. (The function $\frac{1}{1-x}$ is certainly defined on a larger domain, but the series is not). We can say that the geometric series lets us write $\frac{1}{1-x}$ as an infinite series; it is the infinite series representation of the function on the domain $(-1, 1)$.

The theory of Taylor series generalizes this situation. For various functions $f(x)$, we want to a build representation of $f(x)$ as a series. This will be a power series which is identical to $f(x)$, at least for part of its domain. To find the power series, we need to choose a centre point α and find coefficients c_n such that

$$f(x) = \sum_{n=0}^{\infty} c_n(x-\alpha)^n.$$

Definition 3.3.1. A function is called *analytic* at $\alpha \in \mathbb{R}$ if it can be expressed as a power series centered at α with a non-zero radius of convergence. Such a power series is called a *Taylor series* representation of the function. In the case that $\alpha = 0$, a Taylor series is often called a *MacLaurin series*.

We know that power series (and therefore all possible Taylor series) are C^∞. There is a nice theorem that provides the reverse implication.

Theorem 3.3.2. *A function f is C^∞ at a point $\alpha \in \mathbb{R}$ if and only if there exists $R > 0$ such that f is analytic on $(\alpha - R, \alpha + R)$.*

This theorem answers the questions of which functions have Taylor series representations: any function which is infinitely differentiable can be expressed as a series, but no other functions can be so expressed.

3.3.2 Calculating Coefficients

The previous section defined a class of analytic functions, but it didn't tell us how to actually find the series for these functions. We get to choose the centre point α, so we need to know how to calculate

the coefficients c_n. Assuming we have a series expression of $f(x)$, let's look at the values of f and its derivatives. Then we calculate the values of the derivatives at the centre point α.

$$f(\alpha) = \sum_{n=0}^{\infty} c_n(\alpha-\alpha)^n = c_0 + \sum_{n=1}^{\infty} c_n \cdot 0 = c_0 \implies c_0 = f(\alpha)$$

$$f'(\alpha) = \sum_{n=1}^{\infty} c_n n(\alpha-\alpha)^{n-1} = c_1 + \sum_{n=2}^{\infty} c_n \cdot 0 = c_1 \implies c_1 = f'(\alpha)$$

$$f''(\alpha) = \sum_{n=2}^{\infty} c_n n(n-1)(\alpha-\alpha)^{n-2} = 2c_2 + \sum_{n=3}^{\infty} c_n \cdot 0 = 2c_2 \implies c_2 = \frac{f''(\alpha)}{2}$$

$$f^{(3)}(\alpha) = \sum_{n=3}^{\infty} c_n n(n-1)(n-2)(\alpha-\alpha)^{n-3} = 6c_3 + \sum_{n=4}^{\infty} c_n \cdot 0 = 6c_3 \implies c_3 = \frac{f^{(3)}(\alpha)}{6}$$

$$f^{(4)}(\alpha) = \sum_{n=4}^{\infty} c_n n(n-1)(n-2)(n-3)(\alpha-\alpha)^{n-4} = 24c_4 + \sum_{n=5}^{\infty} c_n \cdot 0$$

$$f^{(4)}(\alpha) = 24c_4 \implies c_4 = \frac{f^{(4)}(\alpha)}{24}$$

We generalize the pattern to write a general expression for the nth coefficient.

$$c_n = \frac{f^{(n)}(\alpha)}{n!}$$

Now we have a way to calculate the coefficient in terms of the derivatives of $f(x)$ at the chosen centre point. Therefore, to find a series representation of $f(x)$ centered at α (assuming $f(x)$ is analytic at α), we use this expression above to calculate the coefficients. We summarize this in a proposition.

Proposition 3.3.3. *If f is analytic at α, then the Taylor series for f has this form:*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n$$

The expression for the coefficients c_n allows for another important result.

Proposition 3.3.4. *(Uniqueness of Coefficients) Two power series centered at the same point are equal if an only if every coefficient is equal.*

Proof. Say we have an equation of power series:

$$\sum_{n=0}^{\infty} c_n(x-\alpha)^n = \sum_{n=0}^{\infty} b_n(x-\alpha)^n$$

The coefficients are determined by the derivatives. But the functions are the same, so they must have the same derivatives at α. Therefore, both b_n and c_n must be calculated by $\frac{f^{(n)}(\alpha)}{n!}$, hence $b_n = c_n$. $\square$

Uniqueness of coefficients is very important for doing algebra with series. If two series are equal, we can then pass to the equality of each of the coefficients to get explicit equations. Curiously, since all the coefficients are determined by the derivatives at the centre point, this means that the derivatives at the centre point encode the entire behaviour of the function (inside the radius of convergence). This is a surprising result, since functions can have a wide range of behaviours far away from their centre points.

3.4 Examples

Let's try to calculate some Taylor series for important functions.

Example 3.4.1. We start with the most important function in calculus: e^x. The derivatives of e^x are just e^x. If we centre a series at $x = 0$, then all these derivatives evaluate to 1. Therefore

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \dots$$

We can check that the radius of convergence for this series is $R = \infty$, so this is an expression for e^x which works for all real numbers.

As an aside, this finally allows for the proper definition of the exponential function. For $r = \frac{a}{b}$ a rational number, $a^r = \sqrt[b]{x^a}$, which was well understood. But if r is irrational, we previously had no idea what a^r actually was nor how to calculate it. We worked on faith that the exponential function e^x was well defined for irrational numbers. Now, however, we can use this series. The value of e^π, which was completely opaque and mysterious before, is now given by a series.

$$e^\pi = \sum_{n=0}^{\infty} \frac{\pi^n}{n!} = 1 + \pi + \frac{\pi^2}{2} + \frac{\pi^3}{6} + \frac{\pi^4}{24} + \frac{\pi^5}{120} + \dots$$

Other important properties of the exponential function can be calculated from the series. Let's differentiate the series. (We use a shift in the series in the last step.)

$$\frac{d}{dx} e^x = \frac{d}{dx} \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=1}^{\infty} \frac{1}{n!} n x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = e^x$$

This recovers the fact that the exponential function is its own derivative.

Example 3.4.2. Let's integrate the geometric series (we set the integration constant to zero).

$$\int \sum_{n=0}^{\infty} x^n dx = \int \frac{1}{1-x} dx$$

$$\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} + C = -\ln|1-x| + C$$

$$\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} = -\ln(1-x)$$

This gives a Tayor series for $-\ln(1-x)$ centered at $\alpha = 0$. Integration can be a convenient way to calculate a series, since we didn't have to calculate all the coefficients directly.

Example 3.4.3. We remarked in the previous section that integration was easy for series. Let's look at the function e^{x^2}. It has no elementary anti-derivative, so we unable to integrate it with conventional methods. However, if we put x^2 into the series for the exponential function, we get a series for e^{x^2}.

$$e^{x^2} = \sum_{n=0}^{\infty} \frac{x^{2n}}{n!}$$

Since this is a series, we can integrate it.

$$\int e^{x^2} dx = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)n!} + C$$

This new series is the anti-derivative of e^{x^2}. We knew such a function should exist, and now we have a representation of it as a Taylor series. (The series has infinite radius of convergence).

Example 3.4.4. The Taylor series for sine and cosine are important examples. Centered at $x = 0$, the derivatives of $\sin x$ form a cycle: $\sin x$, $\cos x$, $-\sin x$, and $-\cos x$. Evaluated at $x = 0$, these gives values of 0, 1, 0, and -1. Therefore, we get the following expressions for the coefficient of the Taylor series. (Note we need to group the coefficients into odds and evens, writing $n = 2k$ for evens and $n = 2k+1$ for odds).

$$\begin{aligned}
c_0 &= f(0) = 0 & c_1 &= f'(0) = 1 \\
c_2 &= \frac{f''(0)}{2!} = 0 & c_3 &= \frac{f'''(0)}{3!} = \frac{-1}{3!} \\
c_4 &= \frac{f^{(4)}(0)}{4!} = 0 & c_5 &= \frac{f^{(5)}(0)}{5!} = \frac{1}{5!} \\
c_6 &= \frac{f^{(6)}(0)}{6!} = 0 & c_7 &= \frac{f^{(7)}(0)}{7!} = \frac{-1}{7!} \\
c_8 &= \frac{f^{(8)}(0)}{8!} = 0 & c_9 &= \frac{f^{(9)}(0)}{9!} = \frac{1}{9!} \\
c_{2k} &= 0 & c_{2k+1} &= \frac{(-1)^k}{(2k+1)!}
\end{aligned}$$

Using these coefficients, the Taylor series for sine centered at $\alpha = 0$ is this series:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

The radius of convergence of this series is $R = \infty$, so it expresses $\sin x$ for all real numbers. We can use similar steps to find the Taylor the series for cosine.

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}$$

The radius of convergence of this series is also $R = \infty$.

Example 3.4.5. Consider $f(x) = \ln x$ centered at $\alpha = 1$.

$$f'(x) = \frac{1}{x} \qquad f''(x) = \frac{-1}{x^2}$$
$$f'''(x) = \frac{2}{x^3} \qquad f^{(4)}(x) = \frac{-6}{x^4}$$

We look for a general patterns. There are three pieces: an alternating sign, a factorial multiplication growing in the numerator, an a power growing in the denominator. Careful to match the indices correctly to the first few elements of the pattern.

$$f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{x^n} \qquad f^{(n)}(1) = (-1)^{n-1}(n-1)!$$

Once we have the general pattern, we evaluate it at the centre point and then we put it into the Taylor series form.

$$\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(n-1)!}{n!}(x-1)^n = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}(x-1)^n$$

The radius of convergence is 1, found by ratio test.

Example 3.4.6. Consider $f(x) = \frac{1}{x^2}$ centered at $\alpha = 3$.

$$f'(x) = \frac{-2}{x^3} \qquad f''(x) = \frac{6}{x^4} \qquad f'''(x) = \frac{-24}{x^5}$$

We look for a general pattern as before.

$$f^{(n)}(x) = \frac{(-1)^n(n+1)!}{x^{n+2}}$$

We evaluate at the centre point.

$$f^{(n)}(3) = \frac{(-1)^n(n+1)!}{3^{n+2}}$$

We put this into the general Taylor series form.

$$\ln x = \sum_{n=1}^{\infty} \frac{(-1)^n(n+1)!}{3^{n+2}n!}(x-3)^n = \sum_{n=1}^{\infty} \frac{(-1)^n(n+1)}{3^{n+2}}(x-3)^n$$

The radius of convergence is 3, found by ratio test.

Example 3.4.7. Consider a function which has the following sequence of derivatives at $x = 0$.

$$0, 1, 2, 0, -1, 4, 0, 1, 8, 0, -1, 16, 0, 1, 32, 0, -1, 64, \ldots$$

The pattern has a cycle of threes. All $3n$ terms are 0. All $3n+1$ terms are $(-1)^n$. All $3n+2$ terms are 2^n. Therefore, the series is best expressed in two pieces.

$$f(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(3n+1)!}x^{3n+1} + \sum_{n=0}^{\infty} \frac{2^n}{(3n+2)!}x^{3n+2}$$

The radius of convergence is ∞, found by a ratio test.

3.5 Exponentials and Trigonometry

Taylor series allow us to investigate the strange relationship between exponentials and trigonometry. To understand this relationship, let's look at the series representation of the trigonometric functions.

The Taylor series for sine and cosine were calculated in the previous setion.

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \qquad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+!)!}$$

The Taylor series for hyperbolic sine and hyperbolic cosine can be calculated similarly.

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \qquad \sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+!)!}$$

The similarities are uncanny: the only difference is the presence of the alternating $(-1)^n$ term. (This is yet another reason that trigonometric and hyperbolic functions share similar names and identities). The exponential definition of the hyperbolics gives $e^x = \sinh x + \cosh x$. This identity is relatively easy to see in the series itself: if we combine the terms of the two hyperbolic series, we get all the terms of the exponential series.

For the trigonometric functions, the connection is similar. However, it's not visible when we work solely with the real numbers. We need complex numbers to see the connection. Therefore, let $\imath^2 = -1$, or equivalently let $\pm\imath$ be the roots of $x^2 + 1 = 0$. That is, $\imath$ is a square root of -1. The number $\imath$ is just a constant, so we can include it when we do series. The arithmetic changes a bit, but it behaves normally with multiplication and addition. We're going to calculate $e^{\imath x}$ and see what happens.

$$e^{\imath x} = \sum_{n=0}^{\infty} \frac{(\imath x)^n}{n!} = \sum_{n=0}^{\infty} \frac{\imath^n x^n}{n!}$$

We split this series up into two series, one with the odd terms and one with the even terms.

$$= \sum_{n=0}^{\infty} \frac{\imath^{2n} x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{\imath^{2n+1} x^{2n+1}}{(2n+1)!}$$

Now, anywhere we see i^2, we can replace it with -1.

$$= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{\imath(-1)^n x^{2n+1}}{(2n+1)!}$$

Finally, pulling out i in the second terms, we find some familiar Taylor series.

$$= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} + \imath \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$$

$$e^{\imath x} = \cos x + \imath \sin x$$

This result is called Euler's formula. It is very similar to the hyperbolic identity noted previously, $e^x = \cosh x + \sinh x$, but we need complex numbers to find it. We can rearrange this to get the opposite direction, giving exponential definitions for trigonometry to mirror hyperbolics:

$$\sin x = \frac{e^{\imath x} - e^{-\imath x}}{2\imath} \qquad \cos x = \frac{e^{\imath x} + e^{-\imath x}}{2}$$

There is one very notable case of Euler's formula: let $x = \pi$.

$$e^{\imath\pi} = \cos\pi + \imath \sin\pi = 1$$

Alternatively, we can evaluate the trigonometric terms for someting very lovely.

$$e^{\imath\pi} + 1 = 0$$

This equation is perhaps the cleanest, most elegant relationship in mathematics. It relates, with no other terms, the five most important numbers: 0, 1, π, e and $\imath$.

3.6 Non-Elementary Functions

In addition to finding connections between known functions, Taylor series can help us construct entirely new functions. These are often called non-elementary functions (the elementary functions are those which we already have worked with: polynomials, roots, exponentials, logarithms, trig, and hyperbolics).

Example 3.6.1. The Bessel functions of order $k \in \mathbb{N}$ are given by this series.

$$J_k(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{2^{2n+k}((n+k)!)^2}$$

The Bessel functions are like the trigonometric functions, but the terms in the denominators are larger. They oscilate like trig functions, but with decaying amplitude. They are important for spherical and circular waves, such as sound waves or ripples on a pond.

Example 3.6.2. The Bessel-Clifford functions are givesn by this series.

$$C_k(x) = \sum_{n=0}^{\infty} \frac{\pi(k+n)x^n}{n!}$$

Example 3.6.3. The Polylogarithm functions are given by this series. (Note that for $s = 1$ the polylogarithm is $Li_1(x) = -\ln(1-x)$, the conventional logarithm).

$$Li_s(x) = \sum_{n=0}^{\infty} \frac{x^n}{n^s}$$

These three examples are just the very start of a huge world of non-elementary functions.

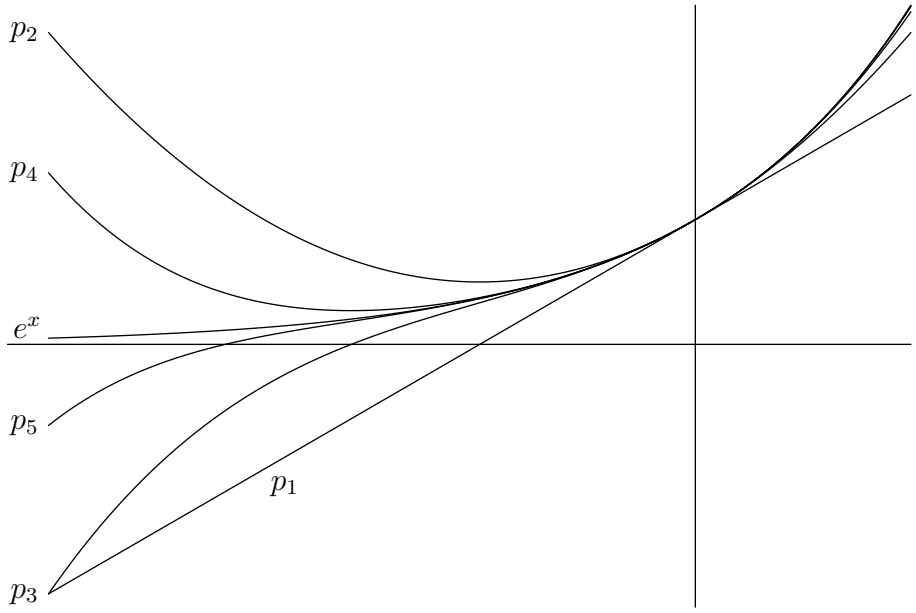

Figure 3.1: Polynomials Approximations to e^x.

3.7 Approximation and Taylor Polynomials

The previous section defined Taylor series for analytic functions. Instead of taking the terms and coefficients all the way to infinity, we could instead truncate the process at some degree. The result is a polynomial which serves as a polynomial approximation to the function.

Definition 3.7.1. If $f(x)$ is analytic, its dth *Taylor polynomials* centered at α is the truncation of its Taylor series, stopping at $(x-\alpha)^d$.

$$f(x) \cong \sum_{n=0}^{d} \frac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n$$

Taylor polynomials give the best possible polynomial approximations to analytic functions.

Example 3.7.2. Look at the exponential function e^x centered at $\alpha = 0$. We have its Taylor series from the previous section. These are its polynomial approximations. Their graphs are shown in Figure 3.1.

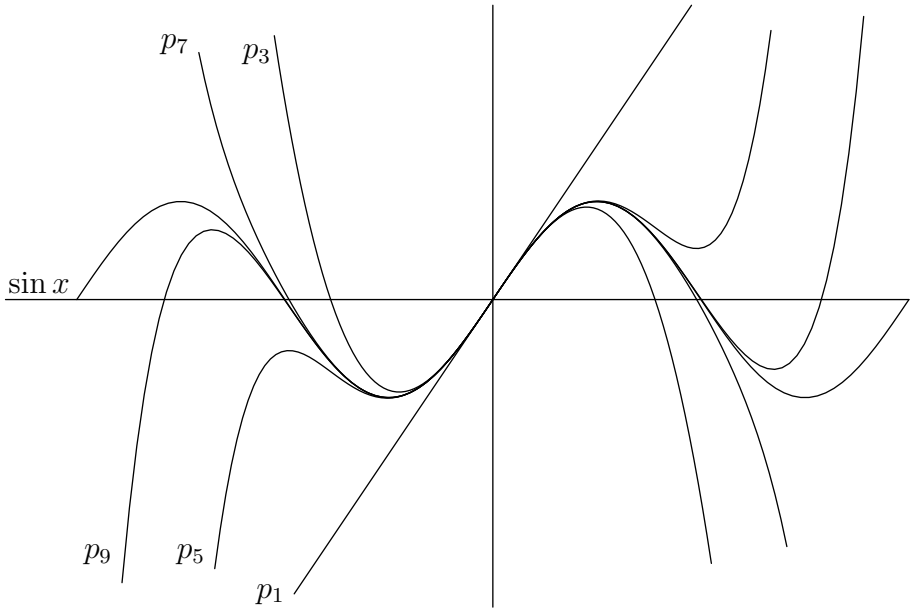

Figure 3.2: Polynomials Approximations to $\sin x$.

$$e^x \cong \sum_{n=0}^{1} \frac{1}{n!}x^n = 1 + x = p_1$$

$$e^x \cong \sum_{n=0}^{2} \frac{1}{n!}x^n = 1 + x + \frac{x^2}{2} = p_2$$

$$e^x \cong \sum_{n=0}^{3} \frac{1}{n!}x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} = p_3$$

$$e^x \cong \sum_{n=0}^{4} \frac{1}{n!}x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} = p_4$$

$$e^x \cong \sum_{n=0}^{5} \frac{1}{n!}x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} = p_5$$

Example 3.7.3. The approximations for sine only have odd exponents, since there are only odd monomials in the Taylor series for sine. These are the first few approximations. Their graphs are

shown in Figure 3.2

$$\sin x \cong \sum_{k=0}^{0} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x = p_1$$
$$\sin x \cong \sum_{k=0}^{1} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} = p_3$$
$$\sin x \cong \sum_{k=0}^{2} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} = p_5$$
$$\sin x \cong \sum_{k=0}^{3} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} = p_7$$
$$\sin x \cong \sum_{k=0}^{4} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} = p_9$$

The main application of approximation is calculating values of transcendental functions. We can't directly calculate their values using basic arithmetic; we need a method. Before the convenience of calculator and computer reference, mathematicians, scientists and engineers carried around large books of tables of values of trig, exponential and logarithmic function.

Polynomials are particularly useful as approximation tools since they involve only the basic operations of arithmetic. Computers can calculate with the basic operations of arithmetic, so computers can understand polynomials. If we want to program a computer or calculator to calculate values of e^x or $\sin x$ or $\ln x$ or some other transcendental function, Taylor series are one of the best techniques.

Example 3.7.4. The logarithm is a transcendental function which can't be directly calculated. We had a Taylor series for the logarithm in the previous section.

$$-\ln(1-x) = \sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} dx$$

Using some clever arithmetic, we can write $\ln 2 = -\ln \frac{1}{2} = -\ln\left(1 - \frac{1}{2}\right)$. If we truncate the series at degree 6, we have this approximation for $\ln 2$.

$$\ln 2 \cong 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \left(\frac{1}{2}\right)^2 + \frac{1}{3}\left(\frac{1}{2}\right)^3 + \frac{1}{4}\left(\frac{1}{2}\right)^4 + \frac{1}{5}\left(\frac{1}{2}\right)^5 + \frac{1}{6}\left(\frac{1}{2}\right)^6$$
$$\ln 2 \cong \frac{1}{2} + \frac{1}{8} + \frac{1}{24} + \frac{1}{64} + \frac{1}{160} + \frac{1}{384}$$
$$\ln 2 \cong \frac{1327}{1920} = 0.6911458333333\ldots = 0.69114583$$

This is not to far off from the value of $\ln 2 = 0.69314\ldots$, accurate to the thousandths place.

Example 3.7.5. There are many ways in mathematics to find approximations to numbers. Recall that the alternating harmonic series also summed to $\ln 2$. If we truncate that series after ten steps, we get this approximation:

$$\ln 2 \cong 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} = \frac{1627}{2520} = 0.645\overline{634920}$$

This expression is a poorer approximation for $\ln 2$; we would need to go much farther down the alternating harmonic series to match the precision of the Taylor series.

3.8 Error Analysis of Series

When we truncate a series to get a Taylor polynomial, we obviously change the function. But how much have we changed it? This change is called the error of the Taylor polynomial, so we ask: what is the error introduced by truncating an infinite series? The study of approximations in mathematics is all about understanding and controlling error.

For Taylor series, there is a theorem that controls the error of Taylor polynomial approximation. This is sometimes called Taylor's Theorem or the Lagrange Error Bound.

Theorem 3.8.1. *Let f be analytic at α with a positive radius of convergence R and let $d \in \mathbb{R}$ with $0 < d < R$. Let $M \in \mathbb{R}_{>0}$ such that on the interval $(\alpha - d, \alpha + d)$ we have*

$$|f^{n+1}(x)| \leq M.$$

The error of the kth Taylor polynomial approximation can be written as

$$R_k = f - \sum_{n=0}^{k} c_n(x-\alpha)^n = \sum_{n=k+1}^{\infty} c_n(x-\alpha)^n.$$

Under all these assumptions, the error satisfies an inequality.

$$|R_n(x)| \leq \frac{M}{(n+1)!}|x-a|^{n+1}$$

This is a difficult theorem to understand. The basic idea is that there exists a balance between three pieces: the order of the approximation, the width of the interval, and the size of the error. Higher order approximations are more accurate, but they are also computationally more difficult; to work on computers, we want to use a higher enough order to be accurate, but a low enough order to allow the computer to finish the calculation quickly. Likewise, a small interval minimizes the error, since error tends to be larger the farther we move away from the centre of the series. We would like to work with a large interval, since it means that one series works for many difficult calculations, but we may not have the accuracy we desire on a large interval. To summarize: we want a low order, a large interval and a small error, but it is difficult to achieve all three. Sometimes, one is sacrificed for the others.

Example 3.8.2. Consider $f(x) = \sin x$. We can take $M = 1$, since the derivatives of $\sin x$ are $\pm \sin x$ and $\pm \cos x$, all of which are bounded in absolutely value by 1. Then we can ask: what order do we need in order to have $\frac{1}{10000}$ precision on $(-10, 10)$? The error calculation looks like:

$$\frac{M}{(n+1)!}|x|^{n+1} \leq \frac{1}{10^4}$$
$$\frac{1}{(n+1)!}|10|^{n+1} \leq \frac{1}{10^4}$$

This is actually quite difficult to solve for n. However, we can try some n. Here are some values for n and the error bound:

$n = 5 \quad |R_6(x)| \leq \frac{12500}{9}$ $\qquad$ $n = 10 \quad |R_{10}(x)| \leq \frac{15625000}{6237}$

$n = 15 \quad |R_{15}(x)| \leq 477$ $\qquad$ $n = 18 \quad |R_{18}(x)| \leq 82.3$

$n = 20 \quad |R_{20}(x)| \leq 19.6$ $\qquad$ $n = 22 \quad |R_{22}(x)| \leq 3.87$

$n = 24 \quad |R_{24}(x)| \leq 0.645$ $\qquad$ $n = 26 \quad |R_{26}(x)| \leq 0.0919$

$n = 28 \quad |R_{28}(x)| \leq 0.0114$ $\qquad$ $n = 30 \quad |R_{30}(x)| \leq 0.00122$

$n = 31 \quad |R_{31}(x)| \leq 0.000381$ $\qquad$ $n = 32 \quad |R_{32}(x)| \leq 0.000116$

$n = 33 \quad |R_{33}(x)| \leq 0.0000339$

So, finally, when $n = 33$ we have the desired accuracy. We need an order 33 polynomial to achieve the desired precision on the desired interval.

What about smaller intervals? Say we only need the interval $\left(\frac{-1}{10}, \frac{1}{10}\right)$. Let's see how accurate various orders are on this interval.

$n = 5 \quad |R_6(x)| \leq 1.39 \times 10^{-9}$

$n = 10 \quad |R_{10}(x)| \leq 2.51 \times 10^{-19}$

$n = 15 \quad |R_{15}(x)| \leq 4.78 \times 10^{-30}$

It is much easier to be precise on this smaller interval.

Example 3.8.3. Consider the function $f(x) = \frac{1}{1-x}$ on $(\frac{-1}{2}, \frac{1}{2})$. Look at the pattern of derivatives.

$$f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}}$$

M bounds the $n+1$st derivative, and the derivatives are increasing functions, so they are maximized at $x = \frac{1}{2}$. This lets us calculate M.

$$M = \frac{(n+1)!}{(1-\frac{1}{2})^{n+2}} = 2^{n+2}(n+1)!$$

We apply the inequality in the Lagrange Error Bound.

$$R_n \leq \frac{M}{(n+1)!}d^{n+1} = \frac{2^{n+2}(n+1)!}{(n+1)!}\left(\frac{1}{2}\right)^{n+1} = 2$$

This is entirely inconclusive, for any order, showing the potential limitations of the Lagrange Error Bound; sometimes, other methdos are required for special situation. For this example, we must work directly with the series to determine its accuracy. Here is a direct calculation of a bound for the error.

$$\begin{aligned} R_n &= \sum_{k=n+1}^{\infty} x^k \leq \sum_{k=n+1}^{\infty}\left(\frac{1}{2}\right)^k = \sum_{k=n+1}^{\infty}\left(\frac{1}{2}\right)^{k-n-1}\left(\frac{1}{2}\right)^{n+1} \\ &= \left(\frac{1}{2}\right)^{n+1}\sum_{k=n+1}^{\infty}\left(\frac{1}{2}\right)^{k-n-1} = \left(\frac{1}{2}\right)^{n+1}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^{k} = \left(\frac{1}{2}\right)^{n+1}\frac{1}{1-\frac{1}{2}} \\ &= \left(\frac{1}{2}\right)^{n+1} 2 = \frac{1}{2^n} \end{aligned}$$

So the error of the nth order can be bounded by $\frac{1}{2^n}$. For $n = 10$, that is less than a thousandth. For $n = 20$, less than a millionth, and so on.

Chapter 4

Vector Geometry

4.1 Introduction

The use of vectors is common in many branches of mathematics. Physically, vectors are motivated by the need to identify points and directions in three-dimensional space. However, the physical motivation is only one among many motivations for vectors. This chapter aims to give a relatively brief introduction to vector algebra. The material is abrieviated from the Linear Algebra course, where vectors are a core topic and are investigated in full detail. Practically, this chapter aims to be sufficient for the needs of vector calculus.

4.2 Vectors and Matrices

We start with the two most important definitions in linear algebra: the vector and the matrix.

Definition 4.2.1. In linear algebra, ordinary numbers (intergers, rational numbers or real numbers) are called *scalars*.

Definition 4.2.2. A *vector* is a finite ordered list of scalars. Vectors can be written either as columns or rows. The scalars in a vector are called its *entries*, *components* or *coordinates*. If our vector has components 4, 15, π and $-e$ (in that order), we write the vector either as a row or a column.

$$\begin{pmatrix} 4 \\ 15 \\ \pi \\ e \end{pmatrix} \qquad \text{or} \qquad (4, 15, \pi, -e)$$

In Linear Algebra, column vectors are used almost exclusively. In this section, we will match that convention and use column vectors. However, in Chapters 6, 8, and 9, we will use row vectors, which are standard for that material.

Definition 4.2.3. A *matrix* is a rectangular array of scalars. If the matrix has m rows and n columns, we say it is an $m \times n$ matrix. The scalars in the matrix are called the *entries*, *components* or *coefficients*.

A matrix is written as a rectangular array of numbers, either in square or round brackets. Here are two ways of writing a particular 3×2 matrix with integer coefficients.

$$\begin{pmatrix} 5 & 6 \\ -4 & -4 \\ 0 & -3 \end{pmatrix} \qquad \begin{bmatrix} 5 & 6 \\ -4 & -4 \\ 0 & -3 \end{bmatrix}$$

The rows of this matrix are vectors in $\mathbb{R}^2$.

$$\begin{pmatrix} 5 \\ 6 \end{pmatrix} \qquad \begin{pmatrix} -4 \\ -4 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

The columns of this matrix are vectors in $\mathbb{R}^3$.

$$\begin{pmatrix} 5 \\ -4 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 6 \\ -4 \\ -3 \end{pmatrix}$$

In these notes, we'll use curved brackets for matrices; however, in many texts and books, square brackets are very common. Both notations are conventional and acceptable.

If we want to write a general matrix, we can write with double indexed unknown coefficients.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

By convention, when we write an aribitrary matrix entry a_{ij} the first index tells us the row and the second index tells us the column. For example, a_{64} is the entry in the sixth row and the fourth column. In the rare occurence that we have matrices with more than 10 rows or columns, we can seperate the indices by commas: $a_{12,15}$ would be in the twelfth row and the fifteenth column. Row and column indexing is very convenient. Sometime we even write $A = a_{ij}$ as short-hand for a matrix when the size is understood or undetermined.

Definition 4.2.4. The rows of a matrix are called the *row vectors* of the matrix. Likewise the columns are called *column vectors* of the matrix.

Definition 4.2.5. The *zero matrix* is the unique matrix (one for every size $m \times n$) where all the coefficients are all zero.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

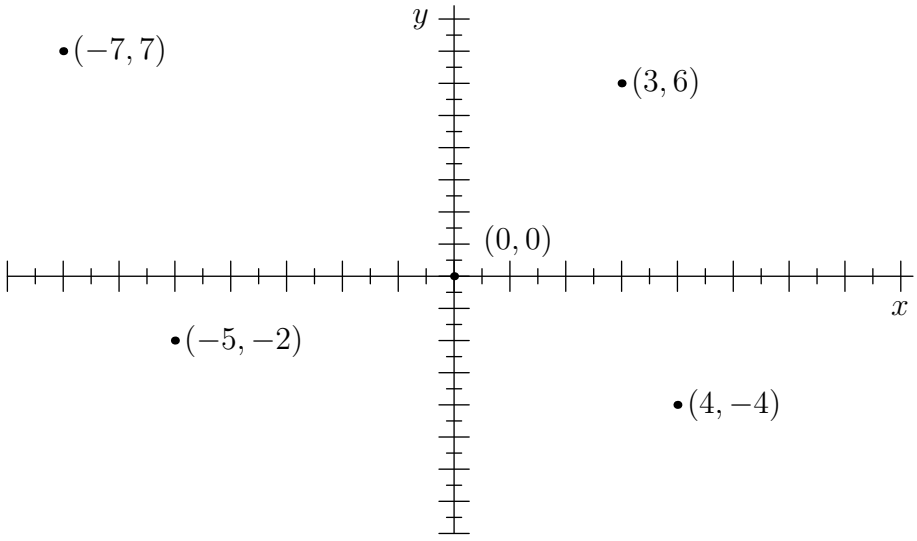

Figure 4.1: Points in the Carteian Plane $\mathbb{R}^2$

Definition 4.2.6. The *diagonal elemenents* of a matrix are the elements where the row and column index are the same. They lie on a diagonal line starting at the top left of the matrix. If an arbitrary matrix is written a_{ij}, the diagonal elements are the elements a_{ii}.

Definition 4.2.7. The *identity matrix* is the unique $n \times n$ matrix (one for each n) where the diagonal entries are all 1 and all other entries are 0.

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

4.3 The Environment

Definition 4.3.1. Let n be a positive integer. *Real Euclidean Space* or *Cartesian Space*, written $\mathbb{R}^n$, is the set of all vectors with n real number entries. In the notation of this chapter, an arbitrary element of $\mathbb{R}^n$ is written as a column.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

We say that $\mathbb{R}^n$ has *dimension* n.

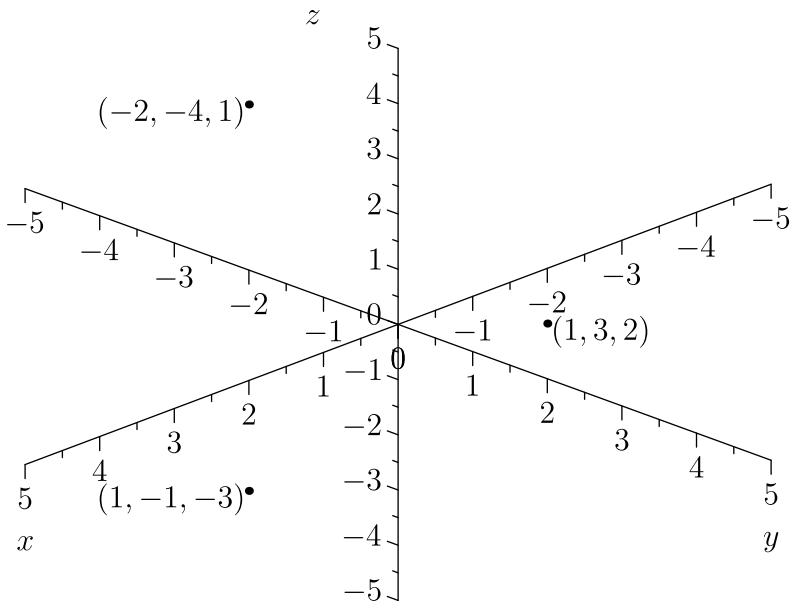

Figure 4.2: Points in Cartesian three-space $\mathbb{R}^3$

Definition 4.3.2. The scalars x_i in a vector are called the *entries*, *coordinates* or *components* of that vector. Specifically, x_1 is the first coordinate, x_2 is the second coordinate, and so on. For $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^4$, we use the letters x, y, z, w instead of x_i.

$$\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \qquad \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} \in \mathbb{R}^4$$

Definition 4.3.3. In any $\mathbb{R}^n$, the *origin* is the unique point given by a vector of zeros. It is also called the zero vector.

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

It is considered the centre point of Cartesian space.

Cartesian space, particularly the Cartesian plane, is a familiar object from high-school mathematics. We usually visualize Cartesian space by drawing axes, one in each independent perpendicular direction. In this visualization, the vector $\begin{pmatrix} a \\ b \end{pmatrix}$ corresponds to the unique point we get moving a units in the direction of the x axis and b units in the direction of the y axis, as in Figure 4.1.

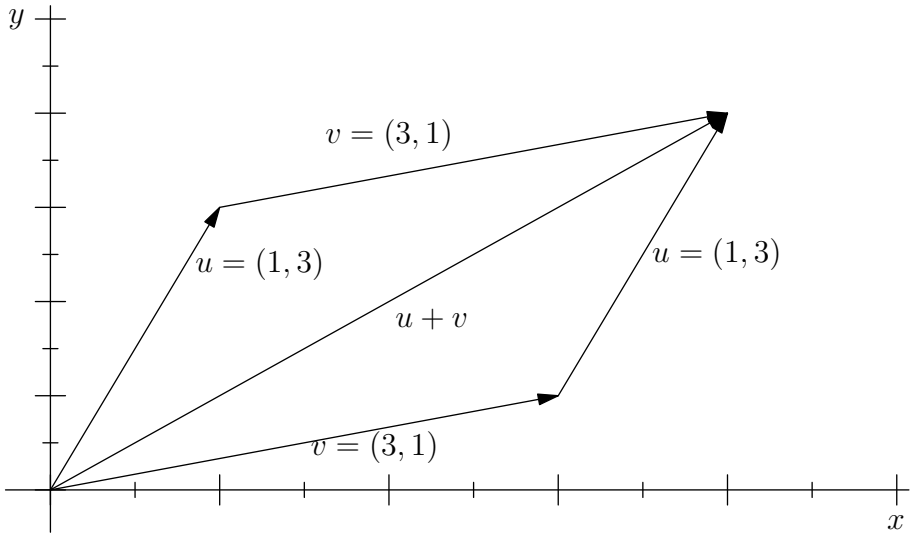

Figure 4.3: Visualizing Vector Addition

As with $\mathbb{R}^2$, the point $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3$ is the unique point we find by moving a units in the x direction, b units in the y direction and c units in the z direction, as shown in Figure 4.2.

When we visualize $\mathbb{R}^2$, we conventionally write the x axis horizontally, with a positive direction to the right, and the y axis vertically, with a positive direction upwards. There are similar conventions for $\mathbb{R}^3$.

Definition 4.3.4. A choice of axis directions in a visualization of $\mathbb{R}^n$ is called an *orientation*.

While we can visualize $\mathbb{R}^2$ and $\mathbb{R}^3$ relatively easily and efficiently, but we can't visualize any higher $\mathbb{R}^n$. However, that doesn't prevent us from working in higher dimensions. We need to rely on the algebraic descriptions of vectors instead of the drawings and visualizations of $\mathbb{R}^2$ and $\mathbb{R}^3$.

In our visualizations of $\mathbb{R}^2$ and $\mathbb{R}^3$, we see the different axes as fundementally different perpendicular directions. We can think of $\mathbb{R}^2$ as the space with two independent directions and $\mathbb{R}^3$ as the space with three independent directions. Similarly, $\mathbb{R}^4$ is the space with four perpendicular, independent directions, even though it is impossible to visualize such a thing. Likewise, $\mathbb{R}^n$ is the space with n independent directions.

4.4 Linear Operations

The previous section built the environment for linear algebra: $\mathbb{R}^n$. We want to know what operations we can perform on $\mathbb{R}^n$; there are several.

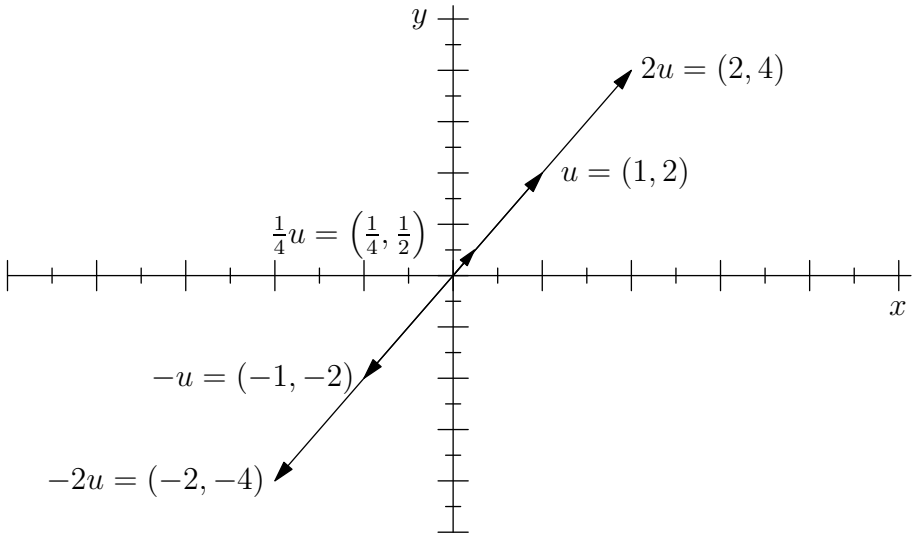

Figure 4.4: Visualizing Scalar Multiplication

Definition 4.4.1. The *sum* of two vectors u and v in $\mathbb{R}^n$ is the sum taken componentwise.

$$u + v = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

Vector addition is visualized by putting the start of one vector at the end of the other vector, as in Figure 4.3. Note that we can only add two vectors in the same dimension. We can't add a vector in $\mathbb{R}^2$ to a vector in $\mathbb{R}^3$.

Definition 4.4.2. If u is a vector in $\mathbb{R}^n$ and $a \in \mathbb{R}$ is a real number, then the *scalar multiplication* of u and a is multiplication by a in each component of u. By convention, scalar multiplication is written with the scalar on the left of the vector.

$$au = a \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} au_1 \\ au_2 \\ \vdots \\ au_n \end{pmatrix}$$

Multiplication stretches of shrink the vector while preserving the direction or flipping if if the scalar is negative. This is visualized in Figure 4.4.

Though there will be other 'multiplications' to come, we generally say that we can't multiply vectors together in any way reminiscent of numbers. Instead, we can only multiply by scalars.

Scalar multiplication also lets us define the difference between vectors.

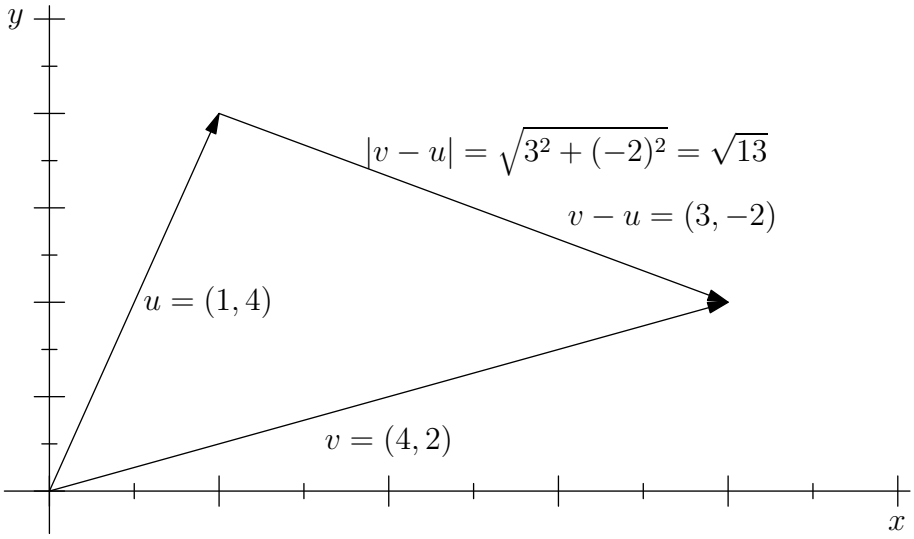

Figure 4.5: Vector Subtraction and Distance Between Vectors

Definition 4.4.3. The difference between two vectors u and v is the vector $u + (-1)v$, defined using addition and scalar multiplication. This works out to be componentwise subtraction.

$$u - v = u + (-1)v = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + (-1) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 - v_1 \\ u_2 - v_2 \\ \vdots \\ u_n - v_n \end{pmatrix}$$

Subtraction is visualized as finding a vector that completes a triangle, as in Figure 4.5.

Definition 4.4.4. Given a set of scalars (such as $\mathbb{R}$), whenever we find a mathematical structure which has the two properties of addition and scalar multiplication, we call the structure *linear*. $\mathbb{R}^n$ is a linear space, because vectors allow for addition and scalar multiplication.

Definition 4.4.5. The *length* of a vector u in $\mathbb{R}^n$ is written $|u|$ and is given by a generalized form of the Pythagorean rule for right triangles.

$$|u| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

This length is also called the *norm* of the vector. A vector of length one is called a *unit vector*.

If we think of vectors as directions from the origin towards a point, this definition of length gives exactly what we expect: the physical length of that arrow in $\mathbb{R}^2$ and $\mathbb{R}^3$. Past $\mathbb{R}^3$, we don't have a natural notion of length. This definition serves as a reasonable generalization to $\mathbb{R}^4$ and higher dimensions which we can't visualize. Note also that $|u| = 0$ only if u is the zero vector. All other vectors have positive length.

Often the square root is annoying and we find it convenient to work with the square of length.

$$|u|^2 = u_1^2 + u_2^2 + \ldots + u_n^2$$

The notions of length and difference allow us to define the distance between two vectors.

Definition 4.4.6. The *distance between two vectors* u and v in $\mathbb{R}^n$ is the length of their difference: $|u - v|$.

You can check from the definition that $|u - v| = |v - u|$, so distance doesn't depend on which comes first. If $|\cdot|$ were absolute value in $\mathbb{R}$, this definition would match the notion of distance between numbers on the number line.

4.5 The Dot Product

Earlier we said that we can't multiply two vectors together. However, there are operations similar to multiplication. The following operation multiplies two vectors, but the result is a scalar instead of another vector.

Definition 4.5.1. The *dot product* or *inner product* or *scalar product* of two vectors u and v is given by the following formula.

$$u \cdot v = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

We can think of the dot product as a scalar measure of the similarity of direction between the two vectors. If the two vectors point in a similar direction, their dot product is large, but if they point in very different directions, their dot product is small. However, we already have a measure, at least in $\mathbb{R}^2$, of this difference: the angle between two vectors. Thankfully, the two measures of difference agree and the dot product can be expressed in terms of angles.

Definition 4.5.2. The *angle* θ between two non-zero vectors u and v in $\mathbb{R}^n$ is given by the equation

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

We assume the angle between any two vectors is $\theta \in [0, \pi]$. If $\theta > \pi$, then we can look at the other side of the angle to get something between 0 and π.

This definition matches with the definition of angles in $\mathbb{R}^2$ and $\mathbb{R}^3$, which we can visualize. However, this serves as a new definition for angles between vectors in all $\mathbb{R}^n$ when $n \geq 4$. Since we can't visualize those spaces, we don't have a way of drawing angles and calculating them with conventional trigonometry. This definition allows us to extend angles in a completely algebraic way.

Definition 4.5.3. Two vectors u and v in $\mathbb{R}^n$ are called *orthogonal* or *perpendicular* or *normal* if $u \cdot v = 0$.

Now that we have some operations, it is useful to see how the various operations interact. The following list shows some interactions between addition, scalar multiplication, length and the dot product. Some of these are easy to establish from the definition and some take more work.

Proposition 4.5.4. *Let u, v, w be vectors in $\mathbb{R}^n$ and let a be a scalar in $\mathbb{R}$.*

$$\begin{aligned}
u + v &= v + u && \textit{Commutative Law for Vector Addition} \\
a(u + v) &= au + av && \textit{Distributive Law for Scalar Multiplication} \\
u \cdot v &= v \cdot u && \textit{Commutative Law for the Dot Product} \\
u \cdot u &= |u|^2 && \\
u \cdot (v + w) &= u \cdot v + u \cdot w && \textit{Distributive Law for the Dot Product} \\
u \cdot (av) &= (au) \cdot v = a(u \cdot v) && \\
|u + v| &\leq |u| + |v| && \textit{Triangle Inequality} \\
|au| &= |a||u| &&
\end{aligned}$$

The last line deserves some attention for the notation. When we write $|a||u|$, $|a|$ is an absolute value of a real number and $|u|$ is the length of a vector. The fact that they have the same notation is frustrating, but these notations are conventional.

4.6 The Cross Product

The dot product is an operation which can be performed on any two vectors in $\mathbb{R}^n$ for any $n \geq 1$. There are no other conventional products that work in all dimensions. However, there is a special product that works in three dimensions.

Definition 4.6.1. Let $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ and $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ be two vectors in $\mathbb{R}^3$. The *cross product* of u and v is written $u \times v$ and defined by this formula:

$$u \times v = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix}$$

The cross product differs from the dot product in several important ways. First, it produces a new vector in $\mathbb{R}^3$, not a scalar. For this reason, when working in $\mathbb{R}^3$, the dot product is often refered to as the *scalar product* and the cross product as the *vector product*. Second, the dot product measures, in some sense, the similarity of two vectors. The cross product measures, in some sense, the difference between two vectors. The cross product has greater magnitude if the vectors are closer to being perpendicular.

If θ is the angle between u and v, the dot product was expressed in terms of $\cos\theta$. This measures similarity, since $\cos 0 = 1$. There is a similar identity for the cross product.

$$|u \times v| = |u||v| \sin\theta$$

This identity tells us that the cross product measures difference in direction, since $\sin 0 = 0$. In particular, this tells us that $|u \times u| = 0$, implying that $u \times u = 0$ (the zero vector is the only vector which has zero length).

We can calculate a combination of the dot and cross products.

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

A similar calculation shows that $v \cdot (u \times v) = 0$. Since a dot product of two vectors is zero if and only if the vectors are perpendicular, the vector $v \times u$ is perpendicular to both u and v. This property turns out to be very useful for describing planes in Sections 6.5 and 8.9. There are two other identities for the cross product that we will use in Chapter 8, which we state here without proof.

Proposition 4.6.2. *Let $u, v, w \in \mathbb{R}^3$ be vectors.*

$$u \times (v \times w) = (u \cdot w)v - (u \cdot v)w$$

Proposition 4.6.3. *Let $u, v, w \in \mathbb{R}^3$ be vectors.*

$$u \cdot (v \times w) = v \cdot (w \times u) = w \cdot (u \times v)$$

An important application of the cross product is describing rotational motion. Linear mechanics describes the motion of an object through space but rotational mechanics describes the rotation of an object independently of its movement through space. A force on an object can cause both kinds of movement, obviously. The table below summarizes the parallel questions of linear motion and rotational motion in $\mathbb{R}^3$.

Linear Motion	Rotational Motion
Straight line in a vacuum	Continual spinning in a vacuum
Direction of motion	Axis of spin
Force	Torque
Momentum	Angular Momentum
Mass (resistance to motion)	Moment of Intertia (resistance to spin)
Velocity	Frequency (Angular Velocity)
Acceleration	Angular Acceleration

How do we describe torque? If there is a linear force applied to an object which can only rotate around an axis, and if the linear force is applied at a distance r from the axis, we can think of the force F and

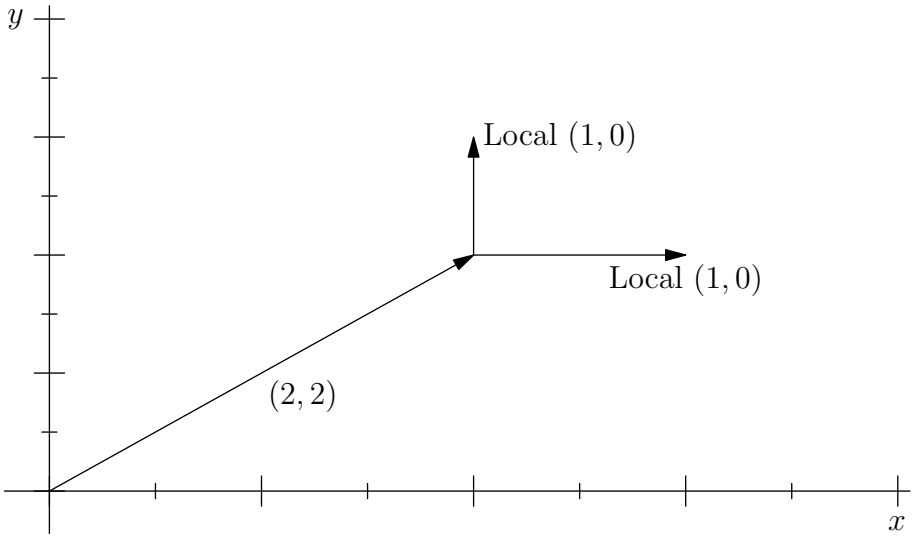

Figure 4.6: Local Direction Vectors

the distance r as vectors. The torque is then $\tau = r \times F$. Notice that $|\tau| = |r||F| \sin\theta$, indicating that linear force perpendicular to the radius gives the greatest angular acceleration. That makes sense. If F and r are colinear, that means we are pushing directly along the axis and no rotation can occur.

The use of cross products in rotational dynamics is extended in many interesting ways. In fluid dynamics, local rotational products of the fluid result in turbulence, helicity, vorticity and similar effects. Tornadoes and hurricanes are particularly extreme examples of vortices and helices in the fluid which is the atmosphere. All the descriptions of the force and motion of these vortices involve cross products in the vectors describing the fluid.

4.7 Local Direction Vectors

We've already spoken about the distinction between elements of $\mathbb{R}^n$ as points and vectors. There is another important subtlety that shows up all throughout vector geometry. In addition to thinking of vectors as directions starting at the origin, we can think of them as directions starting anywhere in $\mathbb{R}^n$. We call these local direction vectors.

For example, at the point $\binom{2}{2}$ in $\mathbb{R}^2$, we could think of the local directions $\binom{1}{0}$ or $\binom{0}{1}$. These are not directions starting from the origin, but starting from $\binom{2}{2}$ *as if that were the origin.*

Using vectors to define local directions is a particularly useful tool. A standard example is camera location in a three dimensional virtual environment. First, you need to know the location of the

camera, which is an ordinary vector starting from the origin. Second, you need to know what direction the camera is pointing, which is a local direction vector which treats the camera location as the current origin.

One of the most difficult things about learning vector geometry is becoming accustomed to local direction vectors. We don't always carefully distinguish between vectors at the origin and local direction vectors; often, the difference is implied and it is up to the reader/student to figure out how the vectors are being used.

4.8 Linear and Affine Subspaces

In addition to vectors, we want to consider various geometric objects that live in $\mathbb{R}^n$.

Definition 4.8.1. A *linear subspace* of $\mathbb{R}^n$ is a non-empty set of vectors L with two properties.

- If $u, v \in L$ then $u + v \in L$.
- If $u \in L$ and $a \in \mathbb{R}$ then $av \in L$.

These are the two basic operations on $\mathbb{R}^n$: we can add vectors and we can multiply by scalars. Linear subspaces are just subsets where we can still perform both operations and remain in the subset.

Geometrically, vector addition and scalar multiplication produce flat objects: lines, planes, and their higher-dimensions analogues. Also, since we can take $a = 0$, we must have $0 \in L$. So linear subspaces can be informally defined as flat subsets which include the origin.

Definition 4.8.2. An *affine subspace* of $\mathbb{R}^n$ is a non-empty set of vectors A which can be described as a sum $v + u$ where v is a fixed vector and u is any vector in some fixed linear subspace L. With some abuse of notation, we can write this as

$$A = u + L.$$

We think of affine subspaces as flat spaces that may be offset from the origin. The vector u is called the *offset vector*. Affine spaces include linear spaces, since we can also take u to be the zero vector and have $A = L$. Affine objects are the lines, planes and higher dimensional flat objects that may or may not pass through the origin.

Notice that we defined both affine and linear subspaces to be non-empty. The empty set $\emptyset$ is *not* a linear or affine subspace.

We need ways to algebraically describe linear and affine substapces. There are two main approaches: loci and spans, but these notes will only focus on loci.

4.9 Loci

Definition 4.9.1. Consider any set of linear equations in the variables $x_1, x_2, \ldots, x_n$. The *locus* in $\mathbb{R}^n$ of this set of equations is the set of vectors which satisfy *all* of the equations. The plural of locus is *loci*.

In general, the equations can be of any sort. The unit circle in $\mathbb{R}^2$ is most commonly defined as the locus of the equation $x^2 + y^2 = 1$. The graph of a function is the locus of the equation $y = f(x)$. However, in linear algebra, we exclude curved objects. We're concerned with linear/affine objects: things which are straight and flat.

Definition 4.9.2. A *linear equation* in variables $x_1, x_2, \ldots x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = c.$$

where a_i and c are real numbers

Proposition 4.9.3. *Any linear or affine subspace of $\mathbb{R}^n$ can be described as the locus of some number of linear equations. Likewise, any locus of any number of linear equations is either an affine subspace of $\mathbb{R}^n$ or the empty set.*

The best way to think about loci is in terms of restrictions. We start with all of $\mathbb{R}^n$ as the locus of no equations, or of the equation $0 = 0$. There are no restrictions. Then we introduce equations. Each equation is a restriction on the available points. If we work in $\mathbb{R}^2$, adding the equation $x = 3$ restricts us to a vertical line passing through the x-axis at $\binom{3}{0}$. Likewise, if we were to use the equation $y = 4$, we would have a horizontal line passing through the y-axis at $\binom{0}{4}$. If we consider the locus of *both* equations, we have only one point remaining: $\binom{3}{4}$ is the only point that satisfies both equations.

In this way, each additional equation potentially adds an additional restriction and leads to a smaller linear or affine subspace. The next three definitions give familiar names for loci with one equation, hence one restriction.

Definition 4.9.4. A *line* in $\mathbb{R}^2$ is the locus of the equation $ax + by = c$ for $a, b, c \in \mathbb{R}$. In general, the line is affine. The line is a linear subspace if $c = 0$.

Definition 4.9.5. A *plane* in $\mathbb{R}^3$ is the locus of the linear equation $ax + by + cz = d$. In general, the plane is affine. The plane is a linear subspace if $d = 0$.

If we think of a plane in $\mathbb{R}^3$ as the locus of one linear equation, then the important dimensional fact about a plane is not that it has dimension two but that it has dimension one less than its ambient space $\mathbb{R}^3$.

Definition 4.9.6. A *hyperplane* in $\mathbb{R}^n$ is the locus of one linear equation: $a_1x_1 + a_2x_2 + \ldots + a_nx_n = c$. It has dimension $n - 1$. It is, in general, affine. The hyperplane is a linear subspace if $c = 0$.

4.9.1 Intersection

Definition 4.9.7. If A and B are sets, their intersection $A \cap B$ is the set of all points they have in common. The intersection of affine subspaces is also an affine subspace. If A and B are both linear subspaces, the intersection is also a linear subspace.

Example 4.9.8. Loci can easily be understood as intersections. Consider the locus of two equations, say the example we have from $\mathbb{R}^2$ before: the locus of $x = 3$ and $y = 4$. We defined this directly as a single locus. However, we could just as easily think of this as the intersection of the two lines given by $x = 3$ and $y = 4$ seperately. In this way, it is the intersection of two loci. Similarly, all loci are the intersection of planes or hyperplanes.

4.10 Standard Forms

Loci can also be defined via dot products. Consider, again, the general linear equation in $\mathbb{R}^n$.

$$a_1u_1 + a_2u_2 + \ldots + a_nu_n = c$$

Let's think of the variables u_i as the components of a variable vector $u \in \mathbb{R}^n$. We also have n scalars a_i which we can treat as a constant vector $a \in \mathbb{R}^n$.

Then we can re-write the linear equation with dot products.

$$a_1u_1 + a_2u_2 + \ldots + a_nu_n = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = a \cdot u = c$$

In this way, a linear equation specifies a certain dot product value (c) with a fixed vector (a). A plane in $\mathbb{R}^3$ is given by the equation $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = c$. The plane is precisely all vectors whose dot product with the vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ is the fixed number c.

If $c = 0$, then we get $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0$. A linear plane is the set of all vectors which are perpendicular to a fixed vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$.

Definition 4.10.1. Let P be a plane in $\mathbb{R}^3$ determined by the equation $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = c$. The vector $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ is called the *normal to the plane*.

Let H be a hyperplane in $\mathbb{R}^n$ determined by the equation

$$u \cdot a = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = c$$

The vector a is called the *normal to the hyperplane.*

If $c = 0$, the plane or hyperplane is perpendicular to the normal. This notion of orthogonality still works when $c \neq 0$. In this case, the normal is a *local* perpendicular direction from a point on the affine plane. Treating any such point as a local origin, the normal points in a direction perpendicular to all the *local direction* vectors which lie on the plane.

Now we can build a general process for finding the equation of a plane in $\mathbb{R}^3$. Any time we have a point p on the plane and two *local direction vectors* u and v which remain on the plane, we can find a normal to the plane by taking $u \times v$. Then we can find the equation of the plane by taking the dot product $p \cdot (u \times v)$ to find the constant c.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot (u \times v) = c$$

If we are given three points on a plane (p, q and r), then we can use p as the local origin and construct the local direction vectors as $q - p$ and $r - p$. The normal is $(q - p) \times (r - p)$. In this way, we can construct the equation of a plane given three points or a single point and two local directions.

4.11 Transformations of Euclidean Spaces

After a good definition of the environment ($\mathbb{R}^n$) and its objects (lines, planes, hyperplanes, etc), the next mathematical step is to understand the functions that live in the environment and affect the objects. First, we need to generalize the simple notion of a function to linear spaces. In algebra and calculus, we worked with functions of real numbers. These functions are rules $f : A \to B$ which go between subsets of real numbers. The function f assigns to each *number* in A a unique *number* in B. They include the very familiar $f(x) = x^2$, $f(x) = \sin(x)$, $f(x) = e^x$ and many others.

Definition 4.11.1. Let A and B be subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. A *function* between linear spaces is a rule $f : A \to B$ which assigns to each *vector* in A a unique *vector* in B.

Example 4.11.2. We can define a function $f : \mathbb{R}^3 \to \mathbb{R}^3$ by $f(x, y, z) = (x^2, y^2, z^2)$.

Example 4.11.3. Another function $f : \mathbb{R}^3 \to \mathbb{R}^2$ could be $f(x, y, z) = (x - y, z - y)$.

Definition 4.11.4. A *linear function* or *linear transformation* from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $f : \mathbb{R}^n \to \mathbb{R}^m$ such that for two vectors $u, v \in \mathbb{R}^n$ and any scalar $a \in \mathbb{R}$, the function must obey two rules.

$$f(u+v) = f(u) + f(v)$$
$$f(au) = af(u)$$

Informally, we say that the function respects the two main operations on linear spaces: addition of vectors and multiplication by scalars. If we perform addition before or after the function, we get the same result. Likewise for scalar multiplication.

This leads to a very restrictive but important class of functions. We could easily define linear algebra as a study of these transformations. There is an alternative and equivalent definition of linear transformation.

Proposition 4.11.5. *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if it sends linear objects to linear objects.*

Under a linear transformation points, lines, planes are changed to other points, lines, planes, etc. A line can't be bent into an arc or broken into two different lines. Hopefully some ideas are starting to fit together: the two basic operations of addition and scalar multiplication give rise to flat objects. Linear transformation preserve those operations, so they preserve flat objects. Exactly *how* they change these objects can be tricky to determine.

Lastly, because of scalar multiplication, if we take $a = 0$ we get that $f(0) = 0$. Under a linear transformation, the origin is always sent to the origin. So, in addition to preserving flat objects, linear transformation can't move the origin. We could drop this condition of preserving the origin to get another class of functions:

Definition 4.11.6. A *affine transformation* from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a transformation that preserves affine subspaces. These transformations preserve flat objects but may move the origin.

Though they are interesting, we don't spend much time with affine transformations. They can always be realized as a linear transformation combined with a shift or displacement of the entire space by a fixed vector. Since shifts are relatively simple, we can usually reduce problems of affine transformations to problems of linear transformations.

Once we understood functions of real numbers we learned to compose them. We do the same for linear transformations.

Definition 4.11.7. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^l$ be linear transformations. Then $g \circ f : \mathbb{R}^n \to \mathbb{R}^l$ is the linear transformation formed by first applying f and then g. Note that the $\mathbb{R}^m$ has to match: f outputs to $\mathbb{R}^m$, which is the input for g. Also note that the notation is written right-to-left: In $g \circ f$, the transformation f happens first, followed by g. This new transformation is called the *composition* of f and g.

4.12 Transformations of $\mathbb{R}^2$ or $\mathbb{R}^3$

It is useful for us to specialize to transformations $\mathbb{R}^2 \to \mathbb{R}^2$, in order to build experience and intuition. As we noted above, linear transformations preserve flat objects. What can we do to $\mathbb{R}^2$ that preserves lines and preserves the origin?

Proposition 4.12.1. *There are five basic types of linear transformation of* $\mathbb{R}^2$*.*

- *Rotations* about the origin (either clockwise or counter-clockwise) preserve lines. We can't rotate around any other point, since that would move the origin. Since we can choose any angle, there are infinitely many such rotations. Also, since rotating by θ radians clockwise is the same as $2\pi - \theta$ radians counter-clockwise, we typically choose to only deal with counter-clockwise rotations. Counter-clockwise rotations are considered positive and clockwise rotations are considered negative.
- *Reflections* over lines through the origin preserve lines. The line of reflection must be a line through the origin or else the reflection will move the origin.
- *Skews* are a little tricker to visualize. A skew is a transformation that takes either veritcal or horizontal lines (but not both) and tilts them diagonally. It changes squares into parallelograms. The tilted lines are still lines, so it is a linear transformation.
- *Dialations* are transformation which stretch or shink in various directions.
- *Projections* are transformations which collapse $\mathbb{R}^2$ down to a line through the origin. Two important exampels are projection onto either axis. Projection onto the x axis sends a point (a, b) to $(a, 0)$, removing the y component. Likewise, projection onto the y axis sends a point (a, b) to $(0, b)$, removing the x component. In a similar manner, we can project onto any line though the origin by sending each point to the closest point on the line. Finally, there is the projection to the origin which sends all points to the origin.

All linear transformation of $\mathbb{R}^2$ are generated by composition of transformations of these five types.

We can also specialize to transformations of $\mathbb{R}^3$, though it is not as easy to give a complete account of the basic types. However, all of the types listed for $\mathbb{R}^2$ generalize.

- Rotations in $\mathbb{R}^3$ are no longer about the origin. Instead, we have to choose an axis. Any line through the origin will do for an axis of rotation. Any rotation in $\mathbb{R}^3$ is determined by an axis of rotation and an angle of rotation about that axis.
- Reflections are also altered: instead of reflecting over a line, we have to reflect over a plane through the origin. Any plane through the origin determines a reflection.
- Skews are similarly defined: one or two directions are fixed and the remaining directions are tilted.

- Dialations are also similar, though we have three possible directions in which to stretch or compress.
- Like $\mathbb{R}^2$, we can project onto the origin, sending everything to zero, or onto a line, sending every point to the closest point on a line. Examples include projection onto the axes. However, we can also project onto planes. Sending (a, b, c) to $(a, b, 0)$, for example, removes the z coordinate; this is projection onto the xy plane.

4.13 Matrix Representation

Matrices are used to encode linear transformations. We define an action to show how a matrix can act on a vector.

Definition 4.13.1. Let $A = a_{ij}$ be a $m \times n$ matrix and let v be a vector in $\mathbb{R}^n$. There is an action of A on v, written Av, which defines a new vector in $\mathbb{R}^m$. That action is given in the following formula.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \ldots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \ldots + a_{2n}v_n \\ \vdots \\ a_{n1}v_1 + a_{n2}v_2 + \ldots + a_{nn}v_n \end{pmatrix}$$

This is a bit troubling to work out in general. Let's see what it looks like slightly more concretely in $\mathbb{R}^2$ and $\mathbb{R}^3$.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + yb \\ cx + dy \end{pmatrix}$$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{pmatrix}$$

In this way, all $m \times n$ matrices determine a method of sending vectors in $\mathbb{R}^n$ to $\mathbb{R}^m$: a function $\mathbb{R}^n \to \mathbb{R}^m$. It is not at all obvious from the definition, but this connection completely describes linear transformations.

Proposition 4.13.2. *If A is a $m \times n$ matrix, then the associated function defined by the matrix action is a linear function $\mathbb{R}^n \to \mathbb{R}^m$. Moreover, all linear functions $\mathbb{R}^n \to \mathbb{R}^m$ can be encoded this way. Finally, each linear function is encoded uniquely, i.e., each $m \times n$ matrix corresponds to a different transformation.*

In this way, the set of linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ is *exactly* the same as the set of $m \times n$ matrices. This is a very powerful result: in order to understand linear transformations of Euclidean space, we only have to understand matrices and their properties.

Example 4.13.3. First, we had the zero matrix: all coefficients are zero.

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0x + 0y + 0z \\ 0x + 0y + 0z \\ 0x + 0y + 0z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The zero matrix corresponds to the transformation that sends all vectors to the origin.

Example 4.13.4. We also had the identity matrix: ones one the diagonal and zeros elsewhere.

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

The identity matrix corresponds to the transformation which doesn't change anything. Appropriately, we called this the identity transformation.

Example 4.13.5. Diagonal matrices only have non-zero entires on the diagonal.

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + 0y + 0z \\ 0x + by + 0z \\ 0x + 0y + cz \end{pmatrix} = \begin{pmatrix} ax \\ by \\ cz \end{pmatrix}$$

This is a dialation: the x direction is stretched by the factor a, the y direction by the factor b and the z direction by the factor c.

4.14 Matrices in $\mathbb{R}^2$

We defined five types of basic transformations in $\mathbb{R}^2$. Let's give them explicit description in terms of matrices.

- The first class was rotations. To be linear, they needed to fix the origin, so they needed to be rotations about the origin. Rotation by θ (counter-clockwise, as per the standard convention) is given by the following matrix.

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

  In particular, rotations by $\pi/2$, π and $3\pi/2$ radians are given by the following matrices, respectively.

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

- The second class was reflections. To preserve the origin, they needed to be reflections over a line through the origin. In general, take $\begin{pmatrix} a \\ b \end{pmatrix}$ to be a non-zero *unit* vector in $\mathbb{R}^2$. Reflection over this line in the direction $\begin{pmatrix} a \\ b \end{pmatrix}$ is given by the following matrix.

  $$\begin{pmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{pmatrix}$$

  In particular, reflections over the x-axis, the y-axis, the line $y = x$ and the line $y = -x$ are given by the following matrices, respectively.

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

- The next set of matrices are skews. We're not going to give a general form, but consider the following two matrices.

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

  The first matrix is a skew by one in the positive y direction, and the second is a skew by one in the positive x direction. Both change squares into parallelograms. We can generalize this slighty: the following matrices are skews in y and x by the factor a, respectively.

  $$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

- Dialations were the next class. To dialate by a in the x direction and by b in the y direction, we use a diagonal matrix.

  $$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

- Finally, we have projections. To project onto the line in the direction $\begin{pmatrix} a \\ b \end{pmatrix}$, we have the following matrix.

  $$\frac{1}{a^2 + b^2} \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}$$

  In particular, projections onto the x axis, the y axis, the line $y = x$ and the line $y = -x$ are given by the following matrices, respectively.

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

  Finally, projection to the origin is given by the zero matrix.

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

4.15 Determinants

Linear transformations are defined by their symmetries: they preserve linear subspaces and linear operations. We use associated matrices to understand these transformations. Matrices are an algebraic description, so that questions of transformation can be turned into algebraic problems. Now that we have the algebraic description of transformations as matrices, we want to investigate what else the matrix can tell us about the transformation. In this section, we are asking two questions in particular.

First, what does the transformation do to the size of objects? I choose the vague word 'size' intentionally because we work in various dimensions. In one dimension, size is length. In two dimensions, size is area. In three dimensions, size is volume. And in higher dimensions, we have some extended notion of volume that fits that dimension; we can call this hyper-volume. It is important to note that size depends on the ambient dimension. A square in $\mathbb{R}^2$ has some area, some non-zero size. But if we have a flat square somewhere in $\mathbb{R}^3$, it has no volume, therefore no substantial size in that space.

Second, what does the transformation do to the orientation of an object? Orientating is slightly trickier to understand than size, so we will define it for each dimension. In $\mathbb{R}$, orientation is direction: moving in the positive or negative direction along the numberline. There are only two directions of movement, so two orientations. If we have a transformation of $\mathbb{R}$, we can ask if it changes or preserves these directions. In $\mathbb{R}^2$, instead of moving in line, we think of moving in closed loops or paths. These paths can be clockwise or counter-clockwise. Then we can ask if a transformation changes clockwise loops into other clockwise loops or into counter-clockwise loops. This gives a notion of orientation in $\mathbb{R}^2$. In $\mathbb{R}^3$, orientation relates to the relative position of positive directions. The axis system in $\mathbb{R}^3$ is, conventionally, given by a right-hand-rule. It we know the x and y directions, the right-hand-rule indicates the positive z direction. Then we can ask where these three directions go under a transformation and if a right-hand-rule still applies. If it does, we preserve the orientation. If it doesn't, and a left-hand-rule would work instead, the transformation reverses orientation. In higher dimension, there are other extentions of the notion of orientation. In each case, the question is binary: a transformation either preserves or reverses orientation..

Definition 4.15.1. Let M be a square $n \times n$ matrix. The *determinant* of M is a real number, written $\det M$, with two properties. Its absolute value $|\det M|$ measures the effect that M has on size (length, area, volume, hyper-volume). Its sign (positive or negative) measures the effect on orientation; if the sign is positive, orientation is preserved, and if the sign is negative, orientation is reversed.

That definition is all well and good, but we need to show that such a thing can be constructed. The next section gives an algorithm for building determinants.

4.16 Calculation of Determinants

The determinant of a (square) matrix is constructed recursively, starting with 2×2 matrices. We'll often use the notation of replacing the curved braces with straight lines to indicate determinants.

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

Example 4.16.1. For the matrix $\begin{pmatrix} 1 & -3 \\ 2 & -1 \end{pmatrix}$, the determinant is $(1)(-1) - (2)(-3) = -1 + 6 = 5$. Therefore, we know that this matrix multiplies all areas by a factor of 5 and preserves orientation.

Example 4.16.2. Recall that rotations in $\mathbb{R}^2$ all have a standard form.

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

The determinant of this expression is $(\cos\theta)(\cos\theta) - (-\sin\theta)(\sin\theta) = \cos^2\theta + \sin^2\theta = 1$. So all rotations preserve area and preserve orientation. This makes sense: spinning a shape around the origin does not change or area, nor does it change clockwise paths around an object into counter-clockwise paths.

Example 4.16.3. Recall we also had a standard form for reflections (we assumed that $\begin{pmatrix} a \\ b \end{pmatrix}$ is unit vector when we define reflections, so $a^2 + b^2 = 1$).

$$\begin{pmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{pmatrix}$$

The determinant here is $((a^2-b^2)(b^2-a^2)-(2ab)(2ab)) = (-a^4-b^4+2a^2b^2-4a^2b^2) = (-a^4-b^4-2a^2) = -(a^2+b^2)^2$ which simplifies to -1. This also makes sense, since area is unchanged but orientation is reversed. A reflection doesn't change area, but we expect that clockwise paths, when flipped, become counter-clockwise paths.

Definition 4.16.4. The algorithm for calculating determinants in general is called *Co-Factor Expansion*. It is a recursive algorithm that reduces determinant calculation to determinants of smaller square matrices, eventually down to 2×2 matrices.

Co-factor expansion procceds in this way: We choose *any* column or row of the matrix. We take the coefficients from that column or row. For each of the coefficients, we multiply that coefficient by the detminant of the matrix formed by removing both the columnn and row containig that coefficient. Then we add up the determinants of these small matrices multiplied by matching coefficients, with a pattern of $\pm$ signs. That pattern of $\pm$ signs is a checkerboard pattern, as in this 5×5 matrix:

$$\begin{pmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \end{pmatrix}$$

That's a hard algorithm to intuit from the formal description. Let's do a number of examples.

Example 4.16.5. Here is a 3×3 example where we do co-factor expansion on the first row.

$$\begin{vmatrix} 5 & -2 & 0 \\ -3 & 3 & -2 \\ 1 & -5 & 3 \end{vmatrix} = (+1)5 \begin{vmatrix} 3 & -2 \\ -5 & 3 \end{vmatrix} + (-1)(-2) \begin{vmatrix} -3 & -2 \\ 1 & 3 \end{vmatrix} + (+1)(0) \begin{vmatrix} -3 & 3 \\ 1 & -5 \end{vmatrix}$$

$$= 5((3)(3) - (-2)(-5)) + 2((-3)(3) - (-2)(1)) + (0)(-3)(-5) - (3)(1)$$

$$= 5(-1) + 2(-7) = -19$$

Example 4.16.6. In this example, we do cofactor expansions along the second column.

$$\begin{vmatrix} 4 & 0 & -8 \\ -1 & -2 & 3 \\ 5 & 0 & 0 \end{vmatrix} = 0(-1)\begin{vmatrix} -1 & 3 \\ 5 & 0 \end{vmatrix} + (+1)(-2)\begin{vmatrix} 4 & -8 \\ 5 & 0 \end{vmatrix} + (0)(-1)\begin{vmatrix} 4 & -8 \\ -1 & 3 \end{vmatrix}$$
$$= 0((-1)(-) - (3)(5)) - 2((4)(0) - (5)(-8)) + 0((4)(3) - (-8)(-1))$$
$$= 0 - 2(40) + 0 = -80$$

Example 4.16.7. Here is a 4×4 example, where we have to do cofactor expansion twice. The first expansion gives us four 3×3 matrices. We do cofactor expansion on each to give the 2×2 matrices that we know how to calculate. We start with cofactor expansion along the first column.

$$\begin{vmatrix} 4 & -1 & -1 & 6 \\ 0 & -3 & 3 & -2 \\ 0 & 1 & 1 & -2 \\ -2 & 0 & 0 & 3 \end{vmatrix}$$
$$= +(+1)(4)\begin{vmatrix} -3 & 3 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 3 \end{vmatrix} + (-1)(0)\begin{vmatrix} -1 & -1 & 6 \\ 1 & 1 & -2 \\ 0 & 0 & 3 \end{vmatrix}$$
$$+ (+1)(0)\begin{vmatrix} -1 & -1 & 6 \\ -3 & 3 & -2 \\ 0 & 0 & 3 \end{vmatrix} + (-1)(-2)\begin{vmatrix} -1 & -1 & 6 \\ -3 & 3 & -2 \\ 1 & 1 & -2 \end{vmatrix}$$
$$= +4\begin{vmatrix} -3 & 3 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 3 \end{vmatrix} - 0 + 0 + 2\begin{vmatrix} -1 & -1 & 6 \\ -3 & 3 & -2 \\ 1 & 1 & -2 \end{vmatrix}$$

Then we do cofactor expansion again on both of the 3×3 matrices. We use the third row of the first matrix, and the first column of the second matrix.

$$= 4\left[(+1)(0)\begin{vmatrix} 3 & -2 \\ 1 & -2 \end{vmatrix} + (-1)(0)\begin{vmatrix} -3 & -2 \\ 1 & -2 \end{vmatrix} + (+1)(3)\begin{vmatrix} -3 & 3 \\ 1 & 1 \end{vmatrix}\right]$$
$$+ 2\left[(+1)(-1)\begin{vmatrix} 3 & -2 \\ 1 & -2 \end{vmatrix} + (-1)(-3)\begin{vmatrix} -1 & 6 \\ 1 & -2 \end{vmatrix} + (+1)(1)\begin{vmatrix} -1 & 6 \\ 3 & -2 \end{vmatrix}\right]$$
$$= 4\left[(0)((3)(-2) - (1)(-2)) + (0)((-3)(-2) - (1)(-2)) + (3)((-3)(1) - (3)(1))\right]$$
$$+ 2\left[(-1)((3)(-2) - (-2)(1)) + (3)((-1)(-2) - (3)(6)) + (1)((-1)(-2) - (3)(6))\right]$$
$$= 4\left[(3)(-3-3)\right] + 2\left[(-1)(-6+2) + (3)(2-6) + (2-18))\right]$$
$$= -72 + 8 - 24 - 32 = -120$$

You may have noticed that we always choose a row or column with a maximum number of zeros for co-factor expansion. This is a useful technique, since it removes extra calculation.

If we did cofactor expansion of an arbitrary 3×3 matrix, we would get a direct expression for the determinant.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei - ahf - bdi + bfg + cdh - ceg$$

We could use this formula for 3×3 matrices, if we wished. We could do the same for larger matrices, but it starts to get quite complicated. The computational complexity of the recursive determinant algorithm grows very quickly. For a 3×3 matrix, we had 3 recursions to 2×2 matrices, each with 2 multiplication terms in the determinant, giving 6 terms. Each term was the multiple of three coefficients, giving 12 total multiplications. For a 4×4, the recursions gives 24 terms, each with 4 coefficients, so $24 \cdot 3 = 72$ multiplications. For a 5×5 matrix, we recurse five times to a 4×4 matrix, for 120 terms each with 5 coefficients, which is $120 \cdot 4 = 600$ multiplications. The pattern continues: there are $n!$ terms in the determinant of a $n \times n$ matrix, each with n coefficient, for $n!(n-1)$ multiplications. This is computationally terrifying, making determinants of large matrices computationally very difficult.

Chapter 5

Non-Linear Coordinate Systems

One way of thinking of transformations of $\mathbb{R}^n$ is as changes of coordinates. In $\mathbb{R}^2$, with standard coordinates x, y, we can think of a linear transformation, say $u = x + v$ and $v = x - v$, as a new system of coordinates. We can describe points, loci, and any other objects in terms of u and v just as well as in terms of x and y. The transformation tells us how to go between the two sets of coordinates.

All the linear transformations mentioned in the previous section (at least those which preserve dimension) are changes of coordinates. In this section, we look at changing to new coordinates in stranger ways; in particular, we look as non-linear transformations.

5.1 Polar Coordinates

In $\mathbb{R}^2$, the most common non-linear coordinate system is polar coordinates. Polar coordinates describe $\mathbb{R}^2$ in terms of circles and rays instead of the conventional lines of Cartesian coordinates. The system has two parameters (coordinates): r and θ. r, the radius, is the distance of a point from the origin. We have $r \in [0, \infty)$, since that distance can't be negative. If we draw a ray from the origin to a point, θ is the angle between that line and the x axis, with the convention that $\theta \in [0, 2\pi)$.

We would like to be able to move between Cartesian and polar coordinates. To that end, we need to describe x and y in terms of r and θ, and vice-versa. The relationships are just trigonometry.

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

The reverse direction is also trigonometry.

$$r = \sqrt{x^2 + y^2} \qquad\qquad \tan\theta = \frac{y}{x} \implies \theta = \arctan\frac{y}{x}$$

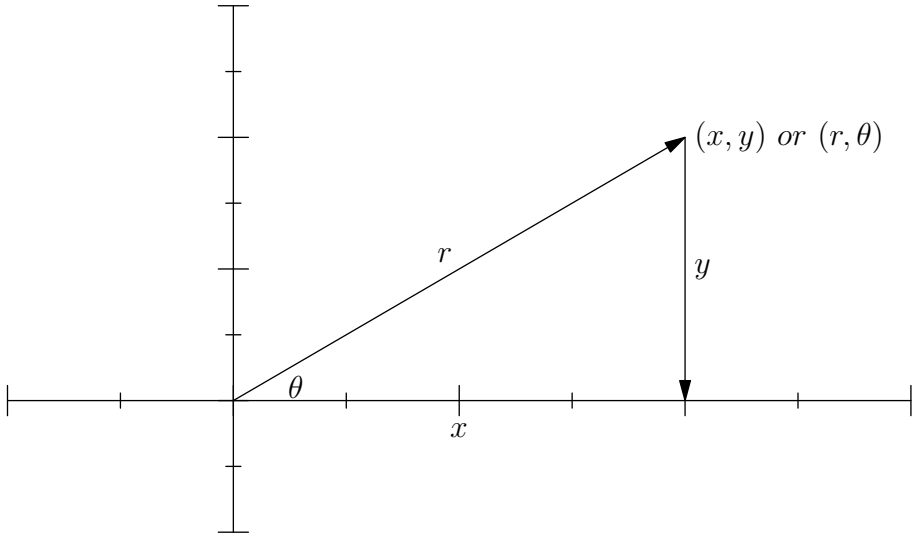

Figure 5.1: Polar Coordinates

If x is zero, then $\tan\theta$ is undefined, but θ will be $\pi/2$ or $3\pi/2$, depending on whether y is positive or negative. If x and y are both zero, at the origin, the angle is not defined at all.

Loci in $\mathbb{R}^2$ are equations in x and y. The simplest such loci were $x = c$, which is a vertical line, and $y = c$, which is a horizontal line. We get loci in $\mathbb{R}^2$ in terms of r and θ as well. $r = c$ is a circle: the shape with any angle and a fixed radius. $\theta = c$ is a ray: the shape with a fixed angle and any radius. Figures 5.2 thought 5.5 show some exammples of loci in polar coordinates.

To translate between loci in polar coordinates and loci in Cartesian coordinates is simply a matter of replacement. The line $x = 4$ in Cartesian coordinates becomes $r\cos\theta = 4$ or $\cos\theta = \frac{r}{4}$. The circle $x^2 + y^2 = 1$ becomes $r^2\cos^2\theta + r^2\sin^2\theta = 1$, which is simply $r^2 = 1 \implies r = 1$. The polar locus $r = \theta$ becomes $\sqrt{x^2 + y^2} = \arctan\frac{y}{x}$.

5.2 Spherical and Cylindrical Coordinates

In $\mathbb{R}^3$, there two are similar coordinate systems. Cylindrical coordinates use polar coordinates in the xy plane and leave the z coordinate unchanaged. The transformations are again given by trigonometry.

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$z = z$$

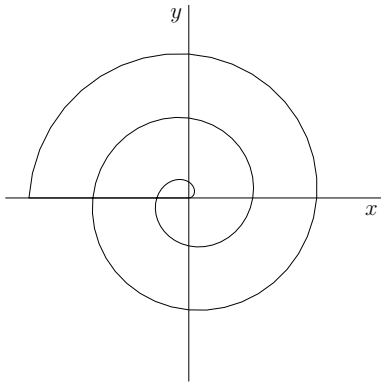

Figure 5.2: $r = \theta$

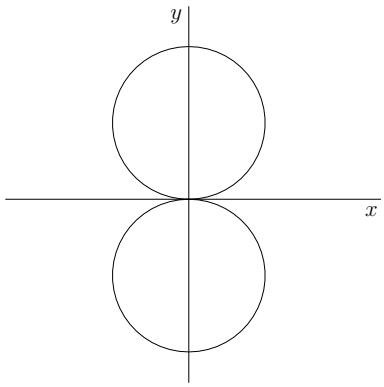

Figure 5.3: $r = |\sin\theta|$

We can invert the transformation.

$$r = \sqrt{x^2 + y^2}$$
$$\theta = \arctan\left(\frac{y}{x}\right)$$
$$z = z$$

These are called cylindrical coordinates since the equation $r = c$ gives rise to a infinitely tall cylinder. $r = c$ in the xy plane is a circle, as before. The z coordinate is left free, so the circle can be located at any z value. That infinitely tall stack of circles is a cylinder.

Spherical coordinates uses a sphere the same way that polar coordinates uses a circle. There is a radius term r, which measure the distance from a point to the origin in $\mathbb{R}^3$. r determines the size of a sphere on which the point is located. After determining a sphere, to find a specific point on the sphere, we use a system which is similar to the system of longitude and latitude on the surface of the earth. θ is the same as longitude, but with $\theta \in [0, 2\pi)$ instead of counting positively in both east and west directions.

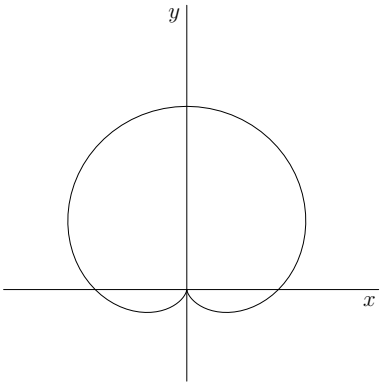

Figure 5.4: $r = 1 + \sin\theta$

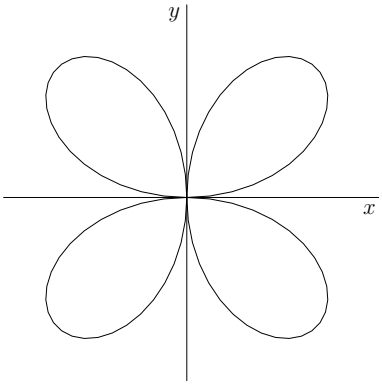

Figure 5.5: $r = 3\sin 2\theta$

The 0 line of longitude is the line that passes through the positive x axis. ϕ is co-latitude instead of latitude: it starts at 0 at the top of the sphere and counts down to π radians at the bottom of the sphere. The transformations involve some tricky trigonometry in $\mathbb{R}^3$.

$$x = r\sin\phi\cos\theta$$
$$y = r\sin\phi\sin\theta$$
$$z = r\cos\phi$$

We can invert the transformation.

$$r = \sqrt{x^2 + y^2 + z^2}$$
$$\theta = \arctan\left(\frac{y}{x}\right)$$
$$\phi = \arctan\left(\frac{\sqrt{x^2 + y^2}}{z}\right)$$

The equation $r = c$ in spherical coordinates gives a sphere.

Chapter 6

Parametric Curves

6.1 Introduction

The major goal of this course is the extension of calculus to functions with multiple variable inputs and/or outputs. In Chapter 4, we introduced linear transformation, which were function $\mathbb{R}^n \to \mathbb{R}^m$, but we would like to also investigate non-linear functions. Parametric curves are the first foray; we allow a single-input function to have a multi-variable non-linear output. We interpreted parametric curves as vectors or positions in some euclidean space; as such, parametric curves are used to talk about motion through space. The calculus of parametric curves is a way to understand the physics of such motion, covering both linear and angular velocity and acceleration in a nice, holistic approach. When considering parametric curves, we should imagine the movement of point-like objects through space under the influence of various forces. Projectiles with gravity and air friction is one imporant example; the motions of planets, moons and satellite under gravity is another.

Considering the motion of stellar objects around a large gravity source (such as planets, asteroids and comets around the sun), we will use the caluculus of parametric curves to derive Kepler's laws of planetary motion from the basic assumptions of Newtonian mechanics. Kepler's laws predate Newton, but they were simply observed, not derived. The fact that Newton's physics, with multivariable calculus, can recover these observations from first principles of motion and gravity is a major accomplishment of that theory. In order to cover Kepler's laws in full, these notes also include descriptions of conics as parametric curves.

6.2 Vector Valued Functions

Definition 6.2.1. Let A be a subset of $\mathbb{R}$. A *vector valued function* is a function $f : A \to \mathbb{R}^n$ for $n \geq 2$. It has a single-variable real input, but outputs a vector in some higher dimensional space. A vector valued function can be written as a vector of individual functions.

$$f(t) = (f_1(t), f_2(t), \ldots, f_n(t))$$

The function f_i are called the *component functions* of the vector-valued function. In $\mathbb{R}^2$, the components will often be written $(x(t), y(t))$, and similarly in $\mathbb{R}^3$, $(x(t), y(t), z(t))$.

The single-variable component functions allow us to quickly extend many notions from single-variable calculus to vector-valued functions.

Definition 6.2.2. The *limit* of a vector valued function is simply the limit of each component.

$$\lim_{t\to a} f(t) = \left(\lim_{t\to a} f_1(t), \lim_{t\to a} f_2(t), \ldots, \lim_{t\to a} f_2(t)\right)$$

Recall that a single variable function $f(t)$ is continuous at $a \in \mathbb{R}$ if

$$\lim_{x\to a} f(x) = f(a).$$

Definition 6.2.3. A vector-valued function is *continuous* if and only if each component function is continuous.

Definition 6.2.4. The *derivative* of a vector valued function $f : A \to \mathbb{R}^n$ is calculated in each component.

$$\frac{df}{dt} = f'(t) = \left(\frac{df_1}{dt}, \frac{df_2}{dt}, \ldots, \frac{df_n}{dt}\right)$$

We state this explicitly, for $f(x) = (x(t), y(t))$ with values in $\mathbb{R}^2$.

$$f'(t) = \frac{d}{dt}f(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right)$$

Likewise, we give the explicit form for $f(t) = (x(t), y(t), z(t))$ with values in $\mathbb{R}^3$.

$$f'(t) = \frac{d}{dt}f(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)$$

Definition 6.2.5. Similarly, the integral of f is defined componentwise.

$$\int f(t)dt = \left(\int f_1(t)dt, \int f_2(t)dt, \ldots, \int f_n(t)dt\right)$$

We state this explicitly, for $f(x) = (x(t), y(t))$ with values in $\mathbb{R}^2$.

$$\int f(t)dt = \left(\int x(t)dt, \int y(t)dt\right)$$

Likewise, we give the explicit form for $f(x) = (x(t), y(t), z(t))$ with values in $\mathbb{R}^3$.

$$\int f(t)dt = \left(\int x(t)dt, \int y(t)dt, \int z(t)dt\right)$$

The result of a derivative or integral of a vector valued function is still a vector. This may seem odd from the viewpoint of first year calculus, where derivatives and integrals measured quantifiable geometric properties, such as slopes of tangent lines and areas under curves. Since the answers here are vectors instead of scalars, we will eventually reconsider those interpretations. A major challenge in extending calculus to several variables is the need to re-adjust our interpretation and intuition concerning derivatives and integrals. The derivative is no longer the slope of a tangent line.

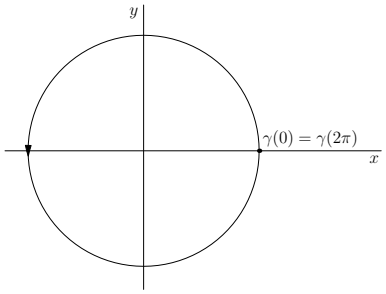

Figure 6.1: The curve $\gamma(t) = (\cos t, \sin t)$

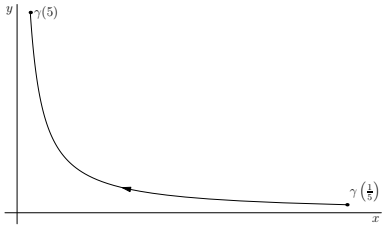

Figure 6.2: The curve $\gamma(t) = \left(\frac{1}{t}, t\right)$

6.3 Parametric Curves

Definition 6.3.1. A parametric curve in $\mathbb{R}^n$ is a continuous function $\gamma : [a, b] \to \mathbb{R}^n$, that is, a continuous vector-valued function defined on an interval. As is convention, we will typically use the symbol γ for an arbitrary parametric curve.

We identify a parametric curve with its image: that is, we think of a curve as a decription of the set of points in $\mathbb{R}^n$ which are the output of the curve. In this interpretation, we think of the curve as describing motion along this set: we start at the point $\gamma(a) \in \mathbb{R}^n$ and move along the curve, ending at $\gamma(b) \in \mathbb{R}^n$ when we get to the end of the domain.

The continuity condition is important, since a parametric curve is a connected path. We could think of vector-valued functions which jump around, but these don't really fit the notion of a curve. Continuity is also necessary for differentiability and we will assume (in order to avoid restating it multiple times) that all our parametric curves are differentiable.

For visualizing parametric curves, we graph only the output or image of the curve. There is never a t axis in any of these graphs; instead, the variable t is the parameter of movement along the curve.

Example 6.3.2. The curve $\gamma(t) = (\cos t, \sin t)$, for $t \in [0, 2\pi]$ traces out a circle, as in Figure 6.1.

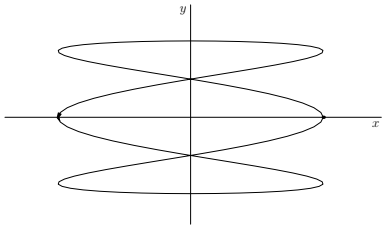

Figure 6.3: The curve $\gamma(t) = (\cos 2t, \sin t)$

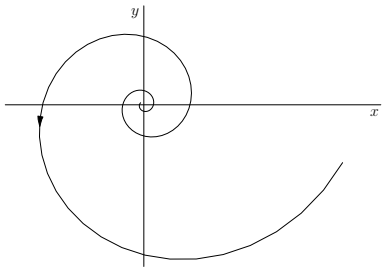

Figure 6.4: The logarithmic spiral

Notice that we defined this curve on the domain $[0, 2\pi]$. If we extend this domain, the curve just starts to retrace over the circle. It's good to observe that parametric curves can self-intersect and trace over themselves many times.

Example 6.3.3. The curve $\gamma(t) = \left(\frac{1}{t}, t\right)$ on the domain $t \in \left[\frac{1}{5}, 5\right]$ traces part of the graph of $f(x) = \frac{1}{x}$, as in Figure 6.2.

Notice that this parametric curve is the graph of a function, specifically the function $f(x) = \frac{1}{x}$ between $x = \frac{1}{5}$ and $x = 5$. Parametric curves where one of the components is just t can be interpreted as graphs of functions.

Example 6.3.4. The curve $\gamma(t) = (\cos 2t, \sin t)$ on the domain $t \in [0, 2\pi]$ oscilates faster in the x direction than in the y direction, as in Figure 6.3.

Example 6.3.5. A spiral in $\mathbb{R}^2$ is a parametric curve of the form $\gamma(t) = (f(t)\cos t, f(t)\sin t)$ where $f(t)$ is a monotonic function. It resembles the circle, but the radius is either increasing or decreasing as the curve traces around the circle. The curve $\gamma(t) = (2e^{\frac{t}{4}}\cos t, 2e^{\frac{t}{4}}\sin t)$ is a logarithmic spiral, as in Figure 6.4.

Example 6.3.6. The curve $\gamma(t) = (t\cos t, t\sin t)$ is the archimedian spiral, as in Figure 6.5.

Example 6.3.7. The curve $\gamma(t) = (t\cos t, t\sin t, t)$ on $[0, 20]$ is a spiral in $\mathbb{R}^3$, as in Figure 6.6.

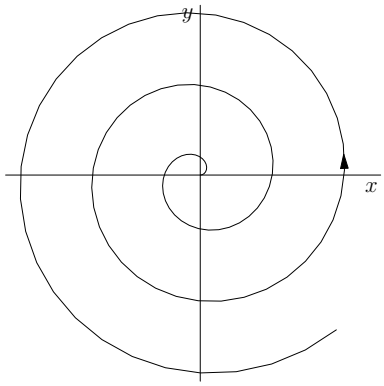

Figure 6.5: The archimedian spiral

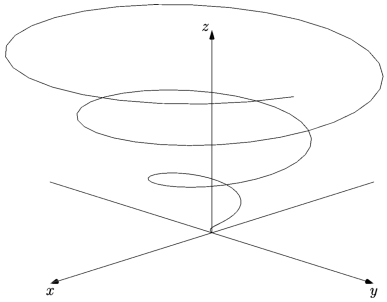

Figure 6.6: The curve $\gamma(t) = (t\cos t, t\sin t, t)$

6.3.1 Parametric Curves in Polar Coordinates

Once we are comfortable with changing coordinate systems, we can use any coordinate system we wish to describe parametric curves. For example, polar coordinates in $\mathbb{R}^2$ are well suited to describing any kind of object that is circular in some sense. Curves in polar coordinates are given as $\gamma(t) = (r(t), \theta(t))$ instead of $\gamma(t) = (x(t), y(t))$. The circle in parametric coordinates is $\gamma(t) = (1, t)$; if we write the components individually, we could write $\theta(t) = t$ and $r(t) = 1$. The logarithmic spiral has components $\theta(t) = t$ and $r(t) = 2e^{\frac{t}{4}}$. The archimedian spiral has components $\theta(t) = t$ and $r(t) = t$.

6.4 Parametrization

6.4.1 Varied Parametrizations

There are many ways to describe the same shape by a parametric curve.

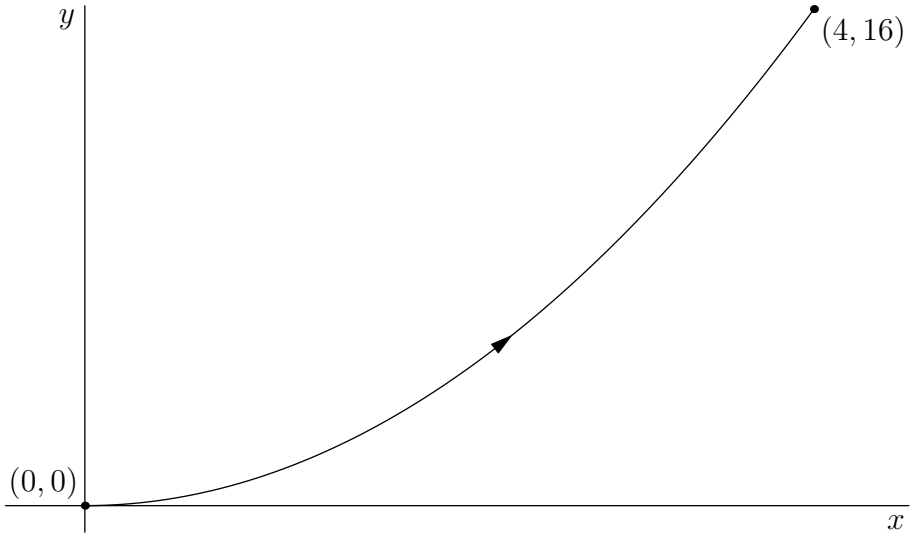

Figure 6.7: The graph of a parabola as a parametric curve

Example 6.4.1. Consider the following curves.

$$\begin{aligned}
\gamma_1(t) &= \left(t^2, t^4\right) && t \in [0, 4] \\
\gamma_2(t) &= \left(t, t^2\right) && t \in [0, 16] \\
\gamma_3(t) &= \left(\sqrt{t}, t\right) && t \in [0, 256] \\
\gamma_4(t) &= \left(5t, 25t^2\right) && t \in \left[0, \frac{16}{5}\right]
\end{aligned}$$

All four of these have exactly the same parabolic image. They all describe the same curve, shown in Figure 6.7.

6.4.2 Reparametrization

Since the same shape can have many different parametrizations, we want a process to switch between them. This process is called reparametrization.

Definition 6.4.2. Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve with coordinates $(\gamma_1(t), \ldots, \gamma_n(t))$. A *reparametrization* of γ is a monotonic increasing function $t = t(u)$ expressing the parameter t in terms of a new parameter u. We replace t by the function $t(u)$ to give a parametric curve in terms of u: $\gamma(u) = (\gamma_1(t(u)), \gamma_2(t(u)), \ldots, \gamma_n(t(u)))$.

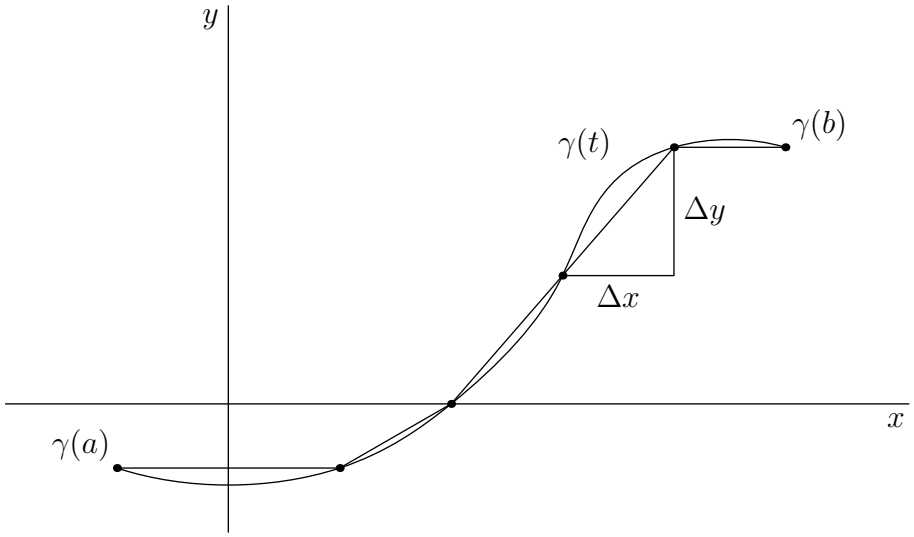

Figure 6.8: Arclength setup by approximation

Example 6.4.3. The unit circle in $\mathbb{R}^2$ is parametrized by $\gamma(t) = (\cos t, \sin t)$. If $t = 3u$ then $\gamma(u) = (\cos 3u, \sin 3u)$ is a reparametrization of the same circle. The first parametrization finishes a revolution in $t \in [0, 2\pi]$, but multiplication by 3 in the second parametrization means that a full revolution is completed in $u \in [0, 2\pi/3]$ – that is, the second parametrization moves along the circle three times as fast.

Many reparametrizations of the circle are possible.

$$\text{If } t = \frac{u}{3} \text{ then } \gamma(u) = \left(\cos\frac{u}{3}, \sin\frac{u}{3}\right)$$
$$\text{If } t = u^2 \text{ then } \gamma(u) = \left(\cos u^2, \sin u^2\right)$$
$$\text{If } t = \sqrt{u} \text{ then } \gamma(u) = \left(\cos\sqrt{u}, \sin\sqrt{u}\right)$$

Even though the shape of the curve is the same, the parametrization affects the rate movement along the curve.

6.4.3 Arc Length

Our goal in this section is to produce a formula describing the length of a parametric curve. We approximate this length by breaking up a curve γ into a series of straight lines. In order to visualize the process, we work in $\mathbb{R}^2$ for the moment, as in Figure 6.8.

For each straight line segment, if we know Δy and Δx, then the length of the segment is $\sqrt{\Delta x^2 + \Delta y^2}$. We approximate the total length of the curve by adding up the lengths of these segments.

$$L \cong \sum \sqrt{\Delta x^2 + \Delta y^2}$$

If we take the limit of this process, breaking the curve into smaller and smaller pieces, we get a Riemann sum which defines a definite integral. In the limit, the Δ terms become the infinitesimals dy and dx.

$$L = \int_a^b \sqrt{dx^2 + dy^2}$$

Both x and y depend on t, so we can change this to a integral in t.

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

The same arguments work in any $\mathbb{R}^n$, leading to the general result.

Proposition 6.4.4. *Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve. Let $\gamma_1(t), \gamma_2(t), \ldots, \gamma_n(t)$ be the components. The length of the parametric curve is calculated by this integral.*

$$L = \int_a^b \sqrt{\left(\frac{d\gamma_1}{dt}\right)^2 + \left(\frac{d\gamma_2}{dt}\right)^2 + \ldots + + \left(\frac{d\gamma_n}{dt}\right)^2}\, dt$$

Example 6.4.5. The asteroid is the parametric curve $\gamma(t) = (\cos^3 t, \sin^3 t)$ for $t \in [0, 2\pi]$, shown in Figure 6.9. We want to calculate its arclength.

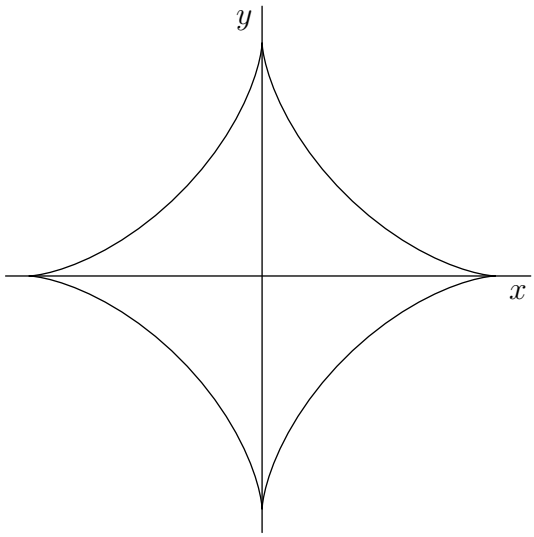

Figure 6.9: The Asteroid

$$
\begin{aligned}
L &= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \\
&= \int_0^{2\pi} \sqrt{(-3\cos^2 t(\sin t))^2 + (3\sin^2 t(\cos t))^2}dt \\
&= \int_0^{2\pi} \sqrt{9\cos^4 t \sin^2 t + 9\sin^4 t \cos^2 t}\, tdt \\
&= \int_0^{2\pi} |3\sin t \cos t| \sqrt{\sin^2 + \cos^2 t}dt \\
&= \int_0^{2\pi} |3\sin t \cos t| dt
\end{aligned}
$$

The absolute value here causes trouble. A convenient way to drop it is to notice that both $\sin t$ and $\cos t$ are positive on $[0, \pi/2]$. That range covers a quarter of the asteroid and the asteroid is symetric, so we can calculate the length of that quarter and multiply by 4.

$$L = 4\int_0^{\frac{\pi}{2}} 3\sin t \cos t dt = 6\int_0^{\frac{\pi}{2}} \sin 2t = -3\cos 2t|_0^{\frac{\pi}{2}} = 6$$

Example 6.4.6. In $\mathbb{R}^3$, consider the helix $\gamma(t) = (\cos t, \sin t, t)$. On the domain $t \in [0, 8\pi]$, the helix makes four revolutions. We calculate its arclength.

$$L = \int_0^{8\pi} \sqrt{\sin^2 t + \cos^2 t + 1}dt = \int_0^{8\pi} \sqrt{2}dt = 8\sqrt{2}\pi$$

6.4.4 Parametrization by Arclength

Our arclength calculation determins the length of the whole curve. However, we can also ask for the length of pieces of the curve.

Definition 6.4.7. Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve with components $(\gamma_1(t), \ldots, \gamma_n(t))$. The *arclength function* $s(t) : [a, b] \to [0, \infty)$ is defined by this integral.

$$s(t) = \int_a^t \sqrt{\left(\frac{d\gamma_1(u)}{du}\right)^2 + \left(\frac{d\gamma_2(u)}{du}\right)^2 + \ldots + + \left(\frac{d\gamma_n(u)}{du}\right)^2}\, du$$

The letter s is conventional notation for the arclength function.

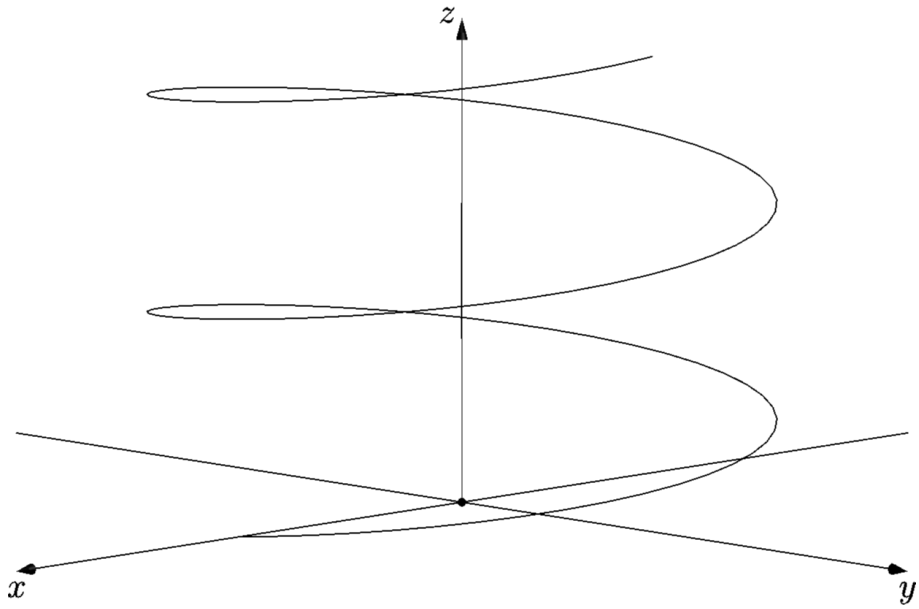

Figure 6.10: The helix $\gamma(t) = (2\cos t, 2\sin t, 4t)$

The arclength function measures the length of the curve as a function of the parameter; it is simply the distance travelled along the curve. For example, if $t \in [0, 10]$, then $s(3)$ is the distance along the curve from $\gamma(0)$ to $\gamma(3)$, $s(5)$ is the distance the curve has covered from $\gamma(0)$ to $\gamma(5)$, $s(8)$ is the distance along the curve from $\gamma(0)$ to $\gamma(8)$ and so on. Since s depends on t outside the integral, we have to choose a temporary variable u for inside the integral; in the integral, t is simply repalced with u.

We want to use the arclength function to reparametrize curve. The process has three steps.

(a) We calculate the arclength fuction $s(t)$ by integration.
(b) We invert the arclength function. Arclength is always an increasing function, so it is always invertible. We write the inverse as $t(s)$.
(c) We use the inverse of the arclength function to reparametrize by replacing t with $t(s)$.

Definition 6.4.8. Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve. Let $s(t)$ be the arclength function, with $t(s)$ its inverse. The reparametrization $\gamma(t(s))$ is called the *parametrization by arclength*. It is the unique parametrization where the parameter is the distance along the curve.

Example 6.4.9. Consider the helix $\gamma(t) = (2\cos t, 2\sin t, 4t)$ defined for $t > 0$, shown in Figure 6.10.

We calculate the arclength function.

$$\begin{aligned} s(t) &= \int_0^t \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2}du \\ &= \int_0^t \sqrt{4\sin^2 u + 4\cos^2 u + 4^2}du = \int_0^t \sqrt{20}du = t\sqrt{20} \end{aligned}$$

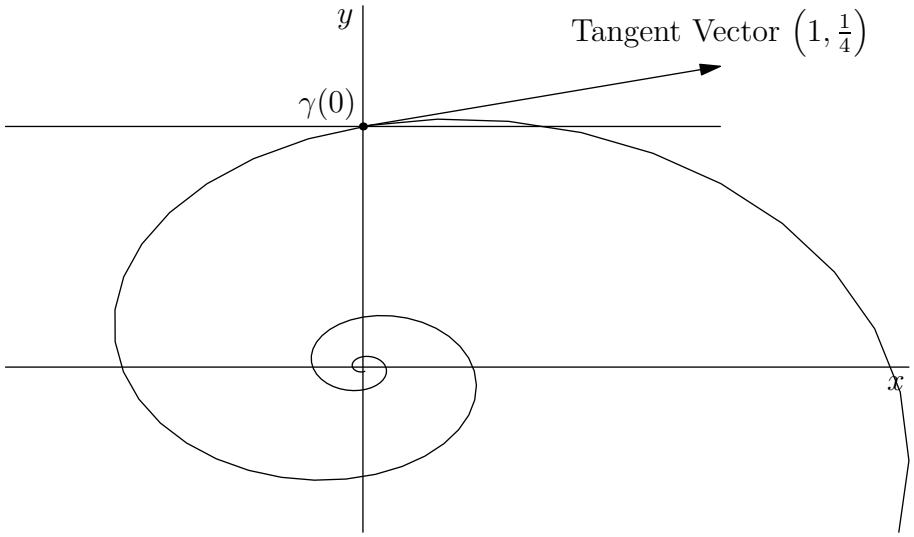

Figure 6.11: A tangent vector to the logarithmic spiral

The arclength function is $s = t\sqrt{20}$; the inverse is $t = \frac{s}{\sqrt{20}}$. We replace t with $t(s)$ in $\gamma(t)$.

$$\gamma(s) = \left(2\cos\left(\frac{s}{\sqrt{20}}\right), 2\sin\left(\frac{s}{\sqrt{20}}\right), \frac{4s}{\sqrt{20}}\right)$$

6.5 Calculus of Parametric Curves

6.5.1 Tangents

Parametric curves are vector-valued functions, so they have derivatives as defined earlier. Since we consider parametric curves as motion through $\mathbb{R}^n$, the derivative has a related interpretation.

Definition 6.5.1. Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve. Its derivative $\gamma'(t)$ is called the *tangent vector* to the parametric curve.

This tangent is no longer the slope of a tangent line from single-variable claculus. It is a vector, which we treat as a local direction vector along the curve. At $\gamma(a)$, $\gamma'(a)$ points in the local direction of movement. The notion of a derivative as a tangent direction replaces old notion of the derivative as the slope of the tangent line.

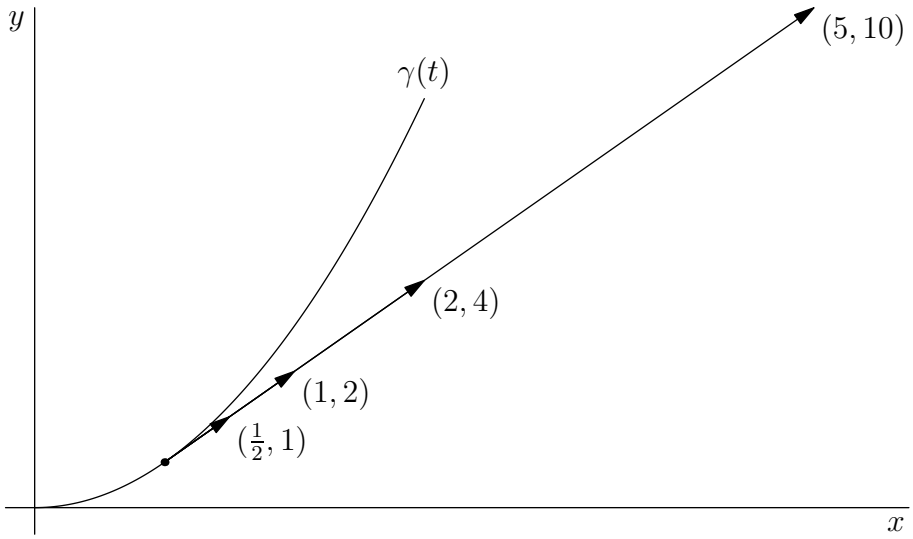

Figure 6.12: Length of tangent vectors depending on parametrization

Example 6.5.2. Consider the logarithmic spiral $\gamma(t) = (e^{\frac{t}{4}} \sin t, e^{\frac{t}{4}} \cos t)$ on the domain $t \in \mathbb{R}$. We calculate its derivative.

$$\gamma'(t) = \left(e^{\frac{t}{4}} \cos t + \frac{1}{4} e^{\frac{t}{4}} \sin t, -e^{\frac{t}{4}} \sin t + \frac{1}{4} e^{\frac{t}{4}} \cos t \right)$$

Evaluating at $t = 0$ at the point of the curve $\gamma(0) = (0, 1)$, the derivative or tangent vector to the curve is $\gamma'(0) = (1, \frac{1}{4})$.

This spiral is moving outwards; we can see how the tangent vector points in the direction of movement as the curves goes through $(0, 1)$. Figure 6.11 shows how we treat the tangent as a local direction vector: it points from $\gamma(0) = (0, 1)$, not from the origin.

Example 6.5.3. Consider the four parametrizations of the graph of the parabola introduced in Example 6.4.1 and their tangents to the point $(1, 1)$.

$$\begin{aligned}
\gamma_1'(t) &= \left(2t, 4t^3\right) && t = 1 \text{ at } (1,1) \implies \gamma_1'(1) = (2, 4) \\
\gamma_2'(t) &= (1, 2t) && t = 1 \text{ at } (1,1) \implies \gamma_2'(1) = (1, 2) \\
\gamma_3'(t) &= \left(\frac{1}{2\sqrt{t}}, 1 \right) && t = 1 \text{ at } (1,1) \implies \gamma_3'(1) = \left(\frac{1}{2}, 1 \right) \\
\gamma_4'(t) &= (5, 50t) && t = \frac{1}{5} \text{ at } (1,1) \implies \gamma_4'(\frac{1}{5}) = (5, 10)
\end{aligned}$$

All of these tangent vectors are multiples of the vector $(1, 2)$, but they have different lengths, as seen in Figure 6.12.

We said the tangent vector indicated the instantaneous direction of motion. The tangent vector also has a length, independent of the direction. The length of the tangent vector measures the speed of the curve going through that point. If the movement along the curve is very fast, the tangent vector will have a large length; if the movement along the curve is slow, the tangent vector will be shorter.

For the parametrizations of the graph of the parabolic, we can see that the direction of the tangent is independent of the parametrization, but the length of the tangent is entirely dependent on the parametrization. The parametrization which move more quickly along the curve have longer tangents.

Often it is useful to only consider the tangent direction.

Definition 6.5.4. If $\gamma'(t)$ is the tangent vector of a parametric curve, the *unit tangent vector* is the unique vector of length one in the same direction as $\gamma'(t)$. It is often written $T(t)$. To calculate the unit tangent, we simply take the tangent vector and divide by its length.

$$T(t) = \frac{\gamma'(t)}{|\gamma'(t)|}$$

Example 6.5.5. In the logarithmic spiral, the tangent vector was $\gamma'(0) = (1, \frac{1}{4})$. We calculate the length of this vector.

$$\left|\left(1, \frac{1}{4}\right)\right| = \sqrt{1 + \frac{1}{16}} = \frac{\sqrt{17}}{4}$$

Then we can calculate the unit tangent.

$$T(t) = \frac{\gamma'(t)}{|\gamma'(t)|} = \frac{4}{\sqrt{17}}\left(1, \frac{1}{4}\right) = \left(\frac{4}{\sqrt{17}}, \frac{1}{\sqrt{17}}\right)$$

Example 6.5.6. In the various parametrizations of the graph of the parabola, the tangent direction was $(1, 2)$, so the unit tangent is $\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$.

6.5.2 Tangents in Non-Linear Coordinates

Tangents are a Cartesian notion. If our curve is given in a different coordinate systems, we still want the word 'tangent' to mean the Cartesian tangent. Cartesian tangents cary the interepretation we want: the direction of motion. Therefore, to find tangent in other coordinate systems, we relate them back to Cartesian coordinates. In polar coordinates, we have $x = r\cos\theta$ and $y = r\sin\theta$. If r and θ depend on t, then we can differentiation these expressions using the product rule.

$$\begin{aligned} x' &= r'\cos\theta - r\sin\theta\theta' \\ y' &= r'\sin\theta + r\cos\theta\theta' \\ \gamma'(t) &= (r'\cos\theta - r\sin\theta\theta', r'\sin\theta + r\cos\theta\theta') \end{aligned}$$

6.5.3 Arclength and Tangents

There are some very natural connections between tangents and arclength which relate to the intepretation of parametric curves as movement through space. Let's return to the arc length formula (in $\mathbb{R}^2$, for simplicity).

$$L = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}dt$$

The tangent vector is $\gamma'(t) = (x'(t), y'(t))$, so this integrand in the arclength formula is nothing more than the length of the tangent vector. This lets us rewrite the arclength formula.

$$L = \int_a^b |\gamma'(t)|dt$$

This hopefully makes intuitive sense – the length of the tangent vector measures the speed of the curve. To get length (the distance travelled), we integrate the speed of movement.

An nice event occurs when we differentiate the arclength function $s(t)$. The fundamental theorem of calculus allows us to differentiate an integral.

$$\frac{d}{dt}s(t) = \frac{d}{dt}\int_a^t |\gamma'(u)|du = |\gamma'(t)|$$

The derivative of the arclength function is the length of the tangent vector, which represents the speed of movement along the curve. To get speed, we differentiate length.

Both of these observations generalize the relationship between distance and speed from single variable calculus. In that forum, the relationship was relatively direct: if $d(t)$ was distance, then $d'(t)$ was speed and if $v(t)$ was speed then $\int v(t)dt$ was distance. The idea here is exactly the same, but since the movement is much more complicated, going through multi-dimensional space, we need more complicated definitions and notations to access the idea.

Example 6.5.7. Recall the parametrization by arclength of the helix in $\mathbb{R}^3$ in Example 6.4.9.

$$\gamma(s) = \left(2\cos\left(\frac{s}{\sqrt{20}}\right), 2\sin\left(\frac{s}{\sqrt{20}}\right), \frac{4s}{\sqrt{20}}\right)$$

Let's calculate the speed in this parametrization by arclength. $|\gamma'(s)|$:

$$\begin{aligned}
|\gamma'(s)| &= \sqrt{x'(s)^2 + y'(s)^2 + z'(s)^2} \\
&= \sqrt{\left(\frac{-2}{\sqrt{20}}\sin\left(\frac{s}{\sqrt{20}}\right)\right)^2 + \left(\frac{2}{\sqrt{20}}\cos\left(\frac{s}{\sqrt{20}}\right)\right)^2 + \left(\frac{4}{\sqrt{20}}\right)^2} \\
&= \sqrt{\frac{4}{20}\left(\cos^2\left(\frac{5}{\sqrt{20}}\right) + \sin^2\left(\frac{5}{\sqrt{20}}\right)\right) + \frac{16}{20}} = \sqrt{\frac{4}{20} + \frac{16}{20}} = 1
\end{aligned}$$

We find that the speed is constantly one. This turns out to be generally true.

Proposition 6.5.8. *Let $\gamma(t) : [a, b] \to \mathbb{R}^n$ be a parametric curve. Our of all the possible reparametrizations of γ, the parametrization by arclength has three properties.*

- *It is the unique parametrization with constant speed of 1.*
- *It is the unique parametrization where all tangent vectors have length 1.*
- *It is the unique parametrization where the parameter measures distance along the curve.*

These properties are very useful and convenient. In the remainder of this chapter, we will make frequent use of the parametrization by arclength; it serves as our default parametrization. We can do this since the parametrization by arclength always exists and is unique. Having a special, default parameter allows us to make good definitions which will be independent of the choice of parametrization. (In practice, the parametrization by arclength can be extremely difficult to actually calculate. For theoretical results, however, this difficulty isn't relevant).

6.5.4 Curvature and Torsion

The tangent vector for a parametric curve measures the speed and direction of movement along the curve. Speed and direction are good information, but not enough to complete describe motion in multiple dimensions. In particular, the tangent vector gives the direction but doesn't say how direction is changing. This section builds up the notions of curvature and torsion, which will complete a full description of th echanging directions of movement. In this section, we exclusively work in $\mathbb{R}^3$.

In $\mathbb{R}^3$, I want to classify three types of movement.

- Straight motion: the direction is fixed and only speed varies.
- Curving motion in a plane: motion is fixed to a plane, but the direction can change and curve in the plane.
- Twisting motion: the most general motion, where in addition to curving in a plane, we can also twist away from the plane.

Throughout this section, let $\gamma(t) : [a, b] \to \mathbb{R}^3$ be a curve in three dimensions. Recall we already have the tangent $\gamma'(t)$ and the unit tangent $T(t)$. If the curve is parametrized by arc length, we will use the parameter s and write $\gamma(s)$. In this case, the tangent is already a unit vector, so $T(s) = \gamma'(s)$. We will usually start our definition using the parmetrization by arclength, to ensure they are independent of parametrizations.

Definition 6.5.9. The *curvature* of $\gamma(s)$ is a scalar $\kappa(s)$ which measure how quickly the curve is turning.

$$\kappa(s) = \left| \frac{dT(s)}{ds} \right|$$

We can think of turning as a change in the direction of the unit tangent $T(s)$. It is a scalar, not showing the new direction but just the magnitude of the change.

There are a couple things to note about this definition. First, κ is the greek letter kappa, even though it looks like our 'k'. Second, $T(s)$ is a vector, so the derivative $\frac{dT(s)}{ds}$ is a vector derivative – the derivative is taken in each component of the vector. Curvature is the scalar length of the resulting vector.

The given definition of curvatuve only works for parametrization by arclength. We would like to calculate κ in terms of an arbitrary parameter t, since arclength can be difficult to work with.

Proposition 6.5.10. *Let $\gamma(t)$ be a parametric curve with an arbitrary parametrization. The curvature $\kappa(t)$ can be calculated in this arbitrary parametrization.*

$$\kappa(t) = \frac{|T'(t)|}{|\gamma'(t)|}$$

Proof. We have the arclength function $s(t)$. The chain rule gives the derivative in a general form.

$$\frac{dT(t)}{dt} = \frac{dT(s(t))}{ds}\frac{ds}{dt}$$

We can rearrange this equation and then take the lengths of the vectors.

$$\frac{dT}{ds} = \frac{\frac{dT}{dt}}{\frac{ds}{dt}} \implies \left|\frac{dT}{ds}\right| = \left|\frac{\frac{dT}{dt}}{\frac{ds}{dt}}\right| \implies \left|\frac{dT}{ds}\right| = \frac{\left|\frac{dT}{dt}\right|}{\left|\frac{ds}{dt}\right|}$$

The left hand side of this expression is exactly κ, the curvature. The numerator on the right hand side can be written $|T'(t)|$, the change in the unit tangent for an arbitrary parametrization. The denominator $\left|\frac{ds}{dt}\right|$ is just the length of the tangent vector $|\gamma'(t)|$. □

Example 6.5.11. Consider the curve in $\mathbb{R}^3$: $\gamma(t) = (at, bt, ct)$ for some constants $a, b, c, \in \mathbb{R}$. This curve is the straight line through $(0, 0, 0)$ in the direction (a, b, c). Let's calculate the curvature of this curve. We work in steps, calculating the various pieces in order.

$$\begin{aligned}
\gamma'(t) &= (a, b, c) \\
|\gamma'(t)| &= \sqrt{a^2 + b^2 + c^2} \\
T(t) &= \frac{\gamma'(t)}{|\gamma'(t)|} = \frac{1}{\sqrt{a^2 + b^2 + c^2}}(a, b, c) \\
T'(t) &= (0, 0, 0) \\
\kappa(t) &= \frac{|T'(t)|}{|\gamma'(t)|} = \frac{0}{\sqrt{a^2 + b^2 + c^2}} = 0
\end{aligned}$$

Unsurprisingly, the straight line has zero curvature.

Example 6.5.12. Consider the unit circle in $\mathbb{R}^2$: $\gamma(t) = (\cos t, \sin t)$.

$$\begin{aligned}
\gamma'(t) &= (-\sin t, \cos t) \\
|\gamma'(t)| &= \sqrt{\sin^2 t + \cos^2 t} = 1 \\
T(t) &= \frac{\gamma'(t)}{|\gamma'(t)|} = (-\sin t, \cos t) \\
T'(t) &= (-\cos t, -\sin t) \\
|T'(t)| &= \sqrt{\cos^2 t + \sin^2 t} = 1 \\
\kappa(t) &= \frac{|T'(t)|}{|\gamma'(t)|} = \frac{1}{1} = 1
\end{aligned}$$

Again unsurprisingly, the unit circle has constant curvature of one; we would expect uniform curvate for a circle. This also give a reference for the scale of curvature: curvature of one is exactly the curvature of the unit circle.

Example 6.5.13. A similar calcluation can be done for the circle of radius a, which is $\gamma(t) = (a\cos t, a\sin t)$. If we repeat all the steps for the unit circle, we get $\kappa(t) = \frac{1}{a}$. Again, this is constant, which makes sense since the circle is equally curved at all points. However, it's interesting to note that the curvature is *inversely* proportional to the radius. A curvature value of κ is interpreted as the curvature of a circle of radius $\frac{1}{\kappa}$.

This hopefully makes sense – a circle with a very large radius doesn't locally have to curve very much. We know this very well, living on the surface of the earth: since the radius of the earth is large, we don't really see the curvature of its surface. However, for circle with very small radius, there is less distance to cover a whole revolution, so the curvature must be larger.

6.5.5 Normals

To complete the description of motion in $\mathbb{R}^3$, we define the normal and binormal vectors. I apologies, on behalf of mathematicians everywhere, for the use of the word 'normal' in this definition (and, franktly, everywhere else it's used in mathematics.) Unfortunately, it's standard and we are stuck with it.

Definition 6.5.14. Let $\gamma(t) : [a, b] \to \mathbb{R}^3$ be a parametric curve in an arbitrary parametrization. The *normal* vector to a curve is written $N(t)$ and defined by the formula

$$N(t) = \frac{T'(t)}{|T'(t)|}.$$

The *binormal* vector to a curve is written $B(t)$ and defined by the formula

$$B(t) = T(t) \times N(t).$$

The following lemma helps us understand the direction of $N(t)$ (and will be useful later in the course as well).

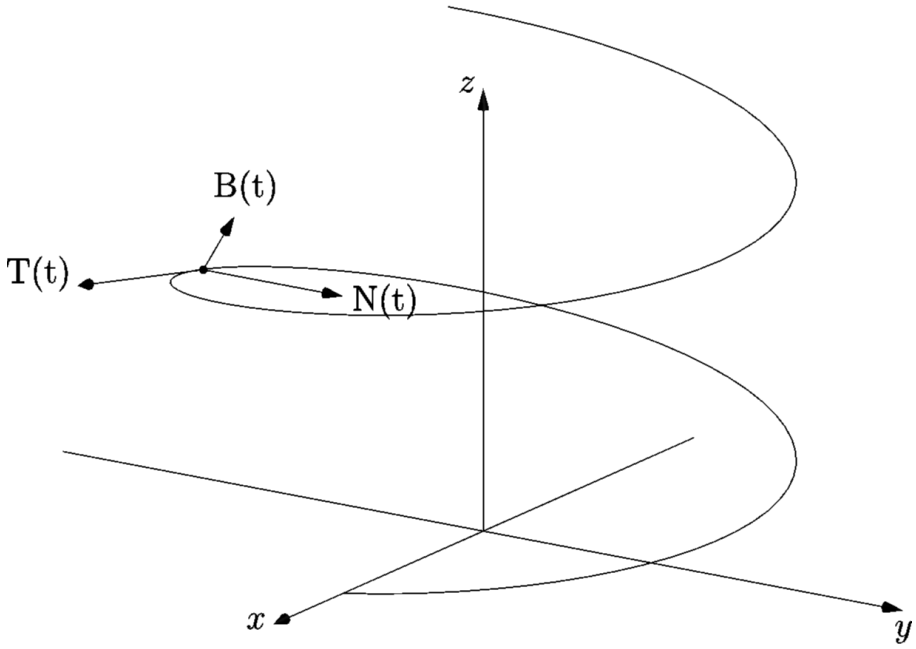

Figure 6.13: The tangent, normal and binormal to a helix in $\mathbb{R}^3$

Lemma 6.5.15. *Let $\gamma(t)$ be a differential parametric curve such that $|\gamma(t)| = 1$ for all t in the domain of the curve. (That is, $\gamma(t)$ is always a unit vector.) Then γ is perpendicular to its derivative.*

$$\gamma(t) \cdot \gamma'(t) = 0$$

Proof. Since $\gamma(t)$ is always a unit vector, the path of $\gamma(t)$ is a path on the unit sphere in $\mathbb{R}^3$. The tangent direction of this path, therefore, must be a tangent to the unit sphere. Unit vector point directly out of the unit sphere; they are perpendicular to tangent vectors of the unit sphere. Therefore, γ and γ' must be perpendicular vectors. □

Lemma 6.5.15 shows that N is perpendicular to T. The cross product of two vectors is perpedicular to both, to all of T, N and B are perpendicular to each other. In addition, T was already a unit vector and N was defined explicitly as a unit vector. The cross product of two perpendicular unit vectors is also a unit vectors, so all three are unit vectors, as seen in Figure 6.13.

Let's complete these definitions with interpretation. We've already spent some time on the unit tangent vector $T(t)$. It represents the direction of movement of the curve at that point. The associated scalar is speed, measuring how fast the object is moving at this point in the curve. The normal vector $N(t)$ represents the direction of curvature. Curvature was a scalar, measuring how curved the curve is at this point. The normal vector adds to this scalar a direction, telling you which way the curve is turning. The two vectors T and N determine a plane.

Definition 6.5.16. The plane determined by T and $N$is called the *osculating plane.*

If a parametric curve isn't twisted, the osculating plane is the plane on which the curve travels. The binormal vector $B(t)$ is perpendicular to the osculating plane. Since planes are best described by their normals, we use $B(t)$ to keep track of the osculating plane.

Our third type of movement was twisting movement, where we allowed a curve path to twist away from its plane of movement. Since the osculating plane is determined by its normal $B(t)$, twisting motion involves the change in $B(t)$.

Definition 6.5.17. Let $\gamma(s) : [0, L] \to \mathbb{R}^3$ be a parametric curve parametrized by arclength. Its *torsion* is written $\tau(s)$ and defined by the formula

$$\tau(s) = \left| \frac{dB(s)}{ds} \right|.$$

Torsion is a scalar that measures the change in B, the change in the normal of the osculating plane. It meausre the rate of twisting, the tendency of curving paths to change their plane of motion.

As with curvature, we want to calculate torsion in an arbitrary parametrization.

Proposition 6.5.18. *Let $\gamma(t) : [a, b] \to \mathbb{R}$ be a parametric curve in an arbitrary parametrization. Then its torsion is calculabed by the formula*

$$\tau(t) = \frac{-1}{|\gamma'(t)|} B'(t) \cdot N(t).$$

Proof. The arclength parameter is related to the parameter t by the arlencth function $s = s(t)$. The chain rule calculates the derivative of the binormal vector.

$$\frac{dB(t)}{dt} = \frac{dB(t(s))}{ds} \frac{ds}{dt}$$

We can rearrange this expression.

$$\frac{dB(s)}{ds} = \frac{\frac{dB(t)}{dt}}{\frac{dt}{ds}} = \frac{B'(t)}{|\gamma'(t)|}$$

Without proof, $B'(t)$ points in the direction of the negative of the normal, $-N(t)$. To get the torsion, which is the length of the vector in the previous formula, we can take the dot product with the unit vector in this direction, which is simply $-N(t)$. □

Now we have three scalars (speed, curvature, torsion) and three matching vectors (tangent, normal and binormal). Together, this information completely describes motion in $\mathbb{R}^3$. For convenience, the next list collects the definitions of all six of these mathematical objects.

$$\begin{aligned}
\gamma(t) &= (x(t), y(t), z(y)) \\
\gamma'(t) &= (x'(t), y'(t), z'(t)) \\
v(t) = |\gamma'(t)| &= \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} \\
T(t) &= \frac{\gamma'(t)}{|\gamma'(t)|} \\
\kappa(t) &= \frac{|T'(t)|}{|\gamma'(t)|} \\
N(t) &= \frac{T'(t)}{|T'(t)|} \\
B(t) &= T(t) \times N(t) \\
\tau(t) &= -\frac{B'(t)}{|\gamma'(t)|} \cdot N(t)
\end{aligned}$$

Example 6.5.19. Let r and h be two positive scalars. Consider the helix $\gamma(t) = (r\cos t, r\sin t, ht)$. In the helix, r is the radius of the circular movement of the helix, and h is the rate of linear movement along the axis of the helix. Let's calculate all the information about the motion along the helix: the speed (v), curvature (κ), torsion (τ),the tangent, the normal and finally the binormal.

$$\begin{aligned}
\gamma'(t) &= (-r\sin t, r\cos t, h) \\
v = |\gamma'(t)| &= \sqrt{r^2\sin^2 t + r^2\cos^2 t + h^2} = \sqrt{r^2 + h^2} \\
T(t) &= \frac{\gamma'(t)}{|\gamma'(t)|} = \frac{1}{\sqrt{r^2+h^2}}(-r\sin t, r\cos t, h) \\
T'(t) &= \frac{1}{\sqrt{r^2+h^2}}(-r\cos t, -r\sin t, 0) \\
|T'(t)| &= \frac{1}{\sqrt{r^2+h^2}}\sqrt{r^2\cos^2 t + r^2\sin^2 t} = \frac{r}{\sqrt{r^2+h^2}} \\
\kappa(t) &= \frac{|T'(t)|}{|\gamma'(t)|} = \frac{\frac{r}{\sqrt{r^2+h^2}}}{\sqrt{r^2+h^2}} = \frac{r}{r^2+h^2} \\
N(t) &= \frac{T'(t)}{|T'(t)|} = \frac{1}{\sqrt{r^2+h^2}} r\sqrt{r^2+h^2}(-r\cos t, -r\sin t, 0) \\
&= \frac{1}{r}(-r\cos t, -r\sin t, 0) = (-\cos t, -\sin t, 0) \\
B(t) &= T(t) \times N(t) = \frac{1}{\sqrt{r^2+h^2}}(-r\sin t, r\cos t, h) \times (-\cos t, -\sin t, 0) \\
&= \frac{1}{\sqrt{r^2+h^2}}(h\sin t, -h\cos t, r)
\end{aligned}$$

$$B'(t) = \frac{1}{\sqrt{r^2+h^2}}(h\cos t, h\sin t, 0)$$
$$\tau(t) = -\frac{B'(t)}{|\gamma'(t)|}\cdot N(t) = \frac{1}{r^2+h^2}(h\cos t, h\sin t, 0)\cdot(-\cos t, -\sin t, 0) = \frac{h}{r^2+h^2}$$

Look at the three scalars.

$$v = \sqrt{r^2+h^2} \qquad \kappa = \frac{r}{r^2+h^2} \qquad \tau = \frac{h}{r^2+h^2}$$

All three, speed, curvature and torsion, are constant here. Let's summarize what kind of shapes of curves we get for constant and zero values of the three scalars describing the motion of parametric curves.

$v=0$	$\kappa=0$	$\tau=0$	$\Longrightarrow$	no movement at all
$v=c$	$\kappa=0$	$\tau=0$	$\Longrightarrow$	straight line movement, no curving or twisting
$v=c_1$	$\kappa=c_2$	$\tau=0$	$\Longrightarrow$	tracing around a circle, constant speed and curvature, no twisting
$v=c_1$	$\kappa=c_2$	$\tau=c_3$	$\Longrightarrow$	helix, constantly moving, curving and twisting

Like the straight line and the circle, the helix is the unique curve which has constant non-zero speed, torsion and cruvature.

6.6 Acceleration and Movement in Space

Now that we understand how parametric curves describe motion in $\mathbb{R}^3$, we can do some physics. This section tries to undertstand what acceleration means in the language of parametric curves. The starting point is simple: acceleration should be the second derivative of the curve.

$$a(t) = \gamma''(t) = \frac{d}{dt}\gamma'(t)$$

This means acceleration is a vector. Let's work out its direction and magnitude, using T, N and B as reference for direction. Recall the definition of the unit tangent.

$$T(t) = \frac{\gamma'(t)}{|\gamma'(t)|}$$

We solve for $\gamma'(t)$.

$$\gamma'(t) = \underbrace{|\gamma'|}_{\text{scalar}}\cdot\underbrace{T(t)}_{\text{direction}}$$

We differentiate this, using the product rule.

$$\frac{d}{dt}\gamma'(t) = \left(\frac{d}{dt}|\gamma'(t)|\right)T(t) + |\gamma'(t)|\frac{d}{dt}T(t) \tag{6.6.1}$$

Recall the definition of the normal.

$$N(t) = \frac{T'(t)}{|T'(t)|}$$

We isolate $T'(t)$.

$$T'(t) = |T'(t)|N(t) \tag{6.6.2}$$

Finally, recall the definition of curvature.

$$\kappa(t) = \frac{|T'(t)|}{|\gamma'(t)|} \implies |T'(t)| = \kappa(t)|\gamma'(t)|$$

We use this expression to replace $|T'(t)|$ in Equation 6.6.2.

$$T'(t) = \kappa(t)|\gamma'(t)|N(t)$$

We put this expression for $T'(t)$ into the second term of Equation 6.6.1, then group and label the terms.

$$a(t) = \underbrace{\frac{d}{dt}|\gamma'(t)|}_{\text{linear acceleration}} T(t) + \underbrace{\kappa(t)|\gamma'(t)|^2}_{\text{angular acceleration}} N(t)$$

Definition 6.6.1. Let $\gamma(t) : [a, b] \to \mathbb{R}^3$ be a parametric curve. Its *linear acceleration* is the vector $\frac{d}{dt}|\gamma'(t)|T(t)$ and its *angular acceleration* is the vector $\kappa(t)|\gamma'(t)|^2N(t)$. Linear acceleration tells us how quickly the speed changes. Angular acceleration tells us how quickly the direction changes.

Example 6.6.2. For the helix, $|\gamma'(t)|$ is constant, so the acceleration is entirely angular.

$$a(t) = 0T(t) + rN(t) = rN(t)$$

This makes sense, since the linear speed doesn't every change; only the direction changes as the helix curves and twists.

6.7 Conics

We are going to need conics to talk about orbital mechanics in the next section. Let's first review the classical definitions of the four types of conics.

Definition 6.7.1. The *circle* is all points equidistant from a center point (focus). It is determined entirely by its centre point and radius.

Parameters	Radius r
Equation at $(0, 0)$	$x^2 + y^2 = r^2$
Equation at (x_0, y_0)	$(x - x_0)^2 + (y - y_0)^2 = r^2$

Definition 6.7.2. The *ellipse* is all points equidistant from *two* points (foci). For the ellipse centered at the origin, the foci are $(\pm c, 0)$ where $c^2 = a^2 - b^2$. The ellipse is determined by a centre point and two axis length, or by the two foci and a distance. (This description of the ellipse assumes $a > b$. If $a < b$, the semimajor and semiminor axes are switched, since the major axis is always larger, the foci would be on the y axis instead of the x, axis and $c^2 = b^2 - a^2$.)

Parameters	Semimajor axis a Semimnior axis b
Equation at $(0,0)$	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$
Equation at (x_0, y_0)	$\dfrac{(x-x_0)^2}{a^2} + \dfrac{(y-y_0)^2}{b^2} = 1$

Definition 6.7.3. The *parabola* is all points which are equidistant to a point (focus) and a line (directrix). For the parabola centered at the origin, the focus is $(0, p)$ and the directrix is the line $y = -p$. The focus and directrix entirely determine the parabola. (This description holds for upward opening parabolae. If the directrix is $y = p$ and the focus $(0, -p)$, then $x^2 = -4py$ is the downward facing parabola. Switching the roles of x and y gives left or right facing parabolae.)

Parameters	Focus - Directrix Distance $2p$
Equation at $(0,0)$	$x^2 = 4py$
Equation at (x_0, y_0)	$(x-x_0)^2 = 4p(y-y_0)$

Definition 6.7.4. The *hyperbola* is all points where the *difference* of the distances to the two foci is constant. The foci, similar to the ellipse, are the points $(\pm c, 0)$ with $c^2 = a^2 + b^2$. The hyperbola has vertices $(\pm a, 0)$, which are the points closest to the origin. In addition, it has asymptotes, which are the lines $y = \pm \frac{b}{a} x$. The hyperbola is determined by its two foci and a distance. (This description is for hyperbolae which open along the x axis. Switching x and y gives hyperbolae which open along the y axis.)

Parameters	Axes $a,\ b$
Equation at $(0,0)$	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$
Equation at (x_0, y_0)	$\dfrac{(x-x_0)^2}{a^2} - \dfrac{(y-y_0)^2}{b^2} = 1$

6.7.1 Eccentricity

The parabola is defined as all points equidistant from a focus and a directrix. All other conics have foci, but the parabola is the only with a directrix. However, there is another method of describing all conics in terms of focus and directrix. To get to that description, consider first the parabola in more

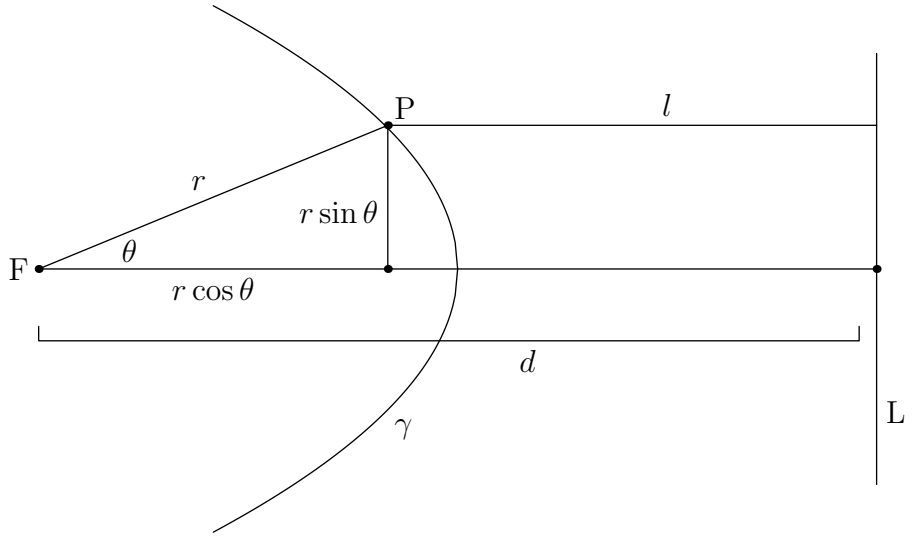

Figure 6.14: Conic definition with directrix L and focus F

detail in the following diagram. In Figure 6.14, we've labeled the conic itself as γ, the focus F and the directrix L and the distance between the focus and the directrix, labeled d.

We've also labeled an angle θ and two distances r and l in this diagram. The point P is, by definition, on the parabola because the distances from P to the focus and to the directrix are equal. In Figure 6.14, $r = l$.

We can generalize the construction by allowing r and l to be non-equal, but insisting that their ration is constant. That is, given a focus F and a directrix L, we consider all points P, such $\frac{r}{l}$ is a constant. $\frac{r}{l} = 1$ will recover the parabola, but other choices of the constant will result, amazingly, in other conics.

Definition 6.7.5. The *eccentricity* of a conic defined via a focus and a directix is the ratio $\frac{r}{l}$ of the distance to the focus over the distance to the directrix.

$$e = \frac{r}{l}$$

We would like to get a description of conics as a parametric curves using the focus, directrix and eccentricity. We will use polar coordinates for this, assuming that F is at the origin. In Figure 6.14, putting F at the origin means that r and θ, as labelled, are exactly polar coordinates.

Proposition 6.7.6. *Let γ be a conic described by a focus at the origin, a directrix $x = d$ and an eccentricity e. Then γ is described by a polar locus,*

$$r = \frac{ed}{1 + e\cos\theta},$$

or by a polar parametric curve using θ as the parameter,

$$\gamma(\theta) = (r(\theta), \theta) = \left(\frac{ed}{1 + e\cos\theta}, \theta\right)$$

Proof. From the diagram and some trigonometry, we can see that $l = d - r\cos\theta$. Eccentricity is the ratio of r to l, so eccentricity has this form:

$$e = \frac{r}{d - r\cos\theta}$$

If we solve for r in this equaiton we get the desired polar locus.

$$r = \frac{ed}{1 + e\cos\theta} \tag{6.7.1}$$

Then we simply choose θ as the parameter for the polar parametric description. □

Proposition 6.7.7. *All four conics can be defined in terms of a focus F, directrix L, and eccentricity e. A conic is thus redefined to be all points where the ratio of the distance to the focus over the distance to the directrix is constant and equal to the eccentricity. The hyperbola has eccentricity $e > 1$, the parabola has eccentricity $e = 1$, the ellipse has eccentricity $e < 1$. The circle is the special limit case $e = 0$, though to realize the shape we have to allow the distance between F and L to approach infinity.*

Proof. We will prove all cases except the limit case of the circle. Our approach is to start with the polar locus and try to recover the standard equations of the conics from the start of this section.

We start with the polar locus and change back to Cartesian coordinates, using the change of coordinates equations: $r = \sqrt{x^2 + y^2}$ and $r\cos\theta = x$.

$$\begin{aligned}
\frac{r}{d - r\cos\theta} &= e \\
\frac{\sqrt{x^2 + y^2}}{d - x} &= e \\
\sqrt{x^2 + y^2} &= ed - ex = e(d - x) \\
x^2 + y^2 &= e^2(d^2 - 2dx + x^2)
\end{aligned}$$

We rearrange this to get a polynomial expression in x.

$$(1 - e^2)x^2 + 2de^2x + y^2 = d^2e^2 \tag{6.7.2}$$

If $e \neq 1$, we divide through by $(1 - e^2)$, then proceed to complete the square in the x term.

$$\begin{aligned}
x^2 + \frac{2de^2}{1 - e^2}x + \frac{y^2}{1 - e^2} &= \frac{d^2e^2}{1 - e^2} \\
x^2 + \frac{2de^2}{1 - e^2}x + \frac{d^2e^4}{(1 - e^2)^2} - \frac{d^2e^4}{(1 - e^2)^2} + \frac{y^2}{1 - e^2} &= \frac{d^2e^2}{1 - e^2} \\
\left(x + \frac{e^2d}{1 - e^2}\right)^2 + \frac{y^2}{1 - e^2} &= \frac{d^2e^2}{1 - e^2} + \frac{d^2e^4}{(1 - e^2)^2}
\end{aligned}$$

We take the right-hand side to common denominator.

$$\left(x + \frac{e^2 d}{1-e^2}\right)^2 + \frac{y^2}{1-e^2} = \frac{d^2e^2(1-e^2) + d^2e^4}{(1-e^2)^2}$$
$$\left(x + \frac{e^2 d}{1-e^2}\right)^2 + \frac{y^2}{1-e^2} = \frac{d^2e^2}{(1-e^2)^2}$$

We divide by the right-hand side.

$$\frac{\left(x + \frac{e^2 d}{1-e^2}\right)^2}{\frac{d^2e^2}{(1-e^2)^2}} + \frac{\frac{y^2}{1-e^2}}{\frac{d^2e^2}{(1-e^2)^2}} = 1$$
$$\frac{\left(x + \frac{e^2 d}{1-e^2}\right)^2}{\frac{d^2e^2}{(1-e^2)^2}} + \frac{y^2}{\frac{d^2e^2}{1-e^2}} = 1$$

The sign in from of the y^2 term is positive if $1 - e^2 > 1$, that is, when $e < 1$. Otherwise, $e^2 - 1$ is positive. In the positive case, we can rewrite the expression as a difference of squares.

$$\frac{\left(x + \frac{e^2 d}{1-e^2}\right)^2}{\frac{d^2e^2}{(1-e^2)^2}} - \frac{y^2}{\frac{d^2e^2}{e^2-1}} = 1$$

In the case $e < 1$, the expression is the sum of positive squares, so we have an ellipse. In the case $e > 1$, the expression is the difference of positive square, so we have a hyperbola. For future convenience, we define the following constants for the $e < 1$ (ellipse) case.

$$x_0 = \frac{-e^2 d}{1-e^2} \qquad a^2 = \frac{e^2 d^2}{(1-e^2)^2} \qquad b^2 = \frac{e^2 d^2}{1-e^2} \tag{6.7.3}$$

Subsituting these cosntant into the previous equation gives an elegant ellipse equation.

$$\frac{(x-x_0)^2}{a^2} \pm \frac{y^2}{b^2} = 1$$

Lastly, in Equation 6.7.2 we excluded $e = 1$. Now assume $e = 1$, which removes the $(1 - e^2)x^2$ term from Equation 6.7.2 entirely. We continue the derivation from that point; the calculation is much easier in this case.

$$2dx + y^2 = d^2 \implies y^2 = d^2 - 2dx \implies y^2 = -2d\left(x - \frac{d}{2}\right)$$

This is the equation for a leftward-opening parabola, so $e = 1$ does recover the parabola.

For all cases except the circle, we recovered the standard equation of the conic. We found a hyperbola when $e > 1$, an ellipse when $e < 1$, and a parabola when $e = 1$. As mentioned in the proposition, the circle is a strange limit case as $e \to 0$, but $e \to 0$ also increases the distance $d \to \infty$. □

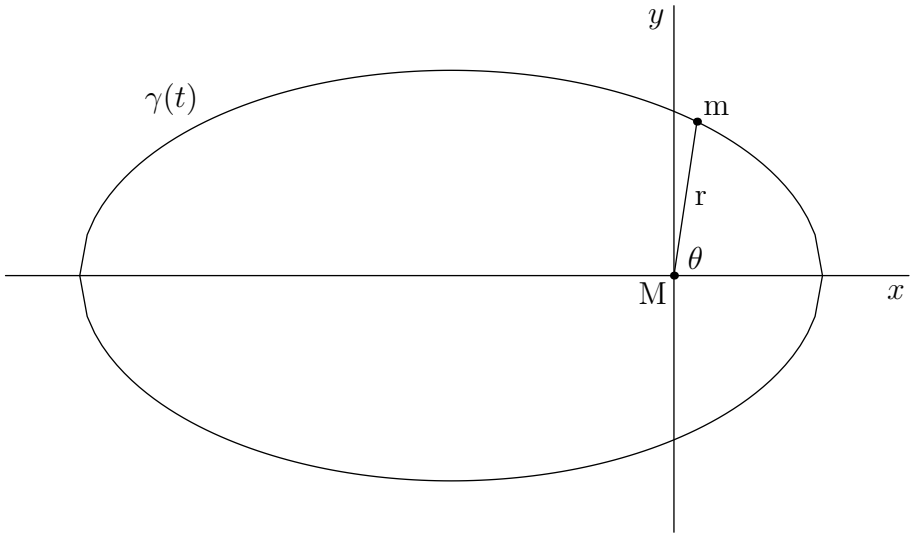

Figure 6.15: An orbital path for Kepler's laws

6.8 Kepler's Laws

Kepler's laws were originally formulated from the observations of planets in the night sky and were written down before Newton. One of the triumphs of Newtonian mechanics is the recovery of Kepler's laws. In this section, we start with Newton's gravity and derive Kepler's laws.

Kepler described three laws of planetary motion.

- Satellites in orbit around a large gravity source have elliptical orbits with the large object at one of the foci of the ellipse.
- The radius of a satellite sweeps out equal area over time.
- The period T of a satellite and the major axis a of the associated ellipse satisfy $T^2 = \alpha a^3$ for some constant α depending only upon the mass of the large object.

6.8.1 Kepler's First Law

The setup for approaching Kepler's laws is show in Figure 6.15.

- We have a large stationary object of mass M, and a small object of mass m in orbit around the larger object. We must assume that $m \ll M$ to allow the larger object to be essentially stationary.

- We place the stationary object of mass M at the origin in $\mathbb{R}^3$. We might work in the plane since the orbits are planar, but it is more useful to think of the orbit sitting in the xy plane in $\mathbb{R}^3$. This will be important, since it turns out to be useful to consider the perperdicular direction to the plane of orbit and use normals and binormals in $\mathbb{R}^3$.
- The curve $\gamma(t) = (r(t), \theta(t), z(t))$ decribes the motion of the orbiting object over time, in cylindrical coordinates (which reduce to polar coordinates in the xy plane). In the third coordinate $z(t)$, we expect that $z(t) = 0$ at all times t (which will produce a curve in the xy plane in polar coordinates), but we don't assume this.
- The curve $\gamma(t)$ is unknown and we wish to derive it from Newtonian physics. We will prove Kelper's laws using the derived curve.
- The force of gravity has magnitude
$$F = \frac{GmM}{|\gamma(t)|^2} = \frac{GmM}{r(t)^2}.$$
- The direction of the force is $-\gamma(t)$, since the force wishes to pull the object back towards the origin. Since we want just direction without changing magnitude, we use the unit vector $u(t) = \gamma(t)/|\gamma(t)|$ for the direction of $\gamma(t)$.
- Newton's first law $F = ma$ applies, so the acceleration is
$$a(t) = \frac{-GM}{|\gamma(t)|^2}u(t) = \frac{-GM}{|\gamma(t)|^3}\gamma(t).$$
- Acceleration is the second derivative of position, therefore Newton's first law becomes
$$\frac{d^2}{dt^2}\gamma(t) = \frac{-GM}{|\gamma(t)|^3}\gamma(t)$$
- We write $h(t) = \gamma(t) \times \gamma'(t)$. $h(t)$ is in the binormal direction, but is not necessarily a unit vector. This seems like a strange definition, but the vector h turns out to be very useful in the proof.

Newton's law is a multivariable differential equation; solving it directly is very difficult and well beyond the scope of this course. Our approach is indirect. We start with two seeminly random calculations. For convenience of notation, we often drop the t variable, writing γ instead of $\gamma(t)$ and likewise for other functions of t.

Lemma 6.8.1.
$$\frac{d}{dt}h(t) = 0$$

Proof.
$$\frac{d}{dt}h(t) = \frac{d}{dt}\gamma \times \gamma' = \frac{d\gamma}{dt} \times \gamma' + \gamma \times \frac{d\gamma'}{dt} = \gamma' \times \gamma' + \gamma \times a$$
The first term $\gamma' \times \gamma' = 0$ since any vector with itself in the cross product is 0. Then γ and the acceleration a are in the same direction (up to ± 1), since the force pulls back towards the origin. Vectors in the same direction likewise have 0 cross product. Therefore, we have $\frac{d}{dt}h = 0$. □

Lemma 6.8.2.

$$h = |\gamma|^2(u \times u')$$

Proof. We can write h this way.

$$h(t) = \gamma \times \gamma' = (|\gamma|u) \times \frac{d}{dt}(|\gamma|u)$$

We use the product rule on the second term.

$$\begin{aligned} &= |\gamma|u \times \left(u\frac{d}{dt}|\gamma| + |\gamma|\frac{d}{dt}u\right) = |\gamma|u \times u\frac{d}{dt}|\gamma| + |\gamma|u \times |\gamma|\frac{d}{dt}u \\ &= |\gamma|\frac{d}{dt}|\gamma|(u \times u) + |\gamma|^2 u \times u' \\ h &= |\gamma|^2(u \times u') \end{aligned}$$

□

Proposition 6.8.3. *(Kepler's First Law) The differential equation*

$$a = \frac{-GM}{|\gamma|^2}u$$

can be solved with a polar parametric curve $\gamma(t) = (r(t), \theta(r))$ which satisfies the equation

$$r(t) = \frac{ed}{1 - e\cos\theta(t)}.$$

This is exactly the same as Equation 6.7.1. By Proposition 6.7.7 such forms describe all conics, so Kepler's First Law states that orbital paths are conics.

Proof. We start with Newton's First Law as a differential equation, using the force of gravity on the right side.

$$a = \frac{-GM}{|\gamma|^2}u$$

Let's take the cross product with h. We use Lemma 6.8.2 on the right side.

$$\begin{aligned} a \times h &= \frac{-GM}{|\gamma|^2}u \times h = \frac{-GM}{|\gamma|^2}u \times \left(|\gamma|^2(u \times u')\right) \\ &= \frac{-GM}{|\gamma|^2}|\gamma|^2\left(u \times (u \times u')\right) = -GM\left(u \times (u \times u')\right) \end{aligned}$$

We use Proposition 4.6.2 to expand the triple cross product on the right side.

$$a \times h = -GM((u \cdot u')u - (u \cdot u)u'))$$

Since u is always a unit vector, we can apply Lemma 6.5.15, which says that $u \times u' = 0$. Also since u is a unit vector, $u \cdot u = |u|^2 = 1$. This deals with both of the dot products.

$$a \times h = -GM(0u - 1u') = GMu' \tag{6.8.1}$$

Then we consider the following derivative.

$$\frac{d}{dt}(\gamma' \times h) = \frac{d\gamma'}{dt} \times h + \gamma' \times \frac{dh}{dt}$$

The first term here is $a \times h$, since a is $\frac{d\gamma'}{dt}$. The second term involves the derivative of h. But Lemma 6.8.1 said that the derivative of h is zero, so this term vanishes.

$$\frac{d}{dt}(\gamma' \times h) = a \times h$$

Then let's replace $a \times h$ with the expression from Equation 6.8.1.

$$\frac{d}{dt}(\gamma' \times h) = GMu'$$

Now we've made progress with our difficult differential equation: we have a time derivative on both sides. We can simply integrate both sides to give $\gamma' \times h = GMu + c$. where c is a *vector* of constants of integration. (That vector corresponds to intial conditions of position and velocity). We take the dot product with γ.

$$\gamma \cdot (\gamma' \times h) = GM(u \cdot \gamma) + \gamma \cdot c = GM|\gamma|u \cdot u + |\gamma||c| \cos\theta = GM + |\gamma||c| \cos\theta$$

Here θ is the angle between γ and c. Since we haven't specified a starting point, we can choose coordinates such that c is in the positive x direction, which means that this θ is the usual θ of polar coordinates and $|\gamma|$ is the usual r of polar coodinates.

$$\gamma \cdot (\gamma' \times h) = GM|\gamma| + |\gamma||c| \cos\theta$$

We solve for $|\gamma|$.

$$|\gamma| = \frac{\gamma \cdot (\gamma' \times h)}{GM + |c| \cos\theta}$$

The expression $|c|/GM$ is a constant, so let's give it a label.

$$e := \frac{|c|}{GM} \tag{6.8.2}$$

Let's switch the $|\gamma|$ for r as well.

$$r = \frac{\gamma \cdot (\gamma' \times h)}{GM + |c| \cos\theta} = \frac{\gamma \cdot (\gamma' \times h)}{1 + e \cos\theta} \left(\frac{1}{GM}\right)$$

The numerator involves $\gamma \cdot (\gamma' \times h)$, which we can rearrange to $h \cdot (\gamma \times \gamma')$ using Proposition 4.6.3. But $(\gamma \times \gamma')$ is the definition of h, so this dot product is $h \cdot h = |h|^2$.

$$r = \frac{|h|^2}{1 + e\cos\theta}\left(\frac{1}{GM}\right) \tag{6.8.3}$$

$|h|^2$ is a constant, since h doesn't change according to Lemma 6.8.1. Let's make another definition.

$$d = |h|^2/|c| \tag{6.8.4}$$

We can then replace the nuemrator of Equation 6.8.3 using d.

$$r = \frac{d|c|}{1 + e\cos\theta}\left(\frac{1}{GM}\right) = \frac{d}{1 + e\cos\theta}\left(\frac{|c|}{GM}\right)$$

Replace $\frac{|c|}{GM}$ by e, since that's how we defined e in Equation 6.8.2.

$$r = \frac{ed}{1 + e\cos\theta}$$

This is the desired form. □

The eccentricity was defined as $e = |c|/GM$. Notice that the eccentricity depends on the constants of integration. That makes sense since those constants determine inital position and velocity. If $e < 1$, then the initial conditions indicate an ellipse. If $e \geq 1$ then we get a hyperbola. The difference is precisely the notion of escape velocity – these initial conditions tell us if the satellite has enough initial energy to escape or be trapped in orbit. If trapped in orbit, the orbits are elliptical. If escaping, the path is parabolic or hyperbolic. We will assume that the orbits are elliptical, i.e., that the original velocity is less than escape velocity, to prove the final two laws.

6.8.2 Kepler's Second Law

Proposition 6.8.4. *(Kepler's Second Law). The area swept out by a line between the M and the satellite is constant in time.*

Proof. Write $A(t)$ for the area swept out in this way. Our goal is to prove that $A'(t)$ is constant. We approach this by looking at the infinitesimal area dA swept out over an infinitesimal time interval dt. Such an area is a small triangle, shown in gray in Figure 6.16. It has side length r and base db, which we can assume is perpendicular to the radius. Therefore, the area dA of the infinitesimal triangle is $\frac{1}{2}rdb$. Then db, as an infinitesimal arclength, is $rd\theta$, so $dA = \frac{1}{2}r^2 d\theta$, which we can integrate (with a temporary internal variable for the integration).

$$A(t) = \int_0^t \frac{1}{2}r(w)^2 d\theta = \int_0^t \frac{1}{2}r(w)^2 \frac{d\theta}{dw} dw$$

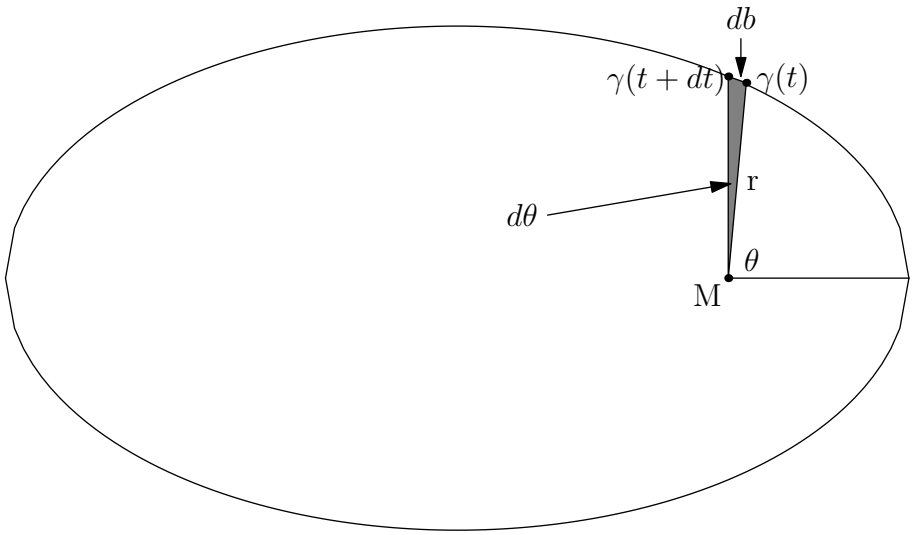

Figure 6.16: Approximate movement along the orbit over time dt.

We calculate the derivative $\frac{dA}{dt}$.

$$\frac{dA}{dt} = \frac{d}{dt}\int_0^t \frac{1}{2}r(w)^2\frac{d\theta}{dw}dw = \frac{1}{2}r(t)^2\frac{d\theta}{dt}$$

Now we return to some of the derivations of the previous section to understand this equation. Recall $r(t) = |\gamma(t)|$. In Cartesian coordinates, γ has this form.

$$\gamma(t) = (r(t)\cos\theta(t), r(t)\sin\theta(t), 0)$$

The the unit vector $u = \gamma(t)/|\gamma(t)|$ is $u(t) = (\cos\theta(t), \sin\theta(t), 0)$. We calculate its derivative.

$$\frac{du}{dt} = (-\sin\theta(t), \cos\theta(t), 0)\frac{d\theta}{dt} \tag{6.8.5}$$

We calculate $u \times u'$.

$$u \times \frac{du}{dt} = (\cos\theta(t), \sin\theta(t), 0) \times (-\sin\theta(t), \cos\theta(t), 0)\frac{d\theta}{dt} = (0, 0, 1)\frac{d\theta}{dt} \tag{6.8.6}$$

Recall Lemma 6.8.2.

$$h = |\gamma(t)|^2 u \times \frac{du}{dt}$$

We replace the cross product on the right with the same expression from Equation 6.8.6.

$$h = |\gamma(t)|^2\frac{d\theta}{dt}(0, 0, 1)$$

We take the magnitude of this vector. (We choose a direction of orbit for that $\frac{d\theta}{dt}$ is positive.)

$$|h| = |\gamma(t)|^2 \frac{d\theta}{dt} = r(t)^2 \frac{d\theta}{dt}$$

Now this looks familiar: if has the same right side as Equation 6.8.5. Therefore, it is equation to the left side of that equation. $\frac{dA}{dt}$.

$$\frac{dA}{dt} = \frac{1}{2}|h|$$

But we know that h is constant by Lemma 6.8.1. Therefore, the rate of change of the area A must be constant. □

6.8.3 Kepler's Third Law

Proposition 6.8.5. *(Kepler's Third Law) The period T of the revolution and the semi-major axis a of the ellipse satisfy $T^2 = \alpha a^3$ for some constant α.*

Proof. We start with the last equation in the previous proof.

$$A(t) = \frac{|h|}{2}t$$

The period T is the time needed to complete one whole orbit. The area swept out over time T should be the whole ellipse. Recall an ellipse with semiaxes a and b has area πab. We make the two areas equal.

$$A(T) = \frac{|h|T}{2} = \pi ab \implies T = \frac{2\pi ab}{|h|} \implies T^2 = \frac{4\pi^2 a^2 b^2}{|h|^2} \tag{6.8.7}$$

From Equation 6.7.3, we have expressions for the semimajor and semiminor axes (a and b) in terms of the eccentricity e and the distance d from the focus to the directrix.

$$a^2 = \frac{e^2 d^2}{(1-e^2)^2} \qquad\qquad b^2 = \frac{e^2 d^2}{1-e^2}$$

We take the square root of a^2 and then the ratio of the two equations.

$$\frac{a}{b^2} = \frac{\frac{ed}{(1-e^2)}}{\frac{e^2 d^2}{1-e^2}} = \frac{1}{ed} \implies ed = \frac{b^2}{a} \tag{6.8.8}$$

Now we can relate ed to G, M and h. Recall the definition of e from Equation 6.8.2 and d from Equation 6.8.3. We calculate the product ed.

$$ed = \frac{|h|^2}{GM} \implies |h|^2 = edGM \tag{6.8.9}$$

We use this to substitute for $|h|^2$ in 6.8.7.

$$T^2 = \frac{4\pi^2 a^2 b^2}{GMed}$$

Then use Equation 6.8.8 to substitute for ed and simplify.

$$T^2 = \frac{4\pi^2 a^2 b^2}{\frac{GMb^2}{a}} = \frac{4\pi^2}{GM} a^3$$

This is the desired result. The proportionality constant is $\frac{4\pi^2}{GM}$, which is certainly constant. □

Chapter 7

Topology of $\mathbb{R}^n$

Topology is a branch of mathematics which studies concepts of shape, near-ness and distance. You might think that such a study woulds just be geometry, and you would be partially correct. Topology is like a coarser version of geometry. It cares about broad notions of near and far instead of precise distances, and about types of shapes instead of precisely defined specifics. Though it lacks the precision of other parts of geometry, it is fundamental to modern mathematics.

We have already seen some topology of $\mathbb{R}$ in single variable calculus, mostly notably in the definition of open and closed intervals: (a, b) and $[a, b]$. Open intervals did not include their endpoints but closed intervals did. Openness and closedness are central to topolgoy. In this brief chapter, we will extend the notions of opennes, closedness and intervals to $\mathbb{R}^n$. We will need these notions to understand multivariable functions in the next section.

7.1 Open and Closed Sets

Definition 7.1.1. Let S be a set. A *topology* on S is a choice of subsets which are open and which subsets are closed.

We need two definitions to allow us to define open and clsoed sets in $\mathbb{R}^n$.

Definition 7.1.2. Let A be a subset of $\mathbb{R}^n$. A point $a \in A$ is called an *interior point* if there exists $\epsilon > 0$ such that all points b with $|a - b| < \epsilon$ are also in A.

A point is an interior point if all nearby points are also in the set. Around an interior point, we can move a little bit in any direction and remain inside the set. This ϵ measures exactly how little the little bit of movement is.

Definition 7.1.3. Let A be a subset of $\mathbb{R}^n$. A point a (not necessarily in the set) is called a *boundary point* if for any $\epsilon > 0$ there exists points b_1 and b_2 such that both $|b_i - a| < \epsilon$, b_1 is in the set, but b_2 isn't. The *boundary* of a set is the set of all its boundary point.

A point is a boundary point if there are nearby points in the set and nearby points not in the set. Nearby means within some small distance, measured by the small positive number ϵ. The two definitions are mutually exclusive: a point cannot be both an interior and a boundary point for a set. (It can, of course, be neither). These two definitions allow us to define open and closed sets in $\mathbb{R}^n$.

Definition 7.1.4. Let A be a subset of $\mathbb{R}^n$. Then A is called *open* is all of its points are interior points. Equivalently, A does not contin any boundary points. Alternatively, A is called *closed* if it contains all of its boundary points (or, equivalently, if it contains its boundary).

This definition properly extends the notions of open and closed intervals in $\mathbb{R}$. The boundary points of an interval are the endpoints. The open intervals lacks their endpoints and the closed intervals included their endpoints. In $\mathbb{R}^n$, the boundaries are much more complicated than endpoints: they can be intricate geometric shapes.

Example 7.1.5. The sphere S^{n-1} in $\mathbb{R}^n$ is all vectors of length one. The closed ball B^n in $\mathbb{R}^n$ is all vectors of length one or less. The sphere $S^{n-1}$is the boundary of the closed ball B^n, and the closed ball is closed because all points of the sphere are all contained in the closed ball.

7.2 Intervals

In addition to open and closed sets in $\mathbb{R}^n$, the is a direct extension of intervals.

Definition 7.2.1. An *open interval* in $\mathbb{R}^n$ is a set of points $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ such that

$$a_1 < x_1 < b_1, \qquad a_2 < x_2 < b_2, \qquad \ldots \qquad a_n < x_n < b_n$$

It is written

$$I = (a_1, b_1) \times (a_2, b_2) \times \ldots \times (a_n, b_n).$$

Definition 7.2.2. An *closed interval* in $\mathbb{R}^n$ is a set of points $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ such that

$$a_1 \leq x_1 \leq b_1, \qquad a_2 \leq x_2 \leq b_2, \qquad \ldots \qquad a_n \leq x_n \leq b_n$$

It is written

$$I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n].$$

Intervals are rectangular objects. In $\mathbb{R}^2$, there are filled-in rectangles, including the hollow bounding rectangle if the interval is closed and excluding it if the interval is open. In $\mathbb{R}^3$, intervals are solid rectangular prisms whose boundaries are rectangular boxes. In higher dimensions, we think of intervals are higher-dimensional rectangular prisms.

Chapter 8

Multivariable Functions and Partial Differentiation

8.1 Definitions

First year calculus dealt with functions $\mathbb{R} \to \mathbb{R}$. Parametric curves dealt with functions $\mathbb{R} \to \mathbb{R}^n$. Multivariable functions are functions $\mathbb{R}^n \to \mathbb{R}^m$. They having an arbitrary number of inputs and outputs. We can distinguish between two main categories, classified by output.

Definition 8.1.1. A *scalar function* is a function $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 8.1.2. A *vector function* is a function $f : \mathbb{R}^n \to \mathbb{R}^m$ for $m > 1$.

Vector functions are the most difficult and complicated, since they have both multiple inputs and multiple outputs. The study of vector functions is the major topic of Calculus IV. For the remainder of these notes, we will restrict our attention to scalar functions.

Example 8.1.3. There are many familiar scalar functions. The potential energy due to gravity on an object with mass m and height h above the surface of the earth is $PE = mgh$ is a two variable function. In a circuit, voltage in terms of current and reisistance is $V = IR$ is another two variable function. The force of gravity between two celestial object $F = -MmG/r^2$ is a three-variable functions which depends on both masses, M and m, as well as the distance between them r. Any quantity which can vary in three dimensional space, such a pressure, temperature, humidity, is a functions of the three variables of location (x, y, z).

Example 8.1.4. Here are some explicit functions $\mathbb{R}^2 \to \mathbb{R}$.

$$f_1(x,y) = x - y$$
$$f_2(x,y) = \sin(xy)$$
$$f_3(x,y) = \frac{1}{2}x^y$$
$$f_4(x,y) = x^2 + y^2 - xy$$
$$f_5(x,y) = \sqrt{x} + \sqrt{y}$$
$$f_6(x,y) = \sqrt{4 - x^2 - y^2}$$

Any other algebraic expression in x and y would also define a function $\mathbb{R}^2 \to \mathbb{R}$.

Definition 8.1.5. The *domain* of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ is the subset of $\mathbb{R}^n$ where the function can be defined.

Domain has the same kind of restrictions as for single-variable functions: no division by zero, no negative even roots, no negative logarithms, etc. However, now that the domain restrictions may apply to any of the input variables. The domains themselves may be very complicated subsets of $\mathbb{R}^n$.

Example 8.1.6. Let's look at the domains of the five explicitly stated functions in the Example 8.1.4

- f_1 and f_4 are polynomials (in two variables), so there are no restrictions. Their domain is $\mathbb{R}^2$.
- f_2 is a sine function, which again imposes no restrictions, so it also has domain $\mathbb{R}^2$.
- f_5 has two square roots, one involving x and one involving y. Therefore, we need $x \geq 0$ and $y \geq 0$ to defined f_5. That domain is the positive x and y quadrant, including the origin and the positive pieces of both axes.
- f_6 also has a square root. To ensure that the square root has a positive argument, we need $4 - x^2 - y^2 \geq 0$, or $x^2 + y^2 \leq 4$. This domain is a a circular disc of radius 2, including its boundary.
- Lastly, f_3, has an strange exponential. This leads to very strange domain behaviour. If y is an integer, x can be any non-zero real number. If y is a fraction, x must be positive is the denominator of y is even, to avoid square roots of negative numbers. If y is irrational, x must be positive. This all leads to a very complicated domain in $\mathbb{R}^2$.

8.2 Geometry and Graphs of Functions

The graph of a single variable $f : \mathbb{R} \to \mathbb{R}$ was a curve in $\mathbb{R}^2$ where one axis was input and one axis was output. The idea generalizes, so a graph has to show both the inputs and output.

Definition 8.2.1. Let $A \subset \mathbb{R}^n$ be the domain of a scalar function $f : A \to \mathbb{R}$. The *graph* of the scalar function is the subset of $\mathbb{R}^{n+1}$ consisting of all point $(x_1, x_2, \ldots, x_n, f(x_1, x_2, \ldots, x_n))$ for $(x_1, x_2, \ldots x_n) \in A$.

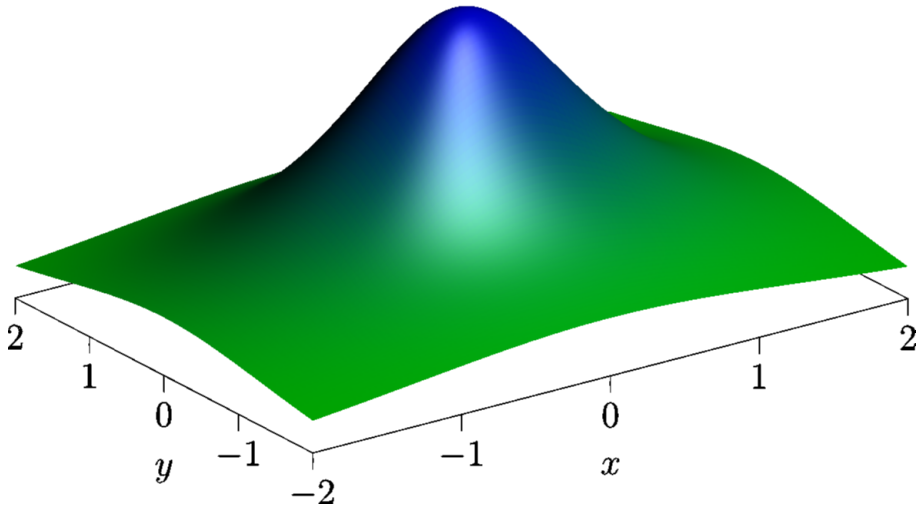

Figure 8.1: The graph of $f(x, y) = \frac{5}{x^2+y^2+1}$

Since we have to show input and outputs, the graph need many dimensions. If $n \geq 3$, then the graph is in $\mathbb{R}^4$ or a higher dimensional space. We can only actually see graphs of scalar function on $\mathbb{R}^2$.

The $\mathbb{R}^2$ case is useful to understand the general situation. In this case, x and y are in input (domain) and z is the output (range). We can think of the graph as a height function: over some point (x, y) in the domain of f, the graph sits at some height $z = f(x, y)$.

We used graphs extensively in single-variable calculus to understand derivatives and integrals: derivative were slopes of tangent lines and integrals were area under curves. We want to generalize this to the new higher-dimensional graphs. For functions $f : \mathbb{R}^2 \to \mathbb{R}$, it's not too difficult to extend the notions. Instead of tangent line, we now have *tangent planes*. Instead of area under the curve, we have volume under the surface. For $f : \mathbb{R}^n \to \mathbb{R}$, we have tangent n-spaces and (n+1)-hyper-volume under the n-dimensions graph surface.

8.2.1 Contour Plots

As an alternative to conventional graphs of function, a nice way to visualize height functions is as topological maps. We refer to these visualizations as contour plots.

Definition 8.2.2. Let $f : A \to \mathbb{R}$ be a scalar function for a domain $A \subset \mathbb{R}^2$. A *countour plot* for f is a plot of curves in $\mathbb{R}^2$ where each curve is a locus of the form $f(x, y) = c$ for some constant c.

A contour plot has a series of implicit curves at constant elevation; the constants c are the elevation. It shows curves where the function takes a specific value. By looking at the relationships of the curves, we can intuit how the function behaves.

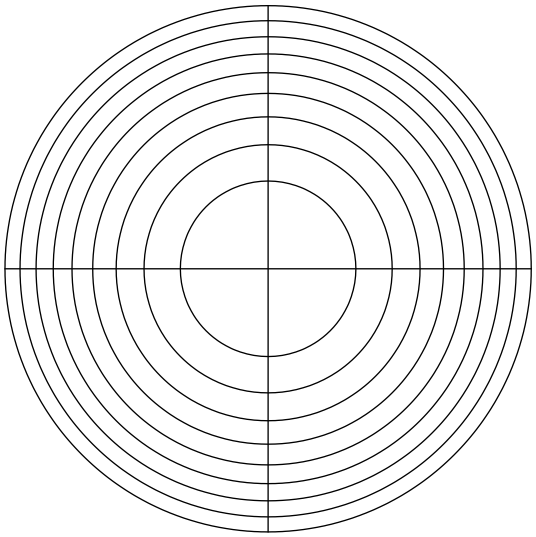

Figure 8.2: The circular contours of $f(x,y) = \frac{5}{x^2+y^2+1}$

Example 8.2.3. Consider $f(x,y) = \frac{5}{x^2+y^2+1}$. It's graph is Figure 8.1. This function has a simple hill at the origin and slopes down in all direction. The contours are loci of the form $\frac{5}{x^2+y^2+1} = c$, which can be rearranged as $\frac{5}{c} - 1 = x^2 + y^2$. These contours are all circles, and are shown in Figure 8.2.

Example 8.2.4. Contour diagrams can show many behaviours. Figure 8.3 shows a pass or saddle point. (The saddle point will be formally defined in Definition 9.1.2.

These contour plots lead us to a general definition.

Definition 8.2.5. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar function. A *level set* for f is a subset of $\mathbb{R}^n$ given by the equation $f(x_1, x_2, \ldots, x_n) = c$ for some $c \in \mathbb{R}$.

Then a countour plot is just a drawing of a variety of level sets of a function $\mathbb{R}^2 \to \mathbb{R}$. It is useful to see where a function is constant. The resulting shapes tell us a great deal about the behaviour of the function. Level sets for $f : \mathbb{R}^3 \to \mathbb{R}$ are called level surfaces.

8.3 Limits of Multivariable Functions

We want to re-establish the tools of calculus for multi-variable functions. As with single-variable functions, our starting point is limits. Let's start with the full, formal, $\epsilon - \delta$ definition.

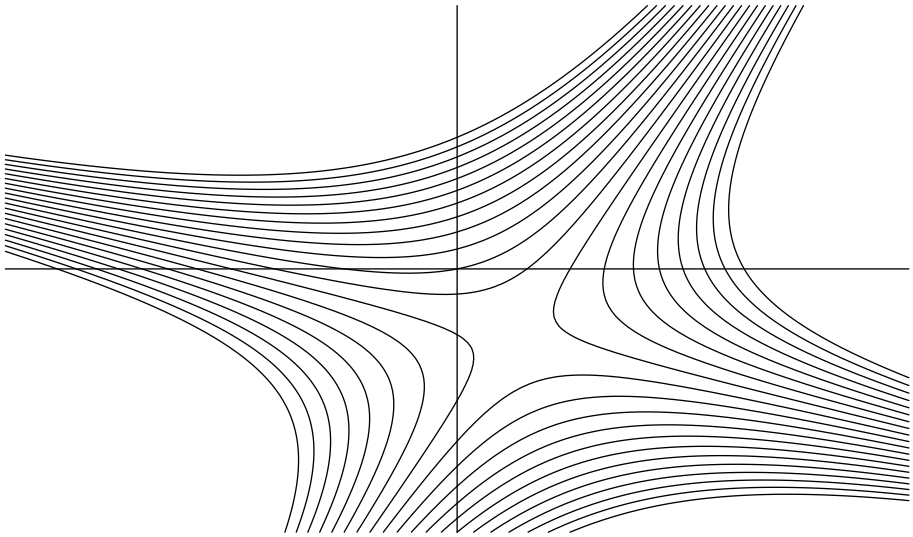

Figure 8.3: The contours of a saddle point

Definition 8.3.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Let $(a_1, a_2, \ldots, a_n)$ be a point in $\mathbb{R}^n$. Then the statement

$$\lim_{(x_1,x_2,\ldots,x_n)\to(a_1,a_2,\ldots,a_n)} f(x_1, x_2, \ldots, x_n) = L$$

means that $\forall \epsilon > 0$ $\exists \delta > 0$ such that if $|(x_1, x_2, \ldots, x_n) - (a_1, a_2, \ldots, a_n)| < \delta$ then $|f(x_1, x_2, \ldots, x_n) - L| < \epsilon$.

The definition is essentially the same as the single-variable definition: as the input gets closer and closer to a specific point, the output gets closer and closer to a fixed value L. The only issue is that 'closer and closer' now happens in $\mathbb{R}^n$ instead of $\mathbb{R}$.

In $\mathbb{R}$, we only had two directions of approach: from the left and from the right. If the behaviour from both sides was the same, we said the limit existed. In $\mathbb{R}^n$ for $n \geq 2$, we have infinitely many ways to approach any given point. We can approach along lines in infinitely many directions out from the point. Even more, we can approach along other paths, such as spiral paths or stranger jagged paths. This makes it much more difficult to determine the behaviour and much more difficult to prove existence of various limits. However, we do have some good news. First, the definition of continuity remains the same.

Definition 8.3.2. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called continuous at $(a_1, a_2, \ldots, a_n)$ if the limit approaching $(a_1, a_2, \ldots, a_n)$ exists and is the same as $f(a_1, a_2, \ldots, a_n)$.

So limits for continuous functions are still reasonable: we just evaluate. But what functions are continuous?

Proposition 8.3.3. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous if and only if it is continuous independently in each of its variables.*

This is the first application of a very important technique: treating the function as a function of only one of the variables and ignoring the other, pretending that they are constant. If we do that, we get n different single variable functions. The proposition says that the function is continuous in the new definition if and only if all of the n different single variable functions are continuous.

This proposition tells us how to recognize continuous functions. Anything involving polynomials, roots, rational functions, trig, exponentials and logarithms is continuous on its domain.

As we mentioned before, proving the existence of limits is difficult. However, the algebraic techniques of first year calculus can still work for calculations. Here are some examples.

Example 8.3.4.

$$\begin{aligned}\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x-y} &= \lim_{(x,y)\to(0,0)} \frac{(x-y)(x+y)}{(x-y)} \\ &= \lim_{(x,y)\to(0,0)} x+y = 0+0 = 0\end{aligned}$$

$$\begin{aligned}\lim_{(x,y)\to(4,1)} \frac{xy-4y^2}{\sqrt{x}-2\sqrt{y}} &= \lim_{(x,y)\to(4,1)} \frac{y(\sqrt{x}-2\sqrt{y})(\sqrt{x}+2\sqrt{y})}{\sqrt{x}-2\sqrt{y}} \\ &= \lim_{(x,y)\to(4,1)} \sqrt{x}+2\sqrt{y} = \sqrt{4}+2\sqrt{1} = 2+2 = 4\end{aligned}$$

Example 8.3.5.

$$\lim_{(x,y)\to(0,0)} \frac{(x+y)^2}{x^2+y^2}$$

Let's approach this along the line $y = mx$. That lets us replace y with mx in the calculation and the limit becuase $x \to 0$.

$$\begin{aligned}\lim_{x\to 0} \frac{(x+mx)^2}{x^2+m^2x^2} &= \lim_{x\to 0} \frac{x^2+2mx+m^2x^2}{x^2+m^2x^2} \\ &= \lim_{x\to 0} \frac{x^2(1+m)^2}{x^2(1+m^2)} = \frac{(1+m)^2}{1+m^2}\end{aligned}$$

This limit depend on the choice of m! We can get infinitely many values (all between 0 and 1) out of this limit depending on which line we use to approach $(0,0)$. With all these possible answers, the limit cannot exist. It is interesting to try to visualize the graph: as we get close to zero, there are pieces of the graph getting close to any number between 0 and 1. Figure 8.4 shows some of this behaviour: approaching from the front of figure leads to 0, but approaching from the sides leads to larger numbers. (The graph is slightly flawed, due to the graphing algorithm. The two cliffs should meet, even though the graph shows a gap between them. Where the two cliffs meet is the line with all the problematic limits.)

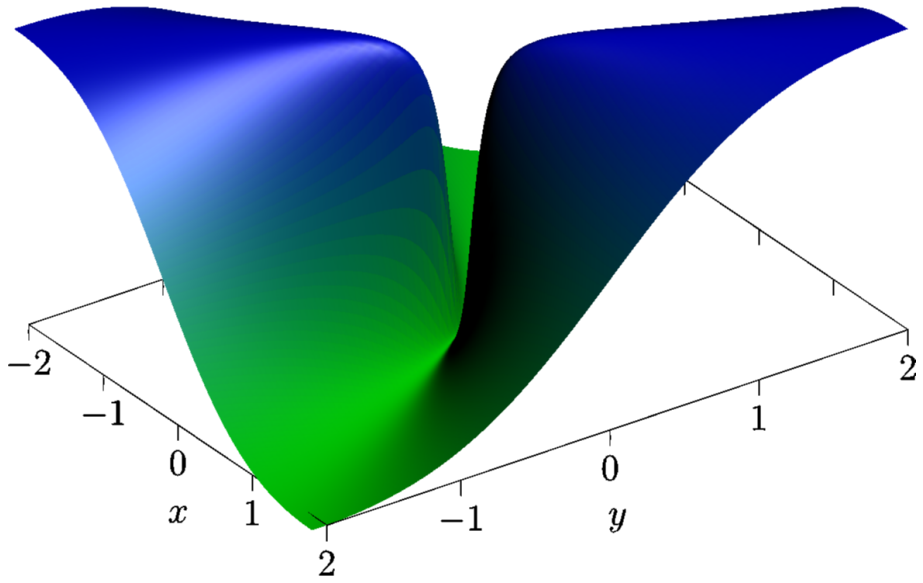

Figure 8.4: The graph of $f(x) = \frac{(x+y)^2}{x^2+y^2}$

Example 8.3.6. In this example, we approach along parabolic paths of the form $x = my^2$ and, like the previous example, get a limit that depends on the path of approach.

$$\lim_{(x,y)\to(0,0)} \frac{3xy^2}{x^2+y^2} = \lim_{(x,y)\to(0,0)} \frac{3xy^2}{x^2+y^2} = \lim_{y\to 0} \frac{3my^4}{m^2y^4+y^4} = \frac{3m}{m+1}$$

Example 8.3.7.

$$\lim_{(x,y)\to(1,2)} \frac{\sqrt{y}-\sqrt{x+1}}{y-x-1} = \lim_{(x,y)\to(1,2)} \frac{y-x-1}{(y-x-1)(\sqrt{y}+\sqrt{x+1})}$$
$$= \frac{1}{\sqrt{2}+\sqrt{2}} = \frac{1}{2\sqrt{2}}$$

Example 8.3.8. Sometime we can make reasonable conclusion independent of path.

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x+y)}{x+y}$$

We can assume any path $y = f(x)$ of approach and still set $u = x + y$, which turns this into a limit of $\sin(u)/u$, which tends to 1.

8.4 Partial Derivatives

The first of several generalizations of derivatives is the partial derivative.

Definition 8.4.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar function. If x is one of the variable, then the derivative of f which pretend that all other variables are constant is called the *partial derivative* of f in the variable x.

The notation for partial derivatives resembles Leibniz notation for ordinary derivatives: $\frac{df}{dx}$. Leibniz style notation is useful since the variable of differentiation is explicit. The partial derivatives of f in terms of x, y, z, or x_i are written with a stylized version of d in Leibniz notation.

$$\frac{\partial f}{\partial x} \qquad \frac{\partial f}{\partial y} \qquad \frac{\partial f}{\partial z} \qquad \frac{\partial f}{\partial x_1}$$

For interpretation, this gives us the rate of change of f with respect to one of its variables. We don't have a holisitic notion of rate of change, but we can see how it changes in each variable.

If we wanted to be formal, pretending that all the other variales are constant is the same as taking limits in one variable. For $f : \mathbb{R}^2 \to \mathbb{R}$, these limits defined the two partial derivatives.

$$\frac{\partial f}{\partial x}(a,b) = \lim_{(x,y)\to(a,b)} \frac{f(a+h,b) - f(a,b)}{h}$$
$$\frac{\partial f}{\partial y}(a,b) = \lim_{(x,y)\to(a,b)} \frac{f(a,b+h) - f(a,b)}{h}$$

Notice that the value of $\frac{\partial f}{\partial x}$ still depends on y. Different y values identify different point in the domain, where the rate of change with respect to x may differ.

Definition 8.4.2. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *differentiable* at $(a_1, a_2, \ldots, a_n)$ if *all* of its partial derivatives $\frac{\partial f}{\partial x_i}$ exists at that point.

As we did for single-variable calculus, we use $\frac{\partial}{\partial x}$ and similar expressions as operators – the things that take derivatives.

There are various notational conventions for partial derivatives. In addition to $\frac{\partial f}{\partial x}$, we can also write this as f_x, $\partial_x f$ and $D_x f$. The first of these is a nice short-hand which we will use frequently.

Example 8.4.3.

$$f(x,y) = x^2 + y^2 \sin x \qquad \frac{\partial f}{\partial x} = 2x + y^2 \cos x \qquad \frac{\partial f}{\partial y} = 2y \sin x$$

$$f(x,y) = \frac{1}{xy + e^{xy}} \qquad \frac{\partial f}{\partial x} = \frac{-1}{x^2 y} + xe^{xy} \qquad \frac{\partial f}{\partial y} = \frac{-1}{xy^2} + ye^{xy}$$

$$f(x,y) = \frac{x^2y^2 + xy}{x^2 + y^2 + 1}$$
$$\frac{\partial f}{\partial x} = \frac{(2xy^2 + y)(x^2 + y^2 - 1) - (x^2 + y^2 + xy)(2x)}{(x^2 + y^2 + 1)^2}$$
$$\frac{\partial f}{\partial y} = \frac{(2x^2y + x)(x^2 + y^2 - 1) - (x^2y^2 + 2y)(2y)}{(x^2 + y^2 + 1)^2}$$

Example 8.4.4. We can also take higher partial derivatives.

$$f(x,y) = 3x^3y^2 - xy^4 + x^2y^3 - 3$$

$$\frac{\partial f}{\partial x} = 9x^2y^2 - y^4 + 2xy^3 \qquad \frac{\partial f}{\partial y} = 6x^3y - 4xy^3 + 2x^2y$$

$$\frac{\partial^2 f}{\partial x^2} = 18xy^2 + 2y^3 \qquad \frac{\partial^2 f}{\partial y^2} = 6x^3 - 12xy^2 + 2x^2$$

$$\frac{\partial^3 f}{\partial x^3} = 18y^2 \qquad \frac{\partial^3 f}{\partial y^3} = -24xy$$

$$\frac{\partial^4 f}{\partial x^4} = 0 \qquad \frac{\partial^4 f}{\partial y^4} = -24x$$

$$\frac{\partial^5 f}{\partial x^5} = 0 \qquad \frac{\partial^5 f}{\partial y^5} = 0$$

Example 8.4.5. In addition, we can mix the partial derivatives, first taking the partial in x and the in y or vice-versa.

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x} = 18x^2y - 4y^3 + 4xy$$

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} = 18x^2y - 4y^3 + 4xy$$

Curiously, we get the same answer from either order of the mixed partial derivatives. This equality is true in general and is called Clairaut's Theorem. In the statement of the proposition, notice the requirement that the domain is an open set. This is an example of the topological conditions invovled in multivariable calculus.

Proposition 8.4.6. *Let $f(x,y)$ be a function $\mathbb{R}^2 \to \mathbb{R}$. If all the partial derivatives f_{xx}, f_{yy}, f_{xy} and f_{yx} exists and are continuous on an open set D, then $f_{xy} = f_{yx}$ on that set.*

8.5 Partial Differential Equations

Now that we have defined partial derivatives, we can introduce what is possibly the most important setting for their use: partial differential equations. We can start with two classic examples. The first is diffusion of heat.

Let's say we have a 1-dimensional rod where length is measured with the variable x. Heat can vary along the rod, so we measure it by a function $u(x)$. However, this distribution of heat can also change over time. Therefore, we should measure the heat distribution both in terms of position x along the rod and time t, $u(x,t)$.

We need to consider several aspects of the situation to give a full account of how heat will diffuse. First, let's look at the mechanics of heat. We make the assumption that heat wants to equalize; with

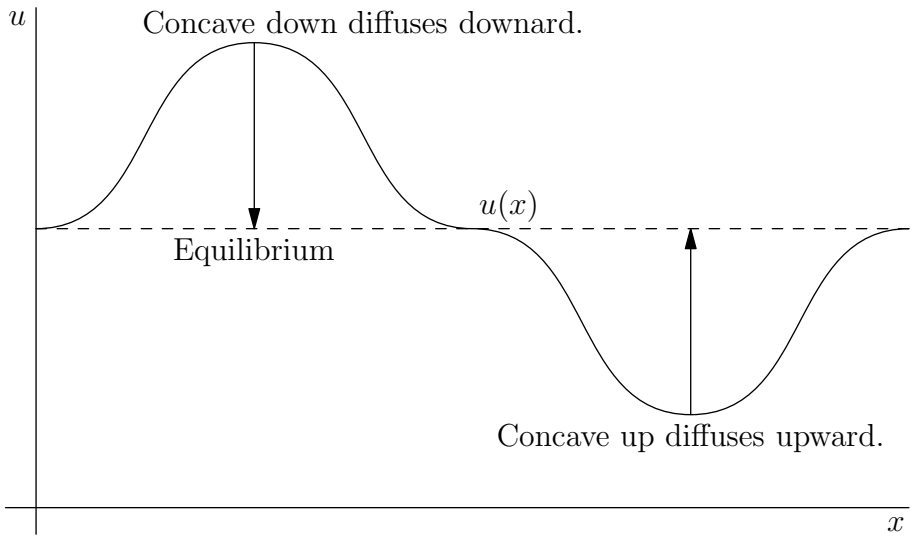

Figure 8.5: Concavity and Heat Diffusion

the absence of external addition of heat, it will diffuse until it equals out everywhere. If addition, we assume that the greater the variance in heat between two adjacent points on the rod, the faster the heat will diffuse. How do we translate this assumptions into mathematics? We need to get a measure of this variance in heat. The measure is local, since heat only diffuses to points adjacent. So how do we meausre how much local variance there is in heat?

Consider the heat picture at some fixed time t_0: $u(x, t_0)$. It is not the value of the heat that determines diffusion, since nearby values can be higher or lower. It is also not the slope of this graph in x that determines the diffusion, since a straight line slope represents and even flow of heat from one end to the other. We can think of heat wanting to return to this even flow, this straight line: so it is the curvature of the graph that disrupts the straight line. Curvature or concavity is measured by the second derivative. Therefore $u_{xx}(x, t_0)$ measures the tendency for the heat to diffuse. Figure 8.5 illustrates how concavity causes heat diffusion.

Diffusion creates change: heat will leave or enter the point. Change is measured by the time derivatives $u_t(x, t)$. So convacity, the second space deriavitve, must be related to the first time derivative. What is the relationship between these? Let's assume the simpliest case for now and make the relationship linear. That means there is a constant α such that

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}.$$

This equation is called the Heat Equation. Though relatively simple, it is one of the most important partial differential equations.

The equation, however, isn't enough to solve the problem. We also need boundary conditions and initial conditions. The boundary conditions tell us what happens at the end of the rod: $u(0, t)$ and $u(l, t)$. In

principle, these could be anything functions of t. For now, let's assume they are constant $u(0,t) = a$ and $u(l,t) = b$.

We also need intial conditions. These tells us the starting heat distribution at at particular time, say $t = 0$, That is $u(x,0) = f(x)$, a single variable function that tells us the original situation.

All together, this information determines the physical system. We can then try to find a function $u(x,t)$ which matches the equation, the boundary conditions and the initial conditions. Such a task is often very difficult to do. However, if both boundary conditions are constant 0 and the initial condition is $f(x) = \sin\left(\frac{\pi x}{l}\right)$ then the function

$$u(x,t) = e^{-\frac{\alpha\pi^2 t}{l^2}} \sin\left(\frac{\pi x}{l}\right)$$

solves the partial differential equation. This is an ideal case: in general the solutions become much more complicated. This initial heat distribution is half a period of a sine wave, and the time dependence is a exponential decay of the amplitude of that sine wave back towards a stable 0-level heat distribution.

This example is archtypical of many partial differential equations. We will almost always have a function which depends on time as well as some other quantities. In physics, these other quantities are usually position. Then the equation is usually organized by taking a time derivative on one side and a position derivative on the other. Then we posit a relationship between the two two derivatives. In addition, boundary conditions tell us what happens at the edges of our environment and initial conditions give us a snapshot of the situation at a fixed moment in time. Then we try to find a multi-variable function that fits all the information.

Another very familiar situation is a wave moving through an elastic medium. We'll assume a 1-dimensional elastic medium (think a wire or string), then $u(x,t)$ measures the displacement of the medium at position x and time t. The physical motivation is similar to the heat equation: the concavity measure the offset of the stiuation from a stable straight line. However, this concavity, instead of causing heat diffusion, causes acceleration on the adjacent points of the elastic medium. The elastic medium doesn't diffuse back to equilibrium, it accelerates, like a spring, back to equilibrium. Acceleration is a second time derivative, so the Wave Equation is

$$\frac{\partial^2 u}{\partial t^2} = \alpha \frac{\partial^2 u}{\partial x^2}.$$

Again, there are boundary conditions and initial condition and, in general, the problem is very difficult to solve. However, if we take the same situation as before, with constant zero boundary conditions at $x = 0$ and $x = l$ and initial wave profile $f(x) = \sin\left(\frac{\pi x}{l}\right)$, then the solution is

$$u(x,t) = \sin\left(\frac{\sqrt{\alpha}\pi t}{l}\right) \sin\left(\frac{\pi x}{l}\right).$$

So, instead of decay to equilibrium, we get an oscillating amplitude, resulting is a very simple standing wave on the wire or string. This is a very simple version: there are no higher harmonics and there is no friction which causes decay over time.

Many other famous equations are relationships between times derivatives and position derivatives. In all of the following examples, the left side is a time derivative and the right is a space derivative.

In classical Newtonian mechanics, for conservative forces like gravity and electromagnetism, the force is often determined by a potential energy field V.

$$F = -\frac{dV(x)}{dx}$$

But $F = ma$ is Newton's first law, and acceleration is a time derivative of position.

$$\frac{d^2x}{dt^2} = \frac{-1}{m}\frac{dV}{dx}$$

The Schrodinger equation is the centre of quantum mechanics: it measures a wave function $\Psi(x, y, z, t)$ in three dimensions and time. The symbol ∇ is a 3-dimemsional differential operator which will be defined later; for now, just know that it is a combination of position derivatives. $\hbar$ is a constant, $\imath$ is a number with $\imath^2 = -1$, m is mass and $V(x, y, z)$ is a potential energy function.

$$\imath\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi$$

Another famous example is the Navier-Stokes equation, which is the fundamental equation of fluid dynamics. There are a number of versions of it, but I'll only write one. The function $v(x, y, z, t)$ is the flow velocity of a three-dimensional fluid. Again, ∇ is a 3-dimensional position differential operator. ρ is the fluid density, p is the pressure, T is something called a stress tensor and f is an external force of the fluid. The equation is

$$\rho\frac{\partial v}{\partial t} = \rho v \cdot \nabla v - \nabla p + \nabla \cdot T + f$$

Solving the Navier-Stokes equation for various initial and boundary conditions is the subject of a whole branch of mathematical physics called fluid dynamics.

8.6 Gradients

After partial derivatives, we want to proceed to define several other generalizations of the derivative. The first is the gradient.

Definition 8.6.1. The gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ is written ∇f and defined as

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

The gradient is a new function $\mathbb{R}^n \to \mathbb{R}^n$. It ouputs the *vector* of partial derivatives of f at any point in its domin. Note that the gradient is a local direction vector *in the domain*.

The best interpretation of the gradient comes from contour plots. Like the gradient, countour plots live in $\mathbb{R}^n$, the domain, and show the level (hyper)surfaces of the function. If, as in $\mathbb{R}^2$, we think of the function as a height function and the countour plot as a topographical map, the gradient shows direction of greatest increase.

If we draw topographical lines on a countour plot, the gradient will always be locally perpendicular to those lines and will point in the direction of greatest increase. Rephrased, this is a useful result: ∇f is always the normal to the level sets of f. If those level sets are hypersurfaces, their tangent planes can be determined by the normal ∇f.

Example 8.6.2. A central example of gradients is found by considering the gravitational potential energy function caused by a mass m at the the origin. Another object of mass m and position (x, y, z) has potential gravitational energy of

$$P = \frac{-GmM}{\sqrt{x^2 + y^2 + z^2}}.$$

By convention, this potential energy is negative. It approaches 0 in the limit at the origin, and approaches $-\infty$ as we get very far from the origin. The gradient of this is

$$\nabla P = \frac{GmM}{\sqrt{(x^2 + y^2 + z^2)^3}}(x, y, z) = \frac{GmM}{x^2 + y^2 + z^2} \frac{1}{\sqrt{x^2 + y^2 + z^2}}(x, y, z).$$

This is precisely the force of gravity. The gradient point in the direction of maximum increase in potential energy with magnitude $\frac{GmM}{r^2}$ where r is the distance between the two objects. This is a common situation we will discuss in Calculus IV: many forces are the result of gradients of potential energy functions.

Example 8.6.3. In another example, consider a function $p(x, y, z)$ which measures the pressure in a rotating cylindrial drum. (Think of a centrifuge). With bounds $z \in [0, 5]$ and $x, y \in [0, \sqrt{3}]$, the function is

$$p(x, y, z) = \frac{1}{z + 1}(x^2 + y^2)$$

The gradient is

$$\nabla p(x, y, z) = \left(\frac{2x}{z + 1}, \frac{2y}{z + 1}, \frac{-(x^2 + y^2)}{(z + 1)^2} \right)$$

This points in the direction of greatest increase in pressure. It is perpendicular to the surfaces of constant pressure. If there is a difference in media in the drum, the lighter medium will be force towards the edges of the drum in the direction of this gradient.

In addition to the gradient, we can think of ∇ itself as the following differential operator.

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \ldots, \frac{\partial}{\partial x_n} \right)$$

This is a vector-valued differential operator: it outputs the vector of partial derivatives ∇f. Now that we have this operator, there are other operations we can use it for. Most of those operations come in Calculus IV, but we can define one such operation here.

Definition 8.6.4. If $f : \mathbb{R}^n \to R$ is a scalar function, the *Laplacian* of f is given by applying ∇ twice. Since ∇ outputs a vector, the second application uses the dot product to output a scalar.

$$\begin{aligned} \nabla^2 f = \nabla \cdot \nabla f &= \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}, \ldots, \frac{\partial}{\partial x_n} \right) \cdot \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \ldots, \frac{\partial f}{\partial x_n} \right) \\ &= \frac{\partial^2 f}{\partial^2 x_1} + \frac{\partial^2 f}{\partial^2 x_2} + \frac{\partial^2 f}{\partial^2 x_3} + \ldots \frac{\partial^2 f}{\partial^2 x_n} \end{aligned}$$

The Laplacian, as a second derivative, measure some kind of multi-dimensional concavity. We considered the heat equation in one dimension; in that equation, concavity measured local displacement from equilibrium. The Laplacian does the same in multiple dimensions. The general Heat Equation is

$$\frac{\partial u}{\partial t} = \alpha \nabla^2 u.$$

Similarly, the general Wave Equation is

$$\frac{\partial^2 u}{\partial t^2} = \alpha \nabla^2 u.$$

8.7 Directional Derivatives

Partial derivatives took one variable and pretended that all other variables were constant. In that way, we got the rate of change in that variable. We could consider $\frac{\partial f}{\partial x}$ the derivative of f *when we move* in the x axis direction. But why do we only need to move in the axis direction? Why can't we move in all directions and consider the rate of change?

Definition 8.7.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and u a *unit* vector in $\mathbb{R}^n$. The directional derivative of f in the direction u is written $D_u f$ and given by a limit definition. Let v be a point in the domain of f.

$$D_u f(v) = \lim_{h \to 0} \frac{f(v + hu) - f(v)}{h}$$

The directional derivative, like the partial derivative, uses a single variable limit: we use the line in the direction u (as a local direction vector form the point v) to give a one-dimensional domain – a copy of $\mathbb{R}^1$. Then we just differentiate along the line. If $u = e_1$, we get $D_{e_1} f = f_{x_1}$. If $u = e_2$ we get $D_{e_2} f = f_{x_2}$ and so on.

Instead of calculating this limit every time, we have a nice tool for calculating directional derivatives.

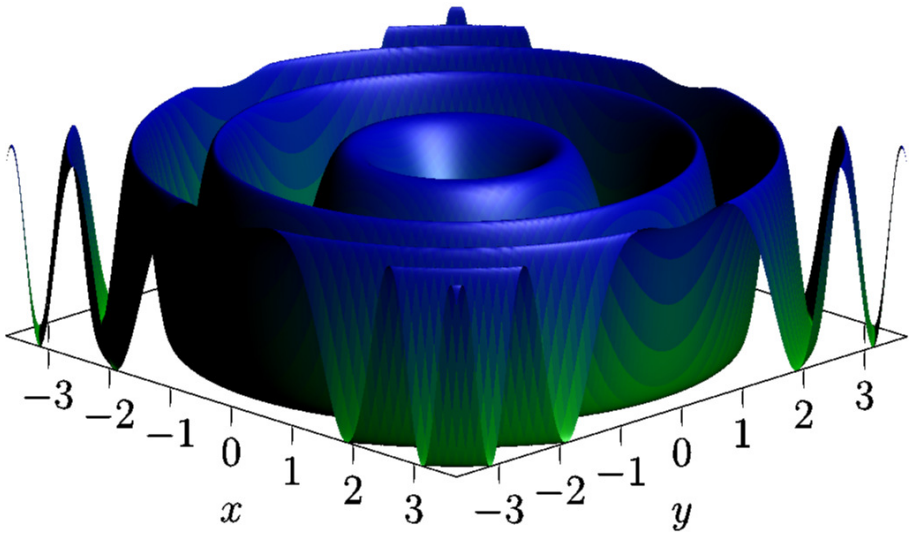

Figure 8.6: The function $f(x, y) = \sin(x^2 + y^2)$.

Proposition 8.7.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and u a* unit *vector in $\mathbb{R}^n$. The directional derivatives $D_u f$ is the dot product of u with ∇f.*

$$D_u f = u \cdot \nabla f$$

If (a, b) or (a, b, c) are unit vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively, we can write the specific form of the proposition for low dimensions.

$$D_{(a,b)} f(x, y) = \frac{\partial f}{\partial x} a + \frac{\partial f}{\partial y} b$$
$$D_{(a,b,c)} f(x, y, z) = \frac{\partial f}{\partial x} a + \frac{\partial f}{\partial y} b + \frac{\partial f}{\partial z} c$$

As we noted above, the directional derivatives in the axis directions give the partial derivatives, so this is an extension of the idea of partial derivatives.

$$D_{(1,0)} f(x, y) = \frac{\partial f}{\partial x}$$
$$D_{(0,1)} f(x, y) = \frac{\partial f}{\partial y}$$
$$D_{(1,0,0)} f(x, y, z) = \frac{\partial f}{\partial x}$$
$$D_{(0,1,0)} f(x, y, z) = \frac{\partial f}{\partial y}$$
$$D_{(0,0,1)} f(x, y, z) = \frac{\partial f}{\partial z}$$

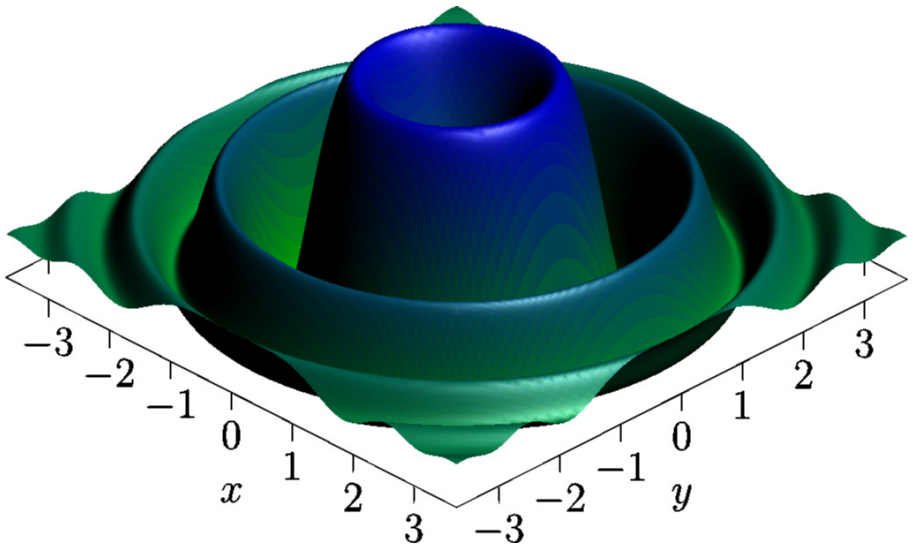

Figure 8.7: The function $f(x,y) = e^{-(x^2+y^2)}\sin(x^2+y^2)$.

Example 8.7.3. Consider this function: $f(x,y) = \sin(x^2+y^2)$, show in Figure 8.6. These are circular sine waves, like riples on a pond which never decrease in amplitude. We have

$$\begin{aligned}
D_{(1,0)}f(x,y) &= 2x\cos(x^2+y^2)\\
D_{(0,1)}f(x,y) &= 2y\cos(x^2+y^2)\\
D_{\left(\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right)}f(x,y) &= \frac{2}{\sqrt{5}}x\cos(x^2+y^2) + \frac{2}{\sqrt{5}}y\cos(x^2+y^2)\\
D_{\left(\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right)}f(\sqrt{\pi},\sqrt{\pi}) &= \frac{2}{\sqrt{5}}\sqrt{\pi}\cos(\pi+\pi) + \frac{2}{\sqrt{5}}\sqrt{\pi}\cos(\pi+\pi) = 4\sqrt{\frac{\pi}{5}}
\end{aligned}$$

If we wanted damped ripples instead, as in Figure 8.7, we would take $f(x,y) = e^{-(x^2+y^2)}\sin(x^2+y^2)$.

$$\begin{aligned}
D_{(a,b)}f(x,y) &= \left[-2xe^{-(x^2+y^2)}\sin(x^2+y^2) + 2xe^{-(x^2+y^2)}\cos(x^2+y^2)\right]a\\
&\quad + \left[-2ye^{-(x^2+y^2)}\sin(x^2+y^2) + 2ye^{-(x^2+y^2)}\cos(x^2+y^2)\right]b\\
D_{(a,b)}f(\sqrt{\pi},\sqrt{\pi}) &= \left[-2\sqrt{\pi}e^{-(\pi+\pi)}\sin(\pi+\pi) + 2\sqrt{\pi}e^{-(\pi+\pi)}\cos(\pi+\pi)\right]a\\
&\quad + \left[-2\sqrt{\pi}e^{-(\pi+\pi)}\sin(\pi+\pi) + 2\sqrt{\pi}e^{-(\pi+\pi)}\cos(\pi+\pi)\right]b\\
&= \frac{2\sqrt{\pi}}{e^{2\pi}}\left[\cos(2\pi)a + \cos(2\pi)b\right] = \frac{2\sqrt{\pi}(a+b)}{e^{2\pi}}
\end{aligned}$$

Finally, look at what happens when we apply the length of a dot product to the directional derivative.

$$|D_u f| = |\nabla f \cdot u| = |\nabla f||u|\cos\theta$$

The cosine term is maximized when the angle $\theta = 0$, that is, when u is the unit vector in the same direction as ∇f. That is, the greatest directional derivative, representing the direction of fastest change, is found in the direction of the gradient. This established the fact, which we claimed earlier, that the gradient points in the direction of greatest change.

8.8 The Chain Rule

We defined partial derivatives to measure rates of change in a particular variable. We extended this to change in any unit direction with directional derivatives. We can extend this even further, but considering the change in a function as we move along a *parametric curve* in the domain.

Let $f(x, y, z) : \mathbb{R}^3 \to \mathbb{R}$, be a potential energy function. Let $\gamma(t) = (x(t), y(t), z(t))$ be a curve moving through $\mathbb{R}^3$. If this is a potential energy function, we want to know how quickly we gain and/or lose energy moving along the path γ. The energy along γ is $f(\gamma(t)) = f(x(t), y(t), z(t))$. The rate of change is $\frac{df}{dt}$. But now f is a composition, $f(\gamma(t))$, so this must be a chain rule calculation. What is the chain rule when we have three (or more) components?.

Proposition 8.8.1. *(The Chain Rule) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar function and $\gamma(t)$ a parametric curve in $\mathbb{R}^n$ inside the domain of f. The derivative of f along γ is*

$$\frac{d}{dt} f(\gamma(t)) = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \ldots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

The total rate of change is the sum of the rates of changes in each of the variables.

For reference, here is the chain rule in $\mathbb{R}^3$.

$$\frac{d}{dt} f((x(t), y(t), z(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

Example 8.8.2. Consider the potential gravitational energy function

$$P = -\frac{GmM}{r} = \frac{-GmM}{\sqrt{x^2 + y^2 + z^2}}$$

If we move along a curve γ, it is nice to know how the potential energy changes. A helical path out of the gravity well might be $\gamma(t) = (\sin t, \cos t, t)$. We differentiate along this path using the chain rule.

$$\begin{aligned}
\frac{dP}{dt} &= \frac{\partial P}{\partial x}\frac{dx}{dt} + \frac{\partial P}{\partial y}\frac{dy}{dt} + \frac{\partial P}{\partial z}\frac{dz}{dt} \\
&= \frac{GmMx}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}\frac{dx}{dt} + \frac{GmMy}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}\frac{dy}{dt} + \frac{GmMz}{(x^2 + y^2 + z^2)^{\frac{3}{2}}}\frac{dz}{dt} \\
&= \frac{GmM\sin t}{(1 + t^2)^{\frac{3}{2}}}\cos t + \frac{GmM\cos t}{(1 + t^2)^{\frac{3}{2}}}(-\sin t) + \frac{GmMt}{(1 + t^2)^{\frac{3}{2}}}1 \\
&= \frac{GmMt}{\sqrt{(1 + t^2)^3}}
\end{aligned}$$

Notice that if we let $|\gamma(t)| = \sqrt{1+t^2}$ at the start, we could have written $P(t) = \frac{-GmM}{|\gamma(t)|} = \frac{-GmM}{\sqrt{1+t^2}}$ and the $P'(t) = \frac{GmMt}{\sqrt{(1+t^2)^3}}$ could have been calculated directly. That would have been easier, but its nice to get confirmation that working with the chain rule leads to the right result.

8.9 Tangent Planes

In single variable calculus, derivatives allowed us to find the slopes and equations of tangent lines to the graph of a function. We want to extend this idea. For functions of two variables, we have graphs which are surfaces in $\mathbb{R}^3$ instead of curves in $\mathbb{R}^2$. These surfaces have tangent planes instead of tangent lines.

If we have such a function $f(x, y)$ let's look at a particular point $(a, b, f(a, b))$ on the graph of the function. We can calculate the partial derivatives $f_x(a, b)$ and $f_y(a, b)$. At the point $(a, b, f(a, b))$, these two partial derivatives give us the rate of change in x and in y. That's the slope of a tangent line in the x direction and a tangent line in the y direction. We'd rather have direction vectors than slopes, but we can construct these. For the x direciton, the y coordinate is 0 since there is no change in y. That gives the vector $(1, 0, f_x(a, b))$. Likewise in the y direction, we have the vector $(0, 1, f_y(a, b))$. These are two tangent local direction vectors.

If we have two vectors on a plane, the normal of the plane is given by the cross product. So we calculate $(1, 0, f_x(a, b)) \times (0, 1, f_y(a, b)) = (-f_x(a, b), -f_y(a, b), 1)$. Therefore, we have the following result.

Proposition 8.9.1. *Let $f(x, y)$ be a function $\mathbb{R}^2 \to \mathbb{R}$. The equation of the tangent plane to f at $(a, b, f(a, b))$ is*

$$z - f(a, b) = f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

Proof. We just calculated the normal at any point $(a, b, f(a, b))$ on the graph of the function. That vector was $(-f_x(a, b), -f_y(a, b), 1)$. From Section 4.10 the equation of the plane is given by the dot product of the variables with the normal. Here is this dot product, with an unknown value c.

$$-f_x(a, b)x - f_y(a, b)y + z = c$$

We have a point on the plane: $(a, b, f(a, b))$. By substitution, we can solve for c.

$$c = f(a, b) - f_x(a, b)a - f_y(a, b)b$$

Putting this c in gives the equation of the plane.

$$-f_x(a, b)x - f_y(a, b)y + z = f(a, b) - f_x(a, b)a - f_y(a, b)b$$

From here, is it just a re-arrangement to get the form in the proposition. □

Example 8.9.2. Consider $f(x,y) = \frac{1}{1+x^2+y^2}$.

$$\frac{\partial f}{\partial x} = \frac{-2x}{(1+x^2+y^2)^2} \qquad \frac{\partial f}{\partial y} = \frac{-2y}{(1+x^2+y^2)^2}$$

At the point $(x,y) = (1,1)$, we have $f_x(1,1) = \frac{-2}{9}$ and $f_y(1,1) = \frac{-2}{9}$. The normal is $\left(\frac{-2}{9}, \frac{-2}{9}, 1\right)$ and the point is $\left(1, 1, \frac{1}{3}\right)$. The tangent plane is

$$\frac{2}{9}x + \frac{2}{9}y + z = \frac{7}{9}.$$

At the point $(x,y) = (0,0)$, we have $f_x(0,0) = 0$ and $f_y(0,0) = 0$. The normal is $(0,0,1)$ and the point is $(0,0,1)$. The tangent plane is

$$z = 1.$$

At the point $(x,y) = (-2,2)$, we have $f_x(-2,2) = \frac{4}{81}$ and $f_y(-2,2) = \frac{-4}{81}$. The normal is $\left(\frac{-4}{81}, \frac{4}{81}, 1\right)$ and the point is $\left(-2, 2, \frac{1}{9}\right)$. The tangent plane is

$$\frac{-4}{81}x + \frac{4}{81}y + z = \frac{1}{9}.$$

The definition of tangent planes for $f : \mathbb{R}^2 \to \mathbb{R}$ extends to as many dimensions as we want. A function $f : \mathbb{R}^3 \to \mathbb{R}$ has a graph in $\mathbb{R}^4$. Its tangent spaces are 3-spaces in $\mathbb{R}^4$. We can understand those 3-spaces in a very similar method. We calcluate the three local tangent directions.

$$v_1 = \left(1, 0, 0, \frac{\partial f}{\partial x}\right) \qquad v_2 = \left(0, 1, 0, \frac{\partial f}{\partial y}\right) \qquad v_3 = \left(0, 0, 1, \frac{\partial f}{\partial z}\right)$$

There isn't a cross-product in $\mathbb{R}^4$, but we can genearlize the pattern in this case to get the normal to the tangent 3-space.

$$\left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, -\frac{\partial f}{\partial z}, 1\right)$$

The equation of the tangent 3-space at $(a,b,c,f(a,b,c))$ is

$$w - f(a,b,c) = \frac{\partial f}{\partial x}(a,b,c)(x-a) + \frac{\partial f}{\partial y}(a,b,c)(y-b) + \frac{\partial f}{\partial z}(a,b,c)(z-c)$$

And we could extend this to $f : \mathbb{R}^n \to \mathbb{R}$, which has a tangent hyperplane in $\mathbb{R}^{n+1}$.

$$v_1 = \left(1, 0, \ldots, \frac{\partial f}{\partial x_1}\right) \qquad v_2 = \left(0, 1, 0, \ldots, \frac{\partial f}{\partial x_2}\right) \qquad \ldots \qquad v_n = \left(0, \ldots, 0, 1, \frac{\partial f}{\partial x_n}\right)$$

The normal to the tangent hyperplane is

$$\left(-\frac{\partial f}{\partial x_1}, -\frac{\partial f}{\partial x_2}, \ldots, -\frac{\partial f}{\partial x_n}, 1\right).$$

The equation of the tangent hyperplane at $(a_1, a_2, a_3, \ldots, a_n, f(a_1, a_2, \ldots, a_n))$ is

$$x_0 - f(a_1, a_2, \ldots a_n) = \frac{\partial f}{\partial x_1}(a_1, a_2, \ldots a_n)(x_1 - a_2) + \frac{\partial f}{\partial x_2}(a_1, a_2, \ldots, a_n)(x_2 - a_2) + \ldots$$
$$\ldots + \frac{\partial f}{\partial x_n}(a_1, a_2, \ldots, a_n)(x_n - a_n).$$

To connect tangent (hyper)planes to tangents to parametric curves and derivatives along those curves, we have the following result.

Proposition 8.9.3. *Let $\gamma(t)$ be a parametric curve in $\mathbb{R}^{n+1}$ and $f : \mathbb{R}^n \to \mathbb{R}$ a differentiable function. Then if $\gamma(t)$ lies on the graph of f, the tangents to $\gamma(t)$ must lie on the tangent planes to the graph of f. (All these tangent vectors are local direction vectors).*

Example 8.9.4. Consider the same function as Example 8.9.2: $f(x, y) = \frac{1}{1+x^2+y^2}$. Then consider the parametric curves $\gamma_1(t) = \left(t, 1, \frac{1}{2+t^2}\right)$ and $\gamma_2(t) = \left(1, t, \frac{1}{2+t^2}\right)$. It is easy to check that both curves lie on the graph of f and both pass through the point $\left(1, 1, \frac{1}{3}\right)$ at $t = 1$. Then we can calculate the tangents to the curves at that point and the plane they span.

$$\begin{aligned}
\gamma_1'(t) &= \left(1, 0 \frac{-2t}{(2+t^2)^2}\right) & \gamma_2'(t) &= \left(0, 1 \frac{-2t}{(2+t^2)^2}\right) \\
\gamma_1'(1) &= \left(1, 0, \frac{-2}{9}\right) & \gamma_2'(1) &= \left(0, 1, \frac{-2}{9}\right) \\
\gamma_1'(1) \times \gamma_2'(1) &= \left(\frac{2}{9}, \frac{2}{9}, 1\right)
\end{aligned}$$

This gives exactly the same normal at the same point $\left(1, 1, \frac{1}{3}\right)$, so the same plane. We can think of tangent planes at the environment for tangents to curves which lie on the graph of the function.

8.10 Linear Approximation

Consider a single variable function $f : \mathbb{R} \to \mathbb{R}$. In Calculus I, we defined the linear approximation of f at the point $(a, f(a)$.

$$f(x) \approx f(a) + f'(a)(x - a)$$

From Calculus II or Chapter 3, you might recognize this as simply the first-order Taylor approximation for f, where we truncate after the linear term. The linear approximation is the line that best approximates f at this point. Its graph is simply the tangent line to f at $(a, f(a))$.

We can rearrange the linear approximaiton to help us generalize to multivariable functions.

$$(f(x) - f(a)) \approx f'(a)(x - a) \tag{8.10.1}$$

In this course, we have become accustomed to think of points in $\mathbb{R}^n$ along with local direction vectors: each point can be thought of as the origin for a system of local directions. We can do the same here: let $(a, f(a))$ be a local origin for a system of local directions. Then switching from $f(x)$ to $f(x) - f(a)$ and from x to $x - a$ is just moving from the usual origin to this new, local, origin.

In this new origin, the function in Equation 8.10.1 is approximated by multiplication by $f'(a)$. This gives us a new interpretation for the single-variable derivative: the derivative is the multiplicative factor for the local linear approximation of f. Locally, (in coordinates pretending that $(a, f(a))$ is the origin), the function is approximated by multiplication by $f'(a)$.

Now lets think of function $f : \mathbb{R}^2 \to \mathbb{R}$. It we want to approximate by linear functions, we need to understand linear functions: $\mathbb{R}^2 \to \mathbb{R}$. Following Section 4.13, we look to matrices. Matrices completely describe linear functions. A 1×1 matrix is just a number, and the linear function $\mathbb{R} \to \mathbb{R}$ is just multiplication by that number. But for $\mathbb{R}^2 \to \mathbb{R}$, we have a 1×2 matrix.

So we ask: what is a linear approximation to the function $f : \mathbb{R}^2 \to \mathbb{R}$. It must be a 1×2 matrix M and fit into a equation mirroring Equation 8.10.1.

$$f(x, y) - f(a, b) \approx M \begin{pmatrix} x - a \\ y - b \end{pmatrix}$$

In local coordinates at $(a, b, f(a, b))$, this is just matrix multiplication by M. So a linear approximation is a matrix multiplication in local coordintaes. What is the matrix M? Well, if we work with partial derivatives, the linear approximation should be formed of linear approximation in x and linear approximation in y.

$$f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial x}(a, b)(y - b)$$

We put this in matrix form.

$$f(x, y) - f(a, b) \approx \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} x - a \\ y - b \end{pmatrix}$$

Definition 8.10.1. The matrix of the linear approximation to a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ is the $1 \times n$ matrix of partial derivatives.

The graph of the linear approximation is the tangent plane that we've already defined at $(a, b, f(a, b))$.

Example 8.10.2. Let's return to the function in Example 8.9.2, $f(x, y) = \frac{1}{1+x^2+y^2}$.

$$M = \begin{pmatrix} \frac{-2x}{(1+x^2+y^2)^2} & \frac{-2y}{(1+x^2+y^2)^2} \end{pmatrix}$$

Look at the point $(0, 0)$.

$$f(x, y) \approx f(0, 0) + \begin{pmatrix} 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1$$

This linear approximation is a constant 1, which makes sense at the top of the small hill. Momentarily, at the peak, nothing is changing and the function doesn't do anything. The linear approximation to doing nothing is appropriately a constant.

Look at the point $(1, 1)$.

$$f(x, y) \approx f(1, 1) + \begin{pmatrix} \frac{-2}{9} & \frac{-2}{9} \end{pmatrix} \begin{pmatrix} x-1 \\ y-1 \end{pmatrix} = \frac{1}{3} - \frac{2(x-1)}{9} - \frac{2(y-1)}{9}$$

Look at the point $(-2, 2)$.

$$f(x, y) \approx f(-2, 2) + \begin{pmatrix} \frac{4}{81} & \frac{-4}{81} \end{pmatrix} \begin{pmatrix} x+2 \\ y-1 \end{pmatrix} = \frac{1}{9} + \frac{4(x+2)}{81} - \frac{4(y-2)}{81}$$

Example 8.10.3. $f(x, y) = x^2 e^{x-y}$.

$$M = \begin{pmatrix} 2xe^{x-y} + x^2 e^x, -x^2 e^{x-y} \end{pmatrix}$$

Look at the point $(2, 2)$.

$$f(x, y) \approx f(2, 2) + \begin{pmatrix} 4 + 4e^2 & -4 \end{pmatrix} \begin{pmatrix} x-2 \\ y-2 \end{pmatrix} = 4 + (4 + 4e^2)(x-2) - 4(y-2)$$

Look at the point $(-1, -1)$.

$$f(x, y) \approx f(-1, -1) + \begin{pmatrix} (2+e) & -1 \end{pmatrix} \begin{pmatrix} x+1 \\ y+1 \end{pmatrix} = 1 + (2+e)(x+1) - (y+1)$$

We extended the idea of a derivative with partial derivatives and directional derivatives. These are both useful, but neither are a universal extension. Both are only pieces of derivatives along certain directions. Gradients were a more universal extension, but they only identify the direction of greatest change. Tangent planes were a good geometric extension, but without an algebraic analogue.

The discussion of linear approximation leads us to a much more universal idea to generalize the derivative. The big idea is this: derivatives, in any dimension, are linear approximations to functions. This idea works in single variables, where multiplication by $f'(a)$ was the linear approximation. The derivatives calculates that factor. For higher dimensions, we can use matrix multiplication instead of just multiplication by a number.

Therefore, we could realistically say that *the derivative* of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is the *matrix of partial derivatives*, which serves as a linear approximation of the function at any point in its domain. It will be useful to give this matrix a name, for future reference.

Definition 8.10.4. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function. Then the $m \times n$ matrix of partial derivatives of f is called the *Jacobian Matrix* of the function.

Chapter 9

Extrema and Optimization

In Calculus I, one of the first applications of derivatives was optimization. The same is true in this course. In this chapter, we'll build the techniques that let us find the maximum and minimum values of multivariable funtions.

9.1 Extrema

Let's first define what we mean by extreme values.

Definition 9.1.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar function. Then a point $p \in \mathbb{R}^n$ is a local maximum of f is there exists $\epsilon > 0$ such that $f(p) \geq f(q)$ for all $q \in B(p, \epsilon)$. Similarly, a point $p \in \mathbb{R}^n$ is a local minimum of f is there exists $\epsilon > 0$ such that $f(p) \leq f(q)$ for all $q \in B(p, \epsilon)$.

This definition clarifies that if a point is a minimum or maximum, is it is a peak or valley in all directions. It needs to be above or below nearby function values in any direction in the domain. In two variables, we also want to classify a new kind of extreme value.

Definition 9.1.2. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. Then a point $p \in \mathbb{R}^n$ is a saddle point of f if there there are two unit directions u and v in $\mathbb{R}^n$ and $\epsilon > 0$ such that $f(p \pm \delta u) \geq f(p)$ and $f(p \pm \delta v) \geq f(q)$ for all $\delta < \epsilon$.

A saddle point is both a minimum and a maximum: it is a minimum in some direction v and a maximum in some other direction u. It is called a saddle point for the saddle-like shape that results from this situation for graphs of two-variable funcitons. For higher dimensions, a saddle point is any point which is a maximum is some number of directions and a minimum is all other direction (for some linearly independent set of directions in the domain).

A key obervation from Calculus I is that maxima and minima where found when $f'(x) = 0$. (Though $f' = 0$ didn't guarantee an exterme value, as in the example of $f(x) = x^3$ at $x = 0$).

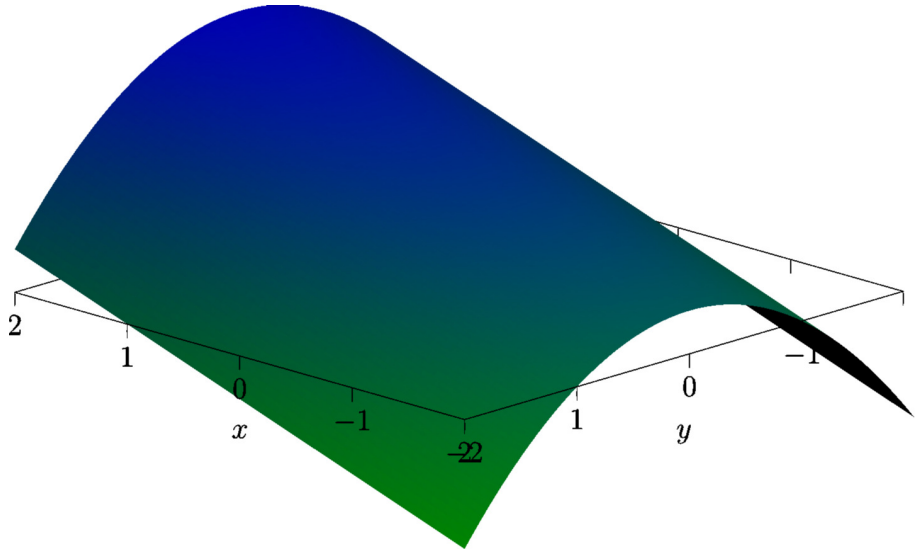

Figure 9.1: The function $f(x, y) = x - y^2 + 3$.

Proposition 9.1.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differential function. If f has a minimum, maximum or saddle point at $p \in \mathbb{R}^n$ and $\nabla f(p) = 0$.*

The gradient measures the direction of greatest change. At a minimum, maximum or saddle point, there is no such direction, so the gradient is zero. As with single-variable calculus, the implication is only one direction. As with single variable function, there may be points where the gradient is zero but the point is neither a minimum, maximum nor saddle point.

The gradient is the vector of partial derivatives, so it is important to note that *all* the partial derivative must vanish. If only some of them vanish, we get interesting behaviour, but not maxima or minima.

Example 9.1.4. For example, consider the function $f(x, y) = x - y^2 + 3$, shown in Figure 9.1. We have $f_y(x, y) = -2y$, which vanishes everywhere along the x axis (when $y = 0$). However, the other partial is $f_x = 1$, which never vanishes. This means that all points $(x, 0)$ are potentiall critical in y but not in x. What you get with this function is a ascending/decending ridge (with slope 1) above the x axis.

Example 9.1.5. Since we work in several dimensions, we can have very complicated sets of maxima/minima. The function $f(x, y) = \cos(x + y)$, shown in Figure 9.2 has a maximum whenever $x + y$ is an even multiple of π. Each of those sets is a whole line, $x + y = 0$, $x + y = 2\pi$ and so on. For functions of two variables, its easy to have lines or curves of maximum or minimum points. In higher dimensions, we can have surfaces or hypersurfaces of maximum or minimum points.

We want to classify critical points: points where $\nabla f = 0$. We can do this informally by looking at each variable individually. If the point is a maximum in all of $x_1, x_2, \ldots, x_n$, then it is a maximum

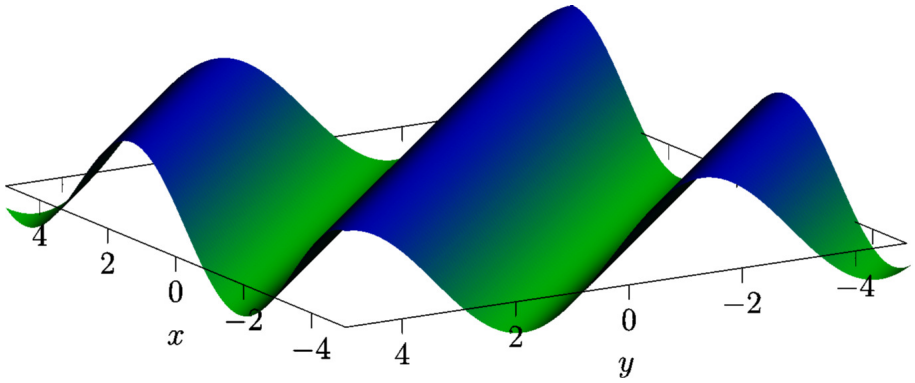

Figure 9.2: The function $f(x, y) = \cos(x + y)$

according to our definition above. Likewise, if it is a minimum in all the variables, it is a minimum. If it is a maxmum in some of the variables and a minimum in others, it is a saddle point. However, if even one of the variable has no max/min behvaiour (like $f(x) = x^3$ at $x = 0$), then the point is neither a maximum, minimum nor saddle point.

9.1.1 Hessian Matrices

This informal approach is reasonable, but it would be nice to have a more formal method for determining the behaviour. In single variable calculus, we have the second-derivative test. If a is a critical point, then it is a maximum if $f''(a)$ is negative, a minimum if $f''(a)$ is positive, and the test is inconclusive if $f''(a) = 0$.

We'd like to generalize this, but we have many second derivatives: all of the possible mixed and non-mixed second partials. One way we can organize all these second partials is in a matrix.

Definition 9.1.6. The matrix of all the second partial derivatives of a scalar function f is called the *Hessian Matrix*.

Here is the Hessian matrix for $f(x, y)$ in two variables.

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial y \partial x} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Here is the Hessian matrix for $f(x, y, z)$ in three variables.

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}$$

Note that the Hessian is not the Jacobian matrix from before; that matrix had only first derivatives. The Hessian matrix only applies to single valued function (outputs in $\mathbb{R}$), is always square, and lists all the possible second partials.

The Hessian matrix captures all of the information about the second derivative of this function, but it is often too unwieldy to be used to determine the behaviour of critical points. However, we have a useful tool from linear algebra to get specific information out of a matrix: the determinant. Let D be the determinant of the Hessian matrix. For $f(x, y)$, D has the following form (using Clairaut's theorm to simplify the mixed partials).

$$D = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2$$

For functions of two variables, the determinant of the Hessian tells us most of what we need to know.

Proposition 9.1.7. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^2 function and let (a, b) be a critical point. Then we have four cases.*

- *If $D(a, b) > 0$ and $\frac{\partial^2 f}{\partial x^2}(a, b) > 0$ then the critical point is a minimum.*
- *If $D(a, b) > 0$ and $\frac{\partial^2 f}{\partial x^2}(a, b) < 0$ then the critical point is a maximum.*
- *If $D(a, b) < 0$ then the critical point is a saddle point.*
- *If $D(a, b) = 0$ then the test is inconclusive.*

This proposition can be generalized to higher dimensions, but it requires more machinery from linear algebra, namely eigenvalues. Clairaut's theorem means that the Hessian is always a symmetric matrix, so it always has a maximal set of real eigenvalues. The general proposition classifies the extrema using those eigenvalues.

Proposition 9.1.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function with H its Hessian matrix. Let $u \in \mathbb{R}^n$ be a critical point and let $H(u)$ be the Hessian evaluated at the point u. Then we have four cases.*

- *If $H(u)$ is not invertible (has determinant 0, has 0 as an eigenvalue), then the test is inconclusive.*
- *If all the eigenvalues of $H(u)$ are positive, then the critical point is a local minimum.*
- *If all the eigenvalues of $H(u)$ are negative, then the critical point is a local maximum.*
- *If the eigenvalues of $H(u)$ are a mix of positive and negative, then the critical point is a saddle point.*

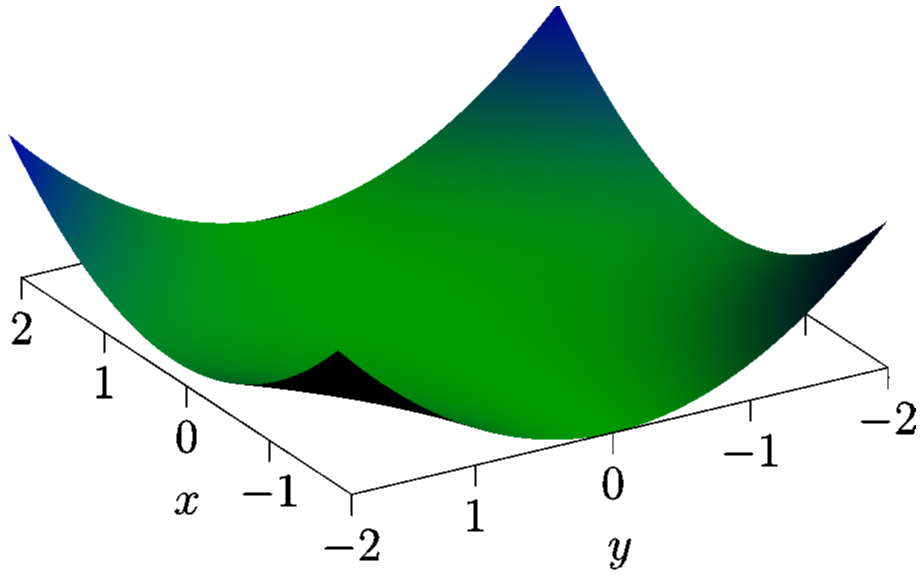

Figure 9.3: The function $f(x,y) = y^2x^2 + y^2$.

9.1.2 Global Extrema

The above analysis was all about local maxima and minima. We can also ask for global maxima and minima. Like the single variable case, we do this by looking at all the local extrema as well as the boundary. The maximum or minimum might be a boundary point which is not a critical point. We do have an existence proposition, which relies on the topology of the domain.

Proposition 9.1.9. *Let C be closed set in $\mathbb{R}^n$ and $f : C \to R$. Then f has at least one global maximum and at least one global minimum, either at a local maximum/minimum or on its boundary.*

In general, finding the maximum and minimum on the boundary can be quite difficult. In the next section, we'll give another technique for finding such maxima and minima.

Example 9.1.10. $f(x,y) = y^2x^2 + y^2$ is shown in Figure 9.3.

$$\frac{\partial f}{\partial x} = 2xy^2$$
$$\frac{\partial f}{\partial y} = 2yx^2 + 2y$$
$$\nabla f(x,y) = 0 \implies (x,y) = (a,0) \qquad \forall a \in \mathbb{R}$$

$$\frac{\partial^2 f}{\partial x^2} = 2y \qquad \frac{\partial^2 f}{\partial y^2} = 2x^2 + 2 \qquad \frac{\partial^2 f}{\partial x \partial y} = 4xy$$
$$D = (2y)(2x^2+2) - 16x^2y^2 = 4yx^2 + 4y - 16x^2y^2 \implies D(a,0) = 0$$

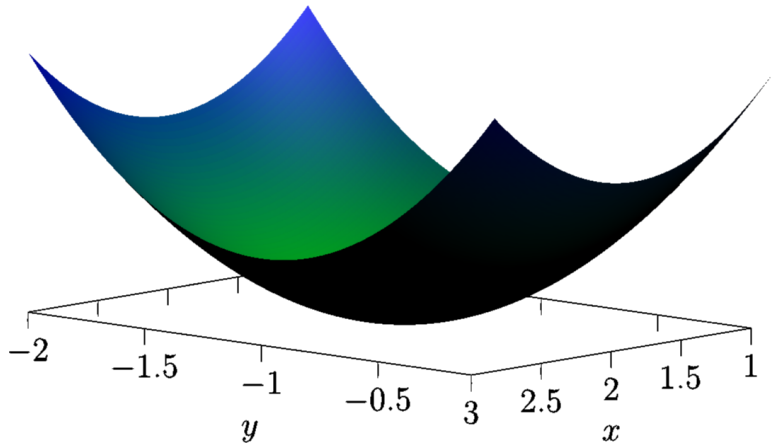

Figure 9.4: The function $f(x, y) = x^2 + 2y^2 - 4x + 4y + 6$.

At all the critical points $(a, 0)$, the second derivative test has $D = 0$, which is inconclusive. We have to investigate directly. We can see that $f(a, 0) = 0$. But $f(a, b)$ for b any small non-zero nubmer takes the value $a^2b^2 + b^2$, which is always positive. Therefore, we can conclude that all the critical points $(a, 0)$ are local minima. In Figure 9.3, we can see that all along the y axis the values stay at 0, which is the lowest output of the function.

Example 9.1.11. $f(x, y) = x^2 + 2y^2 - 4x + 4y + 6$ is shown in Figure 9.4.

$$\begin{aligned}
\frac{\partial f}{\partial x} &= 2x - 4 \\
\frac{\partial f}{\partial y} &= 4y + 4 \\
\nabla f(x, y) = 0 &\implies (x, y) = (2, -1) \\
\frac{\partial^2 f}{\partial x^2} &= 2 > 0 \\
\frac{\partial^2 f}{\partial y^2} &= 4 > 0 \\
\frac{\partial^2 f}{\partial x \partial y} &= 0 \\
D &= 8 > 0
\end{aligned}$$

The point $(2, -1)$ is a local minimum.

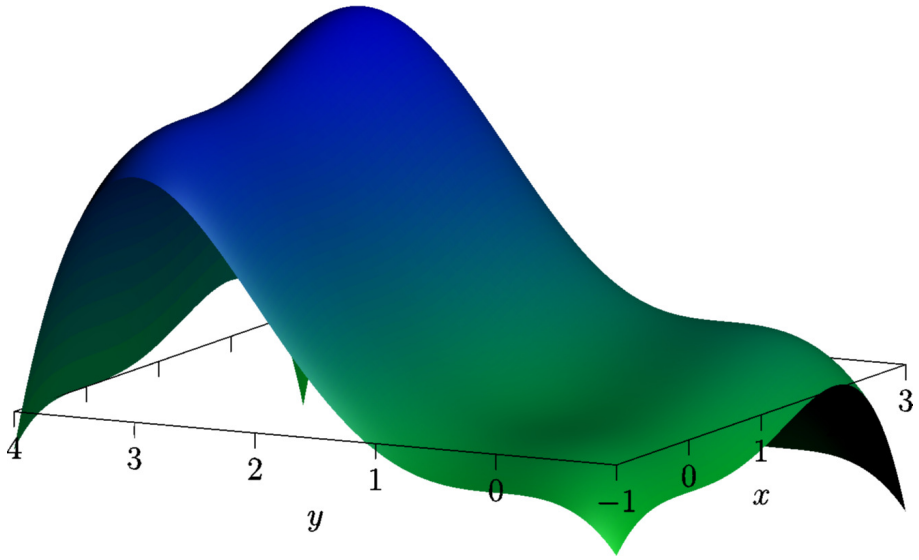

Figure 9.5: The function $f(x,y) = \frac{8}{3}x^3 + 4y^3 - x^4 - y^4$.

Example 9.1.12. $f(x,y) = \frac{8}{3}x^3 + 4y^3 - x^4 - y^4$ is shown in Figure 9.5

$$\begin{aligned}
\frac{\partial f}{\partial x} &= 8x^2 - 4x^3 = 4x^2(2-x) \\
\frac{\partial f}{\partial y} &= 12y^2 - 4y^3 = 4y^2(3-y) \\
\nabla f(x,y) = 0 &\implies (x,y) = (0,0),(0,3),(2,0),(2,3) \\
\frac{\partial^2 f}{\partial x^2} &= 16x - 12x^2 = 3x(4-3x) \\
\frac{\partial^2 f}{\partial y^2} &= 24y - 12y^2 = 12(2-y) \\
\frac{\partial^2 f}{\partial x \partial y} &= 0 \\
D &= 48xy(4-3x)(2-y)
\end{aligned}$$

$$D(0,0) = 0 \qquad D(0,3) = 0 \qquad D(2,0) = 0 \qquad D(2,3) > 0 \qquad \frac{\partial^2 f}{\partial x^2}(2,3) < 0$$

The point $(2,3)$ is a local maximum, but the test is inconclusive at all other points. As we can see in Figure 9.5, there is a maximum above $(2,3)$. At the other three critical points, we can observe a momentary flattening of the graph, but we have neither maxima, minima, nor saddle points. The zero Hessian for these points makes sense; they are like the critical point at $x = 0$ of the cubic $f(x) = x^3$, which has a zero derivative, but no extremum.

Example 9.1.13. Let's try a more geometric optimization problem: what is the largest rectangular prism you can fit in a sphere of radius r? We'll assume that the prism is centrally located in the sphere,

which means that its shape is entirely determined by one of its vertices on the edge of the sphere. If that vertex is (x, y, z), then the volume of the prism is $2x \times 2y \times 2z$.

I'd like to work in spherical coordinates instead of x, y and z. The radius r is fixed, but θ (longitude) and ϕ (colatitude) will vary.

$$\begin{aligned} h &= 2r\cos\phi \\ w &= 2r\sin\phi\cos\theta \\ l &= 2r\sin\phi\sin\theta \\ V &= hwl = 8r^3\cos\phi\sin^2\phi\cos\theta\sin\theta = 4r^3\sin(2\theta)(\cos\phi - \cos^3\phi) \end{aligned}$$

Then we can optimize the function $V(\theta, \phi)$.

$$\begin{aligned} \frac{\partial V}{\partial \phi} &= 4r^3\sin(2\theta)(-\sin\phi + 3\cos^2\phi\sin\phi) \\ \frac{\partial V}{\partial \theta} &= 8r^3\cos 2\theta(\cos\phi - \cos^3\phi) \\ \nabla V = 0 &\implies (\phi, \theta) = \left(\arccos\frac{1}{\sqrt{3}}, \frac{\pi}{4}\right) \\ V &= 4t^3\sin\frac{\pi}{2}\left(\frac{1}{\sqrt{3}} - \frac{1}{3\sqrt{3}}\right) = \frac{8r^3}{3\sqrt{3}} \end{aligned}$$

We didn't do the calcluation, but it is reasonable to check that the critical point represents a maximum. The area we get is the area of cube of side length $\frac{2r}{\sqrt{3}}$, which seems, intuitively, like the right kind of length.

9.2 Constrained Optimization

Let's reconsider the box-in-a-sphere problem of Example 9.1.13. Instead of working inside the sphere, we could just consider a prism centred at the origin with positive octant vertex (a, b, c) and volume $8abc$. That maximized by itself will diverge to ∞, since we have nothing to bound the prism. However, we can consider the *constraint* $x^2 + y^2 + z^2 = r^2$ for some fixed r.

The idea, in general, is to ask when a function $f(x_1, x_2, \ldots, x_n)$ is optimized subject to one (or more) constraints $g(x_1, x_2, \ldots x_n) = c$. Let's first think of this geometrically. If both f and g were functions of two variables, we could draw contour plots. $g(x, y) = c$ would be a fixed countour line for g, and we would be asking for maximum of f along that contour line. We can notice that such a maximum occurs where a countour line for f touches the contour line $g(x, y) = c$ and then backs off. That is, the maximum occurs where the contour lines are tangent.

But the contour lines are level sets, which are defined by their normals. The normals are given by gradients. If the level sets are tangent to each other, that means the gradients must be in the same

direction (up to a sign). That is, this optimization occurs when $\nabla f \parallel \nabla g$. We can express this as $\nabla f = \lambda \nabla g$ for some real constant $\lambda \neq 0$.

This holds in any dinemsion: the level sets will be tangent and there will be a constant λ with $\nabla f = \lambda \nabla g$. To optimize with constraint, we now have a system of equation in $n+1$ variables, the x_i and λ: $\nabla f = \lambda \nabla g$ and $g(x_1, x_2, \ldots, v_n) = c$. To optimize the situation, we try to solve this system. This method is called the method of Lagrange Multipliers.

(Historically, the connection to the mathematician Lagrange comes from a *Lagrangian* of the form $f + \lambda g$. We're essentially looking for critical points of this Lagrangian).

Example 9.2.1. What is the largest rectangle with fixed perimeter p? The rectangle has length l and height h, with area $A(h,l) = hl$. The constraint is the fixed perimeter $2h + 2l = p$, so the function is $g(h,l) = 2h + 2l$. The gradients are $\nabla A = (l, h)$ and $\nabla g = (2, 2)$. The equation $\nabla A = \lambda \nabla g$ is two equations: $h = 2\lambda$ and $l = 2\lambda$. We can substitute these into the constrait $2h + 2l = p$ to get $8\lambda = p$ so $\lambda = p/8$ and $h = l = 2\lambda = p/4$. The area is then $p^2/16$.

Example 9.2.2. Let's actually do the motivating Example 9.1.13 with Lagrange Multipliers.

$$V(x,y,z) = 8xyz \qquad g(x,y,z) = x^2 + y^2 + z^2 = r^2$$

$$\nabla V = (8yx, 8xz, 8xy) \qquad \nabla g = (2x, 2y, 2z)$$

$$\nabla V = \lambda \nabla g$$

$$8yz = \lambda 2x \qquad 8xz = \lambda 2y \qquad 8xy = \lambda 2z$$

We will use the first of the three equations to isolate for x, which we substitute in the second.

$$x = \frac{1}{\lambda} 4yz \qquad \frac{8}{\lambda} 4yz^2 = \lambda 2y$$

We solve for z (cancelling y, so careful with the possibility of $y = 0$.)

$$16z^2 = \lambda^2 \implies z = \frac{\lambda}{4}$$
$$\frac{8}{\lambda} 4y^2 z = \lambda 2z$$

Since everything in this question is symmetric, we apply the same reasoning with different variables.

$$y = \frac{\lambda}{4} \qquad x = \frac{\lambda}{4}$$

We put x, y and z into the constraint to solve for λ.

$$x^2 + y^2 + z^2 = \frac{3\lambda^2}{16} = r^2 \implies \lambda = \frac{4r}{\sqrt{3}}$$

Then we use the value of γ to give the values of the variables.

$$x = y = z = \frac{r}{\sqrt{3}}$$
$$V = \frac{8r^3}{3\sqrt{3}}$$

We recovered the same maximal area as before. We don't have to worry about the $y = 0$ possibility that we noted (or likewise $x = 0$ and $z = 0$), since any solution in that case would lead to zero volume, which is certainly not maximal.

Example 9.2.3. Instead of a sphere, we could ask for the largest box inside an ellipsoid with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Geometrically, this because more difficult. However, using Lagrange Multipliers, we only need to change the constraint from the equation of the sphere to the equation of the ellipse.

$$V(x, y, z) = 8xyz \qquad g(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = r^2$$
$$\nabla V = (8yx, 8xz, 8xy) \qquad \nabla g = \left(\frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2}\right)$$
$$\nabla V = \lambda \nabla g$$
$$8yz = \lambda \frac{2x}{a^2} \qquad 8xz = \lambda \frac{2y}{b^2} \qquad 8xy = \lambda \frac{2z}{c^2}$$

We isolate for x in the first expression and substitute in the second.

$$x = \frac{4z^2yz}{\lambda} \qquad 8\left(\frac{4a^2yz}{\lambda}\right) z = \frac{\lambda 2y}{b^2}$$

We solve for z (again, careful with $y = 0$).

$$16a^2b^2z^2 = \lambda^2 \implies 4abz = \lambda \implies z = \frac{\lambda}{4ab}$$

We have symmetry again in the variables.

$$x = \frac{\lambda}{4bc} \qquad y = \frac{\lambda}{4ac}$$

We substitute x, y and z in the constraint and solve for λ.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 3\left(\frac{\lambda^2}{16a^2b^2c^2}\right) = 1 \implies \lambda = \frac{4abc}{\sqrt{3}}$$

We use the value of λ to determine the other variables.

$$x = \frac{a}{\sqrt{3}} \qquad y = \frac{b}{\sqrt{3}} \qquad z = \frac{c}{\sqrt{3}}$$
$$V = \frac{8abc}{3\sqrt{3}}$$

The volume of the ellipsoid is $\frac{4}{3}\pi abc$, so as a fraction of that volume, this is $\frac{2}{\sqrt{3}\pi}$ times the volume of the ellipse, a very believable fraction. Again, the $x = 0$, $y = 0$ or $z = 0$ cases are discounted since they would lead to zero area and we are looking for a maximum.

Example 9.2.4. We can use Lagrange multipliers with several constraints for higher dimensional problems. Consider a plane $x+y+z=12$ and a parabaloid $z=x^2+y^2$. Their intersection is an ellipse in $\mathbb{R}^3$. I'd like to find its highest point and its lowest point, thinking of z as elevation. In Lagrange Multipliers, rhe function $f(x,y,z)=z$ is simply height, and the two equations are the constraints.

$$\begin{array}{lll} f(x,y,z)=z & g(x,y,z)=x+y+z=11 & h(x,y,z)=x^2+y^2-z=0 \\ \nabla f=(0,0,1) & \nabla g=(1,1,1) & \nabla h=(2x,2y,-1) \\ \nabla f=\lambda\nabla g+\mu\nabla h & & \\ 0=\lambda+2x\mu & 0=\lambda+2y\mu & 1=\lambda-\mu \end{array}$$

This system is relatively straightforward, even with five variables. We can solve for λ in the first two equation and equate the results.

$$2x\mu=\lambda=2y\mu \implies x=y \quad (\mu\neq 0)$$

We put x and y into the h constraint.

$$z=2x^2$$

Then we put this expressionv for z, and replace y with x in the g constraint. This result is an expression solely in x, so we can solve for values of x.

$$2x+2x^2=12 \implies x=-3 \text{ or } x=2$$

We know the other variables in terms of x, so values of x determine values of y and z.

$$\begin{aligned} x=-3 &\implies y=-3 \;\; z=18 \\ x=2 &\implies y=2 \;\; z=8 \end{aligned}$$

We can simply look at the z value, since we are trying to maximize it. We see a maximum at $(-3,-3,18)$ and the minimum at $(2,2,8)$.

Example 9.2.5. As we've seen in these examples, constrained optimization is a particularly useful tool for sovling complicated geometry problems. Here is another such problem. Say we have the equation of a very general ellipse centred at the origin, such as $17x^2+12xy+8y^2=100$. Since this ellipse has an xy term, it is not oriented along any axis – it's on some slanted plane through the origin. So we don't know, from inspection, what is major and minor semi-axes are. However, we can ask for the points further from and closest too the origin. That's asking for the min and max of the funciton $f(x,y)=x^2+y^2$ on the constraint which is the equation of the ellipse.

$$\begin{array}{ll} f(x,y)=x^2+y^2 & g(x,y)=17x^2+12xy+8y^2=100 \\ \nabla f=(2x,2y) & \nabla g=(34x+12y,12x+16y) \\ \nabla f=\lambda\nabla g & \\ 2x=\lambda(34x+12y) \quad 2y=\lambda(12x+16y) & \end{array}$$

We solve for λ in both equation and equate the results.

$$\begin{aligned} \frac{-2x}{34x+12y}=\lambda=\frac{-2y}{12x+16y} \\ 12x^2+16xy=34xy+12y^2 \end{aligned}$$

This is a bit tricky, since there are quadratic terms in both x and y. Turns out, though, we can get use out of isolating y^2.

$$2x^2 - 3xy - 2y^2 = 0 \implies y^2 = x^2 - \frac{3}{2}xy$$

We replace y^2 in the constraint. We are fortunate, since this replacement cancelled off the xy term and leaves us just with an expression in x.

$$\begin{aligned} 12x^2 + 12xy + 8\left(x^2 - \frac{3}{2}xy\right) &= 100 \\ 12x^2 + 8x^2 = 100 &\implies x = \pm 2 \\ x = 2 &\implies 8 - 6y - 2y^2 = 0 \implies y = 1, -4 \\ x = -2 &\implies 8 + 6y - 2y^2 = 0 \implies y = -1, 4 \end{aligned}$$

This gives us the four points $(2, 1)$, $(-2, -1)$, $(2, -4)$ and $(-2, 4)$. The semi-major axis is the distance to the further point, and the semi-minor axis is the distance to the closet point (from the origin). Therefore, the semi-major axis is $\sqrt{20}$ units long along the line in the direction $(2, -4)$ and the semi-minor axis is $\sqrt{5}$ units long in the direction $(2, 1)$.