

# §. Preface

Hello! Welcome to MATH 1ZB3, MATH 1AA3, or whatever course.

I'm Jason Cheng, a second-year (as of time of writing) computer engineering student. This study guide contains, as of time of writing, the whole content of the courses above, and more. It is cited straight from the textbook, everything important you need to know. The textbook used is **Calculus: Early Transcendentals (2020)**. I recommend getting the textbook even if you use this book: it contains tons of examples and diagrams that will help you beyond words can. However, if you can't pay, just use a [SHADOW LIBRARY](#).

If you're wondering who I am to be making this guide, my grades for the first two MATH 1ZB3 tests should help (the final exam hasn't happened yet as of writing).

Test #1 Mark: 18/18 (= 100%)

Test #1 Version: 1

Answers for T1:

Your Answers									
1. D	2. B	3. D	4. C	5. A	6. C	7. A	8. B	9. C	10. E
11. B	12. A	13. D	14. A	15. A	16. A	17. D	18. C		
Version 1 Answer Key									
1. D	2. B	3. D	4. C	5. A	6. C	7. A	8. Bcd	9. Cab	10. E
11. Bcd	12. A	13. D	14. Acd	15. Acd	16. Acd	17. D	18. Cad		

Test #1 Class Average: 64.62%

Test Questions and Answers: If you lost marks on the test then you can see where you went wrong by viewing the tests, and the answers to the tests here: <https://mcmasteru365.sharepoint.com/:f:/r/sites/course-695914-group/Shared%20Documents/General/Test%20Copies%20and%20Keys?csf=1>

Test #2 Mark: 15/17 (= 88.24%)

Test #2 Version: 1

Answers for T2:

Your Answers									
1. C	2. D	3. A	4. B	5. B	6. E	7. C	8. A	9. A	10. D
11. E	12. B	13. C	14. B	15. D	16. A	17. C	18. E		
Version 1 Answer Key									
1. C	2. D	3. A	4. B	5. B	6. E	7. C	8. A	9. A	10. D
11. B	12. B	13. C	14. E	15. D	16. Cab	17. Bcd	18. E		

Test #2 Class Average: 53.60%

(Yes, I know I lowkey fumbled on Test 2).

Flexing out of the way, I suggest checking in the **outline** of whatever PDF reader you are using, and you will see each chapter, subchapter, and sub-subchapter there. Click to skip to specific chapters or parts.

I have included here explanations of my own to help you through specific chapters and concepts that may be difficult. Also note that while I have used large headings to denote content of each test, your course schedule will be different every year- expect some content to be off by a bit, some content to be taught out of order, even some chapters to be skipped entirely.

On the subject of skipped chapters, several subchapters are not part of your course curriculum. You are responsible for knowing which ones you need to know, but I have included warnings specific to MATH 1ZB3 at each problematic chapter.

Last thing: if you are here from my [LINEAR ALGEBRA STUDY GUIDE](#), just know that at this point in my life, I do not feel entirely great about its contents. Yes, it is a very helpful resource, and yes, I know students that have scored 10+s using it (*I guide others to a treasure I cannot possess*), but I think it is not my finest work. It is messy, amateurish, and very barebones. Fun fact: the LASG is 64 pages. This is 70 (on A4 paper), all while covering only course content. I think that speaks to how much harder I'm trying to help you understand the content.

This time, I have a better system in place.

- Any block math, that is any math in its own paragraph block and surrounded by an outline, for example

Math puns are the first sine of madness.

is an important theorem straight from the textbook, one you are expected to know and understand.

- Any block math, for example

$$|\text{You}(x, y)\rangle = \sqrt{p} |:\rangle + \sqrt{1-p} |:(\rangle$$

$$p \rightarrow 1 \text{ as } (x, y) \rightarrow (x_{\text{bed}}, y_{\text{bed}}) \text{ under observation}$$

is considered 'important', as in part of a proof for a theorem, an important byproduct, or an important theorem that isn't from the textbook. (if you don't understand this example, then ask ChatGPT or your nearest physicist or something. Trust me, it's funny)

- Any inline math, for example  $9 + 10 = \int_0^1 126x^5 dx$ , is used for formatting or anything less important.

As usual, `.md` source code is posted in the [SAME PLACE YOU GOT THIS PDF](#) (if you got it elsewhere, I'm not mad at you :D).

## 0. Review

These chapters should be review from MATH 1ZA3 (or equivalent). You should know everything within by heart.

### 7.5: Strategies for Integration

- $\int x^n dx = \frac{x^{n+1}}{n+1} \quad (n \neq -1)$
- $\int \frac{1}{x} dx = \ln |x|$
- $\int e^x dx = e^x$
- $\int b^x dx = \frac{b^x}{\ln b}$
- $\int \sin x dx = -\cos x$
- $\int \cos x dx = \sin x$
- $\int \sec^2 x dx = \tan x$
- $\int \csc^2 x dx = -\cot x$
- $\int \sec x \tan x dx = \sec x$
- $\int \csc x \cot x dx = -\csc x$
- $\int \sec x dx = \ln |\sec x + \tan x|$
- $\int \csc x dx = \ln |\csc x - \cot x|$
- $\int \tan x dx = \ln |\sec x|$
- $\int \cot x dx = \ln |\sin x|$
- $\int \sinh x dx = \cosh x$
- $\int \cosh x dx = \sinh x$
- $\int \frac{dx}{x^2+a^2} = \frac{1}{a} \tan^{-1} \left( \frac{x}{a} \right)$
- $\int \frac{dx}{\sqrt{a^2-x^2}} = \sin^{-1} \left( \frac{x}{a} \right); \quad a > 0$
- $\int \frac{dx}{x^2-a^2} = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right|$  \*This one is easily avoided by using partial fractions.
- $\int \frac{dx}{\sqrt{x^2+a^2}} = \ln \left| x + \sqrt{x^2+a^2} \right|$  \*This one can be avoided with trig. substitution.

Tips for integration:

- Manipulate integrand.
  - Expand. Simplify. Factor. Substitute. Rewrite radicals as powers.
  - Try to work it into a form with an obvious solution.
- Classify according to form.
  - Trig Functions
    - If the function is made up entirely of trig functions, like  $f(x) = \sin^n x \cos^m x$ , consider a **trigonometric substitution**.
  - Rational Functions
    - If the function is a rational function, like  $f(x) = \frac{A+Bx}{C+Dx}$ , consider **partial fraction decomposition**.
  - Radicals
    - For the common identities  $\sqrt{x^2+a^2}$ ,  $\sqrt{x^2-a^2}$ ,  $\sqrt{a^2-x^2}$ , use a **trigonometric substitution**. For radicals of form  $\sqrt[n]{g(x)}$ , use a **substitution**.
- If all else fails, try the above again.
  - There are two main methods of integration: **substitution** and **parts**. Try to make any substitution. Remember that you can use the differential term ( $dx$ ,  $dy$ ,  $dt$ , etc.) as the second term for integration by parts.
    - Try algebraic manipulation.

- You can split terms separated by  $+/−$  into multiple integrals: try to coax the integral into this form.
- Try multiplying the integral by a term that cancels to zero: for example, multiplying by  $\frac{\cos x + 1}{\cos x + 1}$  to create a trigonometric identity.
- If you don't know how to already, learn how to **rationalize denominators**.

# 1. Series, Sequences, and Infinity

This 'section' of the course will span from around its start to the first midterm.

## 7.8: Improper Integrals

**Improper integrals** are definite integrals with an infinite interval or a discontinuous integrand. We will refer to the former as **Type 1** and the latter as **Type 2**.

### 7.8.1: Integrals with Infinite Intervals

This is **Type 1**. Integrals of this type will appear as:

$$\int_b^a f(x) dx$$

where either  $a$  or  $b$  are  $\infty$ .

In order to solve such an integral, we abstract the infinite end to a variable, then take the integral's limit as that variable approaches infinity. Thus,

$$\int_b^\infty f(x) dx = \lim_{a \rightarrow \infty} \int_b^a f(x) dx$$

and similar for integrals bounded from  $-\infty$ .

If the limit exists, we call that integral **convergent**. If not, we call it **divergent**.

Formally, we define

(a) If  $\int_a^t f(x) dx$  exists for every number  $t \geq a$ , then

$$\int_a^\infty f(x) dx = \lim_{t \rightarrow \infty} \int_a^t f(x) dx$$

provided this limit exists (as a finite number).

(b) If  $\int_t^b f(x) dx$  exists for every number  $t \leq b$ , then

$$\int_{-\infty}^b f(x) dx = \lim_{t \rightarrow -\infty} \int_t^b f(x) dx$$

provided this limit exists (as a finite number).

(c) If both  $\int_a^\infty f(x) dx$  and  $\int_{-\infty}^a f(x) dx$  are convergent, then we define

$$\int_{-\infty}^\infty f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^\infty f(x) dx$$

The following important fact is also presented:

$$\int_1^\infty \frac{1}{x^p} dx \text{ is convergent if } p > 1 \text{ and divergent if } p \leq 1.$$

We call a function of this kind a **p-series**.

Tip: This isn't in the textbook, but remember that different infinities are different! If two terms go to infinity, you cannot just 'cancel them out'. Example,  $\frac{\sqrt{x}}{\ln x}$ . Both go to infinity as  $x \rightarrow \infty$ , making  $\frac{\infty}{\infty}$ , but this term actually diverges to  $\infty$ , not 1. Similarly, two infinity-tending terms in a situation like  $\infty - \infty$  do not always equal 0, etc.

To figure it out, use l'Hopital's rule. Eg.  $\frac{\ln x}{x}$ . Differentiate top and bottom,  $\frac{\frac{1}{x}}{1}$ , taking limit to infinity  $\lim_{x \rightarrow \infty} \frac{1}{x} = 0$ , so in this case, the infinities actually do 'cancel out'. But you cannot just assume they will. Try l'Hopital's rule on the example that goes to  $\infty$  above.

## 7.8.2: Integrals with Discontinuities

This is **Type 2**. Integrals of this type will appear as:

$$s \int_a^t f(x) dx$$

where the function  $f(x)$  has a discontinuity at  $t$ , or like this:

$$\int_a^b f(x) dx$$

where there is a discontinuity at  $x = c$  where  $a < c < b$ .

In the first case, you replace  $t$  with the limit as the bound approaches the discontinuity, as in

$$\int_a^t f(x) dx = \lim_{t \rightarrow b^-} \int_a^b f(x) dx \quad \text{or} \quad \int_t^b f(x) dx = \lim_{t \rightarrow a^+} \int_a^b f(x) dx$$

We call the subsequent improper integral **convergent** if the limit exists and **divergent** if it does not.

Formally, we define

(a) If  $f$  is continuous on  $[a, b)$  and is discontinuous at  $b$ , then

$$\int_a^b f(x) dx = \lim_{t \rightarrow b^-} \int_a^t f(x) dx$$

if this limit exists (as a finite number).

(b) If  $f$  is continuous on  $(a, b]$  and is discontinuous at  $a$ , then

$$\int_a^b f(x) dx = \lim_{t \rightarrow a^+} \int_t^b f(x) dx$$

if this limit exists (as a finite number).

The improper integral  $\int_a^b f(x) dx$  is called **convergent** if the corresponding limit exists and **divergent** if the limit does not exist.

(c) If  $f$  has a discontinuity at  $c$ , where  $a < c < b$ , and both  $\int_a^c f(x) dx$  and  $\int_c^b f(x) dx$  are convergent, then we define

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

The Type II p-series acts differently than the Type I. It is one of the biggest mistakes students make (McLean said that).

$$\int_0^1 \frac{1}{x^p} dx \text{ is convergent if } p < 1 \text{ and divergent if } p \geq 1.$$

or the exact opposite of the Type I.

## 7.8.3: The Comparison Test

This is a way to find out if an improper integral converges or diverges without finding its exact value.

If  $f, g$  are continuous functions with  $f(x) \geq g(x) \geq 0$  for  $x \geq a$ ,

(a) If  $\int_a^\infty f(x) dx$  is convergent, then  $\int_a^\infty g(x) dx$  is convergent.

(b) If  $\int_a^\infty g(x) dx$  is divergent, then  $\int_a^\infty f(x) dx$  is divergent.

Basically, if you know a function  $f$  to be either convergent or divergent, you can guarantee another function  $g$  to be one of them based on the relationship between their  $y$ -values. If  $f$  converges and  $g$  is underneath  $f$  for all  $x$  (at least, all relevant  $x$  values in the integral), it will also converge. If  $f$  diverges and  $g$  is above  $f$  for all  $x$ , it will also diverge.

## 11.1: Sequences

Sequences are lists of numbers following some specific pattern. As such, they can be represented as functions or in sigma notation. (Note: sequences and series are NOT the same! This is a common mistake! Most of the theorems in Ch. 11 will only work for one or the other so do not mix them up. Additionally, if a sequence converges, it does NOT guarantee its series converges! However, since convergent series will always have convergent sequences, this may lead you to form incorrect assumptions when in reality you got lucky.)

### 11.1.1: Infinite Sequences

An infinite sequence is a sequence where there is a term  $a_n$  for every positive integer  $n$ .

If we graph a lot of values of an infinite sequence, it may appear to begin to converge at a specific value.

A sequence  $\{a_n\}$  has the limit  $L$  and we write

$$\lim_{n \rightarrow \infty} a_n = L \quad \text{or} \quad a_n \rightarrow L \text{ as } n \rightarrow \infty$$

if we can make the terms  $a_n$  as close to  $L$  as we like by taking  $n$  sufficiently large.

If  $L$  exists, we say the sequence **converges** (or is **convergent**).

Otherwise, we say the sequence **diverges** (or is **divergent**).

Again, do not confuse this with the convergence of a series.

If we define the general term  $a_n$  to be expressed by a function  $f(n) = a_n$ , we can say that  $\lim_{n \rightarrow \infty} a_n = L$ . Find below some useful sequence laws. These are the same as the limit laws you already know but in a slightly different format.

Suppose that  $\{a_n\}$  and  $\{b_n\}$  are convergent sequences and  $c$  is a constant. Then:

1.  $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$
2.  $\lim_{n \rightarrow \infty} (a_n - b_n) = \lim_{n \rightarrow \infty} a_n - \lim_{n \rightarrow \infty} b_n$
3.  $\lim_{n \rightarrow \infty} ca_n = c \lim_{n \rightarrow \infty} a_n$
4.  $\lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n$
5.  $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}$  if  $\lim_{n \rightarrow \infty} b_n \neq 0$
6.  $\lim_{n \rightarrow \infty} a_n^p = [\lim_{n \rightarrow \infty} a_n]^p$  if  $p > 0$  and  $a_n > 0$

The squeeze theorem is also helpful. It is like the comparison test, but you determine the exact value of a limit.

Essentially, if a sequence is 'squished' between two other sequences above and on the bottom that both converge to the same value, the sequence in between will also converge to that value. This can be very powerful if you know how to use it.

If  $a_n \leq b_n \leq c_n$  for  $n \geq n_0$  and  $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = L$ , then  $\lim_{n \rightarrow \infty} b_n = L$ .

The following formula can also prove useful:

If  $\lim_{n \rightarrow \infty} |a_n| = 0$ , then  $\lim_{n \rightarrow \infty} a_n = 0$

A sequence's convergeability cannot be removed by function composition, that is if you pipe every term of a convergent sequence through a function  $f(a_n)$ , it will still be convergent. Formally,

If  $\lim_{n \rightarrow \infty} a_n = L$  and the function  $f$  is continuous at  $L$ , then

$$\lim_{n \rightarrow \infty} f(a_n) = f(L)$$

This may be important: these are what will happen for sequences of the form  $r^n$ :

$$\lim_{n \rightarrow \infty} r^n = \begin{cases} 0 & \text{if } -1 < r < 1 \\ 1 & \text{if } r = 1 \\ \text{does not exist} & \text{if } r = -1 \\ \text{diverges} & \text{if } |r| > 1 \end{cases}$$

### 11.1.2: Monotonic/Bounded Sequences

A sequence is called **increasing** if  $a_{n+1} > a_n$  for all terms  $n$ . A sequence is **decreasing** if  $a_{n+1} < a_n$  for all terms  $n$ . If a sequence is either increasing or decreasing, it is **monotonic**.

A sequence  $a_n$  is **bounded above** by  $M$  if:

$$a_n \leq M \text{ for all } n \geq 1$$

Or it is **bounded below** by  $m$  if:

$$m \leq a_n \text{ for all } n \geq 1$$

If a sequence is bound above and below, then it is called a **bounded sequence**.

The **Monotonic Sequence Theorem** states that

Every bounded, monotonic sequence is convergent.

...

In particular, an increasing and bounded above sequence converges and a decreasing and bounded below sequence converges.

## 11.2: Series

A series is the sum of a sequence's terms. Please do remember not to confuse the two.

### 11.2.1: Infinite Series

Sigma notation is most commonly used for series. An infinite series, where all the infinite terms of an infinite sequence are summed, is notated like such:

$$\sum_{n=1}^{\infty} a_n \quad \text{or} \quad \sum a_n$$

A partial sum is a sum of the terms of a sequence up to a term  $n$ , such that  $s_n = a_1 + a_2 + \cdots + a_n$ .

Given the series  $\sum_{n=1}^{\infty} a_n$  and its  $n$ -th partial sum

$$s_n = a_1 + a_2 + \cdots + a_n,$$

If the sequence of partial sums  $\{s_n\}$  converges and  $\lim_{n \rightarrow \infty} s_n = L$  where  $L$  is a real number,

$\{a_n\}$  is **convergent** and  $L$  is the **sum** of the series, expressed as

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

But if  $\{s_n\}$  **diverges**, the series is **divergent**.

Something you should take away from this is that a series is just **the limit of the sequence of partial sums of a sequence**. This can be helpful to visualize the concept.

### 11.2.2: Sum of a Geometric Series

The geometric series is a series of form

$$\sum_{n=1}^{\infty} ar^{n-1} \quad \text{where } a \neq 0$$

This series is **very common**, and all you must know about it is that

<p>A geometric series is convergent if</p> $ r  < 1$ <p>If a geometric series is convergent, its sum is</p> $s_n = \frac{a}{1 - r} \quad  r  < 1$ <p>If <math> r  &gt; 1</math>, the series is divergent.</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Note that some questions may pose a geometric series with a shifted index, like  $ar^n$ , and you will need to shift it back/front accordingly to bring it to the provided form, or else the equations will not work. If you are provided a shifted index, you may alternatively just use the fact that  $a$  is not necessarily the value in the format given, but the first term of the geometric sequence. Find the first term and use that instead, and  $r$  should be unaffected by a shifted index, and you'll get the same answer.

### 11.2.3: Tests for Convergence

Remember a series is divergent if its sequence of partial sums is divergent.

<p>If the series <math>\sum_{n=1}^{\infty} a_n</math> is convergent, then <math>\lim_{n \rightarrow \infty} a_n = 0</math></p>
--------------------------------------------------------------------------------------------------------------------------------

Essentially, if a series is convergent, its sequence will converge to 0. **Common mistake:** this is not true in reverse! This is a very very very common mistake to make. Kindly do not make it. All convergent series will have a 0-sequence, but not all 0-sequences have a convergent series. It's the square/rectangle thing.

<p>If <math>\lim_{n \rightarrow \infty} a_n</math> does not exist or <math>\neq 0</math>, the series is divergent.</p>
------------------------------------------------------------------------------------------------------------------------

Remember the common mistake, this does not mean that the series will be convergent if the limit is 0. Here is a helpful graphic:

$$\lim_{n \rightarrow \infty} a_n = \begin{cases} \text{D.N.E :} & \text{Series is guaranteed divergent.} \\ \neq 0 : & \text{Series is guaranteed divergent.} \\ 0 : & \text{We have no idea. (The series could be convergent, but we don't know based on this.)} \end{cases}$$

The first three limit laws listed above for sequences are also correct for series. In particular, these are the addition, subtraction, and multiplication laws.

<p>If <math>\sum a_n</math> and <math>\sum b_n</math> are convergent series, then:</p>
----------------------------------------------------------------------------------------

<p>(a) <math>\sum_{n=1}^{\infty} ca_n = c \sum_{n=1}^{\infty} a_n</math></p> <p>(b) <math>\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n</math></p> <p>(c) <math>\sum_{n=1}^{\infty} (a_n - b_n) = \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} b_n</math></p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 11.3: The Integral Test and Estimates of Sums

### 11.3.1: The Integral Test

<p>Suppose <math>f</math> is a continuous, decreasing, positive function on <math>[1, \infty)</math> and let <math>a_n = f(n)</math>. Then the series <math>\sum_{n=1}^{\infty} a_n</math> converges if and only if the improper integral <math>\int_1^{\infty} f(x) dx</math> is convergent.</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<p>(i) If <math>\int_1^{\infty} f(x) dx</math> is convergent, then <math>\sum_{n=1}^{\infty} a_n</math> is convergent.</p> <p>(ii) If <math>\int_1^{\infty} f(x) dx</math> is divergent, then <math>\sum_{n=1}^{\infty} a_n</math> is divergent.</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Here are the integral test's caveats in a more legible form.

- The function must be ultimately decreasing for  $x > N$  for some number  $N$ .
- The function must be entirely positive.
- The function must be completely continuous.
- The three conditions above must only be true on the integral's interval.
- You can start the improper integral with other bounds like  $\int_n^\infty$ .

Note that although you get a numeric result if it is convergent, that result is **not** the sum itself.

The p-series was discussed back in 7.8.1, but only its role in an improper integral. That rule can also be applied to the p-series as a series.

$$\sum_{n=1}^{\infty} \frac{1}{n^p} dx \text{ is convergent if } p > 1 \text{ and divergent if } p \leq 1.$$

**Common mistake:** Please remember that an improper integral is not the same thing as a series summation. All the improper integral does is prove the sum of the area underneath the sequence exists as  $x$  gets large.

$$\sum_{n=1}^{\infty} a_n \neq \int_1^{\infty} f(x) dx$$

### 11.3.2: Estimating the Sum of a Series

The partial sum at any term  $n$  can be used as a approximation of the actual sum,  $s$ . The difference between an approximation of the sum and the sum itself,  $s - s_n$  is known as  $R_n$  or the **remainder** of the approximation.

$$\begin{array}{l} \text{Suppose } f(k) = a_k, \text{ where } f \text{ is a continuous, positive, decreasing function} \\ \text{for } x \geq n \text{ and } \sum_{n=1}^{\infty} a_n \text{ is convergent. If } R_n = s - s_n, \\ \int_{n+1}^{\infty} f(x) \leq R_n \leq \int_n^{\infty} f(x) dx \end{array}$$

Or in simpler terms, iterating through partial sums of each term, each successive approximation will be more precise than the one before. This is useful when a question asks for accuracy to certain decimal point. Note that the improper integrals are NOT the approximations themselves. Here is some explanation on what these integrals actually are:

- Recall that  $R_n$  is not an integral, but a sigma infinite summation of the sequence.
- Since  $f$  is decreasing, the sum of values at integer points can be compared to integrals:
  - The sum from  $k = n + 1$  to  $\infty$  is roughly comparable to the integral of  $f(x)$  from  $x = n + 1$  to  $\infty$ .
    - This can be a **common mistake**: people do not understand how the sum of a series from  $n \rightarrow \infty$  is not the same as the integral on the same range. Why?
      - On sigma notation, we add each term for **every positive integer** value of  $n$ .
      - For the integral, we add each term for **every positive** value of  $n$ .
      - Don't confuse all of this with the convergence of the sequence!
- The integral  $\int_{n+1}^{\infty} f(x) dx$  underestimates the sum because the rectangle heights (like  $f(k)$ ) are greater than or equal to  $f(x)$  for  $x \in [k, k + 1]$ .
  - But if the integral is summing so many more values while the sum is summing only values from each positive integer, how is this possible? If the integral is summing  $f(59.999)$ ,  $f(60)$ ,  $f(60.001)$  while the summation is summing only  $f(60)$ , how is the summation larger than the integral?
    - Well, remember the condition of the integral test: the function **must be decreasing**. If we create a rectangle spanning width 1, it will **EXPAND PAST THE SLOPE OF THE DECREASING LINE**, since we only take a sample every positive integer.
    - In contrast, the integral is much smoother, since we take infinite samples, and thus is always smaller than the partial sum.
- The final step to understanding all of this is that the rightmost term  $\int_n^{\infty} f(x) dx$  is the **only** term that starts at  $n$ . The leftmost term and the remainder both start at  $n + 1$ . While the subscript  $R_n$  seems to imply it is the sum of the terms  $n$  to  $\infty$ , look at its definition: it only sums the terms from  $n + 1$  to  $\infty$ . As such, this theorem states that if we shift back



the index to  $n$ , the extra area afforded to the integral in the area  $x = (n, n + 1)$  is enough to make it so the integral is **greater** than all that extra area over the curves given to the summation's rectangles.

- [HERE IS ANOTHER GRAPH](#) that demonstrates this. In order from top to bottom are the left to right terms in the inequality.

By adding  $s_n$  to all sides, we get

$$s_n + \int_{n+1}^{\infty} f(x) \leq s \leq s_n + \int_n^{\infty} f(x) dx$$

This can be used to provide upper/lower bounds for  $s$ .

## 11.4: The Comparison Tests

The comparison tests work similarly for series as they do for sequences.

### 11.4.1: Direct Comparison Test

Suppose that  $\Sigma a_n, \Sigma b_n$  are series with positive terms.

- (i) If  $\Sigma a_n$  is convergent and  $b_n \leq a_n$  for all  $n$ ,  $\Sigma b_n$  is also convergent.
- (ii) If  $\Sigma a_n$  is divergent and  $b_n \geq a_n$  for all  $n$ ,  $\Sigma b_n$  is also divergent.

- Since you need a known convergent/divergent series to use the direct comparison tests, these two families are most common.
  - P-series
    - Converges when  $p > 1$ , diverges when  $p \leq 1$
  - Geometric series
    - Converges when  $|r| < 1$ , diverges when  $|r| \geq 1$

### 11.4.2: Limit Comparison Test

The direct comparison test is useless in the case where your known series is divergent and larger than your target series, or when your known series is convergent and smaller than your target series. In these cases where you still want to use a comparison (like if the series is very similar to one of the good families above), you use this.

Suppose that  $\Sigma a_n, \Sigma b_n$  are series with positive terms. If

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = c$$

where  $c$  is finite and  $c > 0$ , then either both series converge or both diverge.

### 11.4.3: Estimating Sums

If we can say that every term of the series  $\{a_n\}$  is smaller or equal to than every corresponding term of the sequence  $\{b_n\}$ , then the remainder of  $\{a_n\}$ ,  $R_n$ , will be smaller or equal to the remainder of  $\{b_n\}$ ,  $T_n$ .

If we look at the remainder estimate for the integral test, we have

$$\int_{n+1}^{\infty} f(x) \leq R_n \leq \int_n^{\infty} f(x) dx$$

which we can adapt to this method of sum estimation:

$$R_n \leq T_n \leq \int_n^{\infty} f(x) dx$$

We can use this to approximate the sum of a series given  $n$ . Remember all these methods to estimate sums do not give the sum itself, unless you are given the partial sum  $s_n$ . You estimate the sum by finding  $R_n$ , or how far the partial sum at  $n$  is from the real sum.

## 11.5: Alternating Series and Absolute Convergence

### 11.5.1: The Alternating Series

An **alternating series** is a series of form  $a_n = (-1)^{(n-1 \mid n)} b_n$  where  $b_n$  is positive.

If the alternating series

$$a_n = (-1)^{(n-1 \mid n)} b_n \quad (b_n > 0)$$

satisfies the conditions:

(i)  $b_{n+1} \leq b_n$  for all  $n$

(ii)  $\lim_{n \rightarrow \infty} b_n = 0$

Then the series is guaranteed to be convergent.

This is obvious if you think about it.

### 11.5.2: Estimating Sums of Alternating Series

If  $s = \sum (-1)^{n-1} b_n$ , where  $b_n > 0$ , is the sum of an alternating series that satisfies

(i)  $b_{n+1} \leq b_n$  and (ii)  $\lim_{n \rightarrow \infty} b_n = 0$

then:

$$|R_n| = |s - s_n| \leq b_{n+1}$$

Essentially, the remainder is always smaller than the **value** of the first neglected term. Note this does not apply to other kinds of series. Do not try.

### 11.5.3: Absolute and Conditional Convergence

A series  $\sum a_n$  is called **absolutely convergent** if the series of absolute values  $\sum |a_n|$  is convergent.

Note that if a convergent series only has positive terms, it will be absolutely convergent too.

A series  $\sum a_n$  is called **conditionally convergent** if it is convergent but not absolutely convergent.

The following is evident:

If a series  $\sum a_n$  is absolutely convergent, it is also convergent.

## 11.6: The Ratio and Root Tests

### 11.6.1: Ratio Test

(i) If  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$ , then  $\sum_{n=1}^{\infty} a_n$  is absolutely convergent (and convergent).

(ii) If  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1$  or  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty$ , then  $\sum_{n=1}^{\infty} a_n$  is divergent.

(iii) If  $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$ , the Ratio Test is inconclusive.

The ratio test usually will fail for certain types of sequences. More on this in 11.7.

### 11.6.2: Root Test

This is most conveniently used when there are  $n$ -th powers involved.

(i) If  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = L < 1$ , then the series  $\sum_{n=1}^{\infty} a_n$  is absolutely convergent (and therefore convergent).

(ii) If  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = L > 1$  or  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \infty$ , then the series  $\sum_{n=1}^{\infty} a_n$  is divergent.

(iii) If  $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1$ , the Root Test is inconclusive.

If the root test fails, the ratio test will also fail. However, if the ratio test fails, there are a small number of situations where the root test will still succeed. In most other cases, the root test really isn't worth using.

- For example, when there is a  $n$ -th power **overall**, as in over the whole thing and not just a term or two.

## 11.7: Strategy for Testing Series

The strategy around testing a series for convergence/divergence is based on its form.

1. **Test for Divergence:** If you can see that  $\lim_{n \rightarrow \infty} a_n$  may be different from 0, then apply the Test for Divergence.
2. **p-Series:** If the series is of the form  $\sum 1/n^p$ , then it is a p-series, which we know to be convergent if  $p > 1$  and divergent if  $p \leq 1$ .
3. **Geometric Series:** If the series has the form  $\sum ar^{n-1}$  or  $\sum ar^n$ , then it is a geometric series, which converges if  $|r| < 1$  and diverges if  $|r| \geq 1$ . You may need to do some algebraic manipulation.
4. **Comparison Tests:** If the series has a form that is similar to a p-series or a geometric series, then a comparison test is a good choice.
  - In particular, if  $a_n$  is a rational function or an algebraic function of  $n$  (involving roots of polynomials), then use a p-series.
    - Break it down to a p-series form, then compare. You should keep the highest power of  $p$ .
  - The comparison tests apply only to series with positive terms, but if  $\sum a_n$  has some negative terms, then we can apply a comparison test to  $\sum |a_n|$  and test for absolute convergence instead.
5. **Alternating Series Test:** If the series takes the form  $\sum (-1)^{n-1} b_n$  or  $\sum (-1)^n b_n$ , then the Alternating Series Test is a strong candidate.
  - Note that if  $\sum b_n$  converges, the original series is absolutely convergent and thus converges.
6. **Ratio Test:** This test is often effective for series containing factorials or other products, including terms with a base raised to the  $n$ -th power.
  - Be aware that for p-series and rational/algebraic functions of  $n$ , the limit of  $|a_{n+1}/a_n|$  approaches 1 as  $n \rightarrow \infty$ . Therefore, the Ratio Test is generally not useful for these types of series.
    - Use a Comparison Test instead.
7. **Root Test:** If the  $n$ -th term  $a_n$  can be expressed in the form  $(b_n)^n$ , then the Root Test can be a helpful tool.
8. **Integral Test:** When  $a_n = f(n)$  and the integral  $\int_1^{\infty} f(x)dx$  is straightforward to evaluate, the Integral Test can be applied effectively, provided its hypotheses are met.
  - Remember the function must be ultimately decreasing, positive, and continuous on the interval.

## 11.8: Power Series

### 11.8.1: Power Series

A power series is a series of form

$$\sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + \dots$$

where the  $c$ s are constants and the  $x$ s are variables. For values of  $x$ , the series may either converge or diverge. The goal of solving a power series is usually to find for what values of  $x$  the series will converge, the **interval of convergence**.

A series of the form

$$\sum_{n=0}^{\infty} c_n(x-a)^n = c_0 + c_1(x-a) + c_2(x-a)^2 + \dots$$

is called a **power series centered at  $a$** .

To solve a power series, you usually use either the **ratio test** or the **root test**.

## 11.8.2: Interval of Convergence

For a power series  $c_n(x-a)^n$ , there are only three possibilities:

- (i) The series converges only at  $x = a$ .
- (ii) The series converges for all  $x$ .
- (iii) There is a positive number  $R$  for which the series converges if  $|x-a| < R$  and diverges if  $|x-a| > R$ .

$R$  is the **radius of convergence** of the power series. In case (i),  $R = 0$ , and in (ii),  $R = \infty$ . The **interval of convergence** is all values for which the power series converges. In case (i), the interval of convergence is a single point  $x = a$ , and in (ii), the radius of convergence is  $x = (-\infty, \infty)$ . In case (iii), the inequality can be transformed:

$$|x-a| < R \longrightarrow a-R < x < a+R$$

to form the interval of convergence.

In case (iii), when  $x = a \pm R$ , the power series may converge or diverge. Thus, if you find yourself in this situation, **you must test both endpoints** to make sure they either converge or diverge. Eg. On the interval  $2-5$ , you must test both 2 and 5 to figure out if it is  $[2, 5]$ ,  $(2, 5]$ ,  $\dots$

## Appendix E: Sigma Notation

You should probably know what sigma notation is before starting the course, especially if you read the prior chapters (sorry but I'm not putting the chapters in strictly learning order), but contained are some formulas and facts that may be important.

If  $a_m, a_{m+1}, \dots, a_n$  are real numbers and  $m, n$  are integers such that  $m \leq n$ , then

$$\sum_{i=m}^n a_i = a_m, a_{m+1}, \dots, a_n$$

This notation means the following:

- Starting from  $i = m$ , for every integer value of  $i$ , the term  $a_i$  is added to a cumulative sum.
- The summing stops when  $i = n$ , thus the last value is  $a_n$ .

Several rules similar to those from the Limit Laws hold.

If  $c$  is any constant (not dependent on  $i$ ), then:

$$\begin{aligned} \text{(a)} \quad & \sum_{i=m}^n ca_i = c \sum_{i=m}^n a_i \\ \text{(b)} \quad & \sum_{i=m}^n (a_i + b_i) = \sum_{i=m}^n a_i + \sum_{i=m}^n b_i \\ \text{(b)} \quad & \sum_{i=m}^n (a_i - b_i) = \sum_{i=m}^n a_i - \sum_{i=m}^n b_i \end{aligned}$$

The following formulas are the most useful part of this section.

Let $c$ be a constant and $n$ a positive integer.	
(a) $\sum_{i=1}^n 1 = n$	(b) $\sum_{i=1}^n c = nc$
(c) $\sum_{i=1}^n i = \frac{n(n+1)}{2}$	(d) $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$
(e) $\sum_{i=1}^n i^3 = \left[ \frac{n(n+1)}{2} \right]^2$	

At least memorize the last three, since the first two are at least somewhat derivable on the fly.

## Mathematical Induction

The textbook is ████ when it comes to induction, so I recommend [THIS VIDEO](#) by my mathematics GOAT Professor Leonard to learn about it. Here is what the textbook has to say:

Let  $S_n$  be a statement about the positive integer  $n$ . If we prove that:

- $S_1$  is true.
- $S_{k+1}$  is true whenever  $S_k$  is true.

Then  $S_k$  is true for all positive integers  $n$ .

Just do a lot of practice problems and you'll probably be fine.

## 2. Power Series, Differentials, and Parametrics

The following will cover around the material from Test 2.

### 11.9: Representations of Functions as Power Series

#### 11.9.1: Representations of Functions using Geometric Series

Similar to how we used trigonometric identities to break down integrals into more easily solved forms, we can do the same to help convert functions into power series using a geometric series.

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n \quad |x| < 1$$

The idea is to turn a function into this form, where we replace  $x$  by something else, and turn it into a power series. Remember that not necessarily all of the  $x$  terms must be removed in the power series:  $x^3 \sum_{n=0}^{\infty} x^n$  is perfectly valid. This is because  $x$  does not rely on  $n$ , so we can remove it without issue. This means you can reshape the sequence, factor out  $x$  terms, do whatever to get it in the format required.

#### 11.9.2: Differentiation and Integration of Power Series

If a function can be a power series, we might like to integrate or differentiate it. A way to do this is to differentiate/integrate each term of the power series separately.

If the power series  $\sum c_n(x-a)^n$  has a radius of convergence  $R > 0$ , then the function

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots = \sum_{n=0}^{\infty} c_n(x-a)^n$$

is differentiable (and therefore continuous) on the interval

$$(a-R, a+R)$$

and the following is true.

$$(i) \quad f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$$

$$(ii) \quad \int f(x) dx = C + c_0(x-a) + c_1 \frac{(x-a)^2}{2} + c_2 \frac{(x-a)^3}{3} + \cdots \\ = C + \sum_{n=0}^{\infty} c_n \frac{(x-a)^{n+1}}{n+1}$$

Here is an alternate symbolic phrasing of the two true statements above:

$$(i.alt) \quad \frac{d}{dx} \left[ \sum_{n=0}^{\infty} c_n(x-a)^n \right] = \sum_{n=0}^{\infty} \frac{d}{dx} [c_n(x-a)^n] \\ (ii.alt) \quad \int \left[ \sum_{n=0}^{\infty} c_n(x-a)^n \right] dx = \sum_{n=0}^{\infty} \int [c_n(x-a)^n] dx$$

There are a few considerations you should keep in mind:

- These only work with **power series**. Do not rely on this for other types of series.
- The **radius of convergence**  $R$  is unchanged after differentiation/integration, but its **interval of convergence** may change.
  - Specifically, the new differentiated/integrated function may now diverge/converge at one of the interval's endpoints.

## 11.10: Taylor and Maclaurin Series

In the last chapter, we used a geometric series to turn a function into a power series. Using what we will learn in this chapter, we can express some more functions as power series.

### 11.10.1: Definitions of Taylor and Maclaurin Series

If  $f$  has a power series expansion at  $a$ , or

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots \quad |x-a| < R$$

then its coefficients are given by

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

In other words,

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

This series is the **Taylor expansion** of  $f(x)$  about  $a$ . When  $a = 0$ ,

$$\begin{aligned}
 f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \\
 &= f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots
 \end{aligned}$$

This series is called the **Maclaurin series**.

Note that a Taylor series's sum is not necessarily equal to  $f$ , like with a geometric series. If the function does not have a valid power series representation at  $a$ , then its Taylor series will just be an approximation. Additionally, the power series expansion of  $f$  at  $a$  is unique.

### 11.10.2: When is a Function Represented by its Taylor Series?

We call the sum of polynomials up to the  $n$ -th term of a Taylor series  $\frac{f^{(n)}(a)}{n!}(x-a)^n$  by the name  $T_n(x)$ , or the  $n$ -th degree Taylor polynomial of  $f$  at  $a$ . Let the quantity  $R_n(x)$  be the remainder of a Taylor series' sum and the function itself at  $n$  terms, i.e  $f(x) = T_n(x) + R_n(x)$ . If we can prove that  $R(x)$  approaches 0 at infinity, we can say that the Taylor series represents the function. Formally,

$$\begin{aligned}
 &\text{If } f(x) = T_n(x) + R_n(x), \text{ where } T_n \text{ is the } n\text{-th degree Taylor polynomial of } f \text{ at } a, \text{ and} \\
 &\quad \lim_{n \rightarrow \infty} R_n(x) = 0 \\
 &\text{for } |x - a| < R, \text{ then } f \text{ is equal to the sum of the Taylor series on } |x - a| < R.
 \end{aligned}$$

The following is called **Taylor's Inequality** and is useful to prove a function is equal to the sum of its Taylor series on an interval.

$$\begin{aligned}
 &\text{If } |f^{(n+1)}(x)| \leq M \text{ for } |x - a| \leq d, \text{ then the remainder } R_n(x) \text{ of the Taylor series satisfies} \\
 &\quad |R_n(x)| \leq \frac{M}{(n+1)!} |x - a|^{n+1} \quad \text{for } |x - a| \leq d
 \end{aligned}$$

**Common mistake:** In Taylor's Inequality, the numerator is of form  $f^{(n+1)}(x)$ , while in the regular Taylor/Maclaurin series, it's  $f^n(a)$  instead. Don't mix this up.

Taylor's inequality is much harder to grasp than the Alternating Series Estimate or the Integral Estimate even though they may seem similar. I will try to explain in a way different from the textbook.

Let's say we have a function  $f(x)$ . Then, we try to approximate it with its  $n$ th Taylor polynomial. It is clear that this is not a perfect approximation of the function as you move further away from the point it is centered on,  $a$ . We may not be able to find the exact value of the difference between the function and your approximation, but we sure can define how big it can be in the worst possible scenario. We are interested in the difference between the functions specifically in a domain of  $x$  including the center point,  $[x_1, x_2]$ .

**We do not know what  $x$  is.** This is important. We have a range of  $x$ -values we are interested in,  $[x_1, x_2]$ , but we could be at any  $x$  in that range. As such, we need to **consider the worst-case at every turn**. Even if a question asks you for the lowest possible value of  $R_n$ , you must find the 'maximum' value of  $R_n$  by choosing your value of  $x$  to get the highest answer. This is because we are doing the worst-case. If we do not maximize its value, our answer will not work for some of the  $x$ -values in the domain, which will be wrong.

Let's take a look at the last part in the inequality,  $|x - a|^{n+1}$ . We want our  $R_n$  to be as big as possible by choosing a suitable  $x$ , so this term must be as large as possible, i.e the term furthest on the edge of  $[x_1, x_2]$  assuming  $a$  is at the interval's center. Therefore, **we choose the value of  $x$  that maximizes  $|x - a|$ .**

Then, about  $M$ . While it may be tempting to just paste-in-place  $f^{(n+1)}(x) = M$  in to get

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

remember we want to optimize this whole expression to be as large as possible. There is another way to phrase all this

called the **Lagrange form** which is a bit clearer:

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} |x - a|^{n+1}$$

which establishes that the argument of the  $f^{(n+1)}$  is **not just some**  $x$ , but whichever value in the domain  $[x_1, x_2]$  that maximizes the value of the function. In essence, we can rewrite Taylor's inequality as

$$|R_n(x)| \leq \left[ \max_{\xi \in [x_1, x_2]} \left( \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \right| \right) \right] \left[ \max_{x \in [x_1, x_2]} (|x - a|)^{n+1} \right]$$

to emphasize that we are manipulating the values of  $x$  we use in order to maximize the value of  $R_n$ , which, reminder, we need to do, or else our solution will not cover every  $x$ -value in  $[x_1, x_2]$ .

$n$  is provided in the problem, of course, and you can just plug it in as is. (Editorial note: while I usually reserve the `\boxed` element for theorems out of the textbook, I'm using it here because I believe it to be clearer than the textbook version of Taylor's Inequality.)

This fact may also come up when solving these kinds of problems, as well as other limit problems:

$$\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$$

It can be found that  $e^x$  is equal to the sum of its Maclaurin series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{for all } x$$

Fun fact: this is a way to find an expression for the constant  $e$ , as if we set  $x = 1$ :

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

### 11.10.3: Maclaurin Series of Important Functions

You should try to derive these yourself as an exercise, but here they are in one block for reference. And yes, you should be expected to memorize all of these. The  $\binom{k}{n}$  in the bottom is called **general binomial coefficient**, and it is **not** the same thing as a combinatorial.

$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots$	for $ x  < 1$	$R = 1$
$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$	for all $x$	$R = \infty$
$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$	for all $x$	$R = \infty$
$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$	for all $x$	$R = \infty$
$\tan^{-1} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$	for $ x  < 1$	$R = 1$
$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$	for $ x  < 1$	$R = 1$
$(1+x)^k = \sum_{n=0}^{\infty} \binom{k}{n} x^n = 1 + kx + \frac{k(k-1)}{2!} x^2 + \frac{k(k-1)(k-2)}{3!} x^3 + \dots$	for $ x  < 1, k \in \mathbb{R}$	$R = 1$

### 11.10.4: New Taylor Series from Old



If we know the Taylor series expression of a function, we can reverse-engineer a function from its expression by manipulating the series expansion we know. The previous table of known Maclaurin series expansions is useful for doing this.

Idk man that's all there is to it this is another reason I recommend you get the textbook. The textbook, the Cengage Early Transcendentals (ISBN: 978-1-337-61392-7), has examples and stuff in this section you should read. It can be expensive, which is why I recommend annas-archive.org.

### 11.10.5: Multiplication and Division of Power Series

Since series expansions are polynomials, we can multiply and divide them. So if you get asked to find the series expansion for a function that looks like a multiplication/division of two functions whose expansions you know (again, refer to the table above), you can easily find the series expansion that way.

## 11.11: Applications of Taylor Polynomials

### 11.11.1: Approximating Functions by Polynomials

The main use for Taylor polynomials is to approximate a function  $f$  around a point  $a$  with polynomials. The more terms you add, the better your approximation becomes. The remainder  $R_n(x) = f(x) - T_n(x)$  comes to mind again, since it defines how 'good' of an approximation your  $n$ -term Taylor series is to the function itself. You can also use Taylor's inequality to estimate your error.

### 11.11.2: Applications to Physics

If you want to simplify an equation in physics, you can use a Taylor series to get a simpler approximation, especially if the interval of interest is small. That's it.

## 8.2: Area of a Surface of Revolution

### 8.2.1: Finding the Area of a Surface of Revolution

A **surface of revolution** is formed when a curve is rotated about a line.

The method we use to find the area of a surface of revolution is based on splitting up the curve into small segments, then rotating each around the axis to get a very polygon-looking approximation. We find the surface area of each revolved segment and sum them all up. If we do this with a curve split into an infinite number of small revolved segments, we get the whole surface of revolution.

When rotating the curve  $y = f(x)$ ,  $a \leq x \leq b$  about the **x-axis**, the area of the surface of revolution is

$$S = \int_a^b 2\pi f(x) \sqrt{1 + [f'(x)]^2} dx$$

In other terms,

$$S = \int_a^b 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

You don't HAVE to know exactly how it works. But, you **really, really, should**, so here's my version about it. If you recall Chapter 8.1: Arc Length, the formula for arc length from  $x = a \rightarrow x = b$  is

$$s = \int_a^b \sqrt{1 + (f'(x))^2} dx$$

and if we take the derivative (oversimplification but it's practically what we're doing), we get

$$ds = \sqrt{1 + (f'(x))^2} dx$$

We needed to do this because we want to get small elements of the curve ( $ds$ ) and revolve each of them.

Now, let's revolve our small curve element around the axis. You can look into how this is calculated, but a revolved curve such as this is called a **frustum**, and the formula for a frustum's surface area is

$$\begin{aligned} A &= 2\pi \cdot r \cdot (\text{slant height}) \\ &= 2\pi \cdot f(x) \cdot ds \\ \text{Substituting } ds: \\ &= 2\pi \cdot f(x) \cdot \sqrt{1 + (f'(x))^2} \end{aligned}$$

Since we're revolving around the  $x$ -axis in this case, the radius of the revolution is the  $y$ -value, i.e  $f(x)$ .

Let's sum every one of these frustum area elements continuously by integrating over the interval.

$$S = \int_a^b 2\pi f(x) \sqrt{1 + (f'(x))^2} dx$$

This is the final equation we can use. However, if we're revolving around the  $y$ -axis instead of the  $x$ -axis, the radius would be the  $x$ -value instead.

$$S = \int_a^b 2\pi x \sqrt{1 + (f'(x))^2} dx$$

And since most  $y$ -axis problems have their equations set up like  $x = g(y)$ , we can write the equation like this:

$$S = \int_a^b 2\pi g(y) \sqrt{1 + (g'(y))^2} dy$$

Nevertheless, the most concise general form for any case can be reached by substituting  $ds = \sqrt{1 + (f'(x))^2} dx$  (or equivalent) to get

$$S = \int_a^b 2\pi r ds$$

Note that there are technically 4 variations in total: two for revolving around  $x$  and  $y$  axes, and each can be expressed in terms of  $x$  or  $y$ .

While you could just memorize these equations, it is much easier to understand how they are derived so you can fix any issues or adapt to any kind of question, since it can get confusing if you just memorized a bunch of equations.

For comparison purposes, here are all four variations.

Axis of Revolution	Function Form	Radius	Differential $ds$	Surface Area Formula
$x$ -axis	$y = f(x)$	$y$	$\sqrt{1 + (dy/dx)^2} dx$	$\int_a^b 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$
$x$ -axis	$x = g(y)$	$y$	$\sqrt{1 + (dx/dy)^2} dy$	$\int_c^d 2\pi y \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy$
$y$ -axis	$y = f(x)$	$x$	$\sqrt{1 + (dy/dx)^2} dx$	$\int_a^b 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$
$y$ -axis	$x = g(y)$	$x$	$\sqrt{1 + (dx/dy)^2} dy$	$\int_c^d 2\pi x \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy$

## 8.3: Applications to Physics and Engineering

### 8.3.1: Hydrostatic Pressure and Force

Pressure is given by  $P = \rho g d$  where  $\rho$  is density,  $g$  is the gravitational constant, and  $d$  is depth. Normally, we use these values:

- $\rho$  is 1000 kg/m<sup>3</sup> in water, which is most common. If they use another liquid, they'll tell you the density.

- $g$  is  $9.81 \text{ m/s}^2$  unless you're in space or something, in which case they'll tell you what  $g$  is in that case.
- $d$  is the special one.

Finding pressure at a singular point is easy. What if you want to find pressure at a rectangle? What if you want to find pressure at a larger shape? Note that the only mentioned problem type is when you're trying to find pressure against a vertical object.

In these cases, you use the fact that **pressure is the same in all directions**. This means to find the pressure acting on a rectangle, we can multiply the pressure by the rectangle's area. But since pressure increases as we go down, we can only do this **when our rectangle spans negligible depth**. In other words, its height is small. So, if we want to find the pressure acting on a vertical shape, like a trapezoid, we do the following:

1. Create a coordinate system.
  1. Pick a sensible origin, one that makes it convenient when you start integrating and such later.
2. Model the shape as a bunch of rectangles.
3. Remember the formula for pressure on a rectangle is  $P = \rho g d \cdot A$ 
  1. Adapted for an infinitesimally thin rectangle where height is  $y$  and width is  $w(y)$ ,  $P = \rho g d \cdot (w(y) \cdot dy)$
4. Integrate on the water-covered interval defined by your coordinate system,  $P = \int_a^b \rho g d w(y) dy$

### 8.3.2: Moments and Centers of Mass

#### Attention

This subchapter will not be tested in this course but is here for posterity.

We want to find a point  $P$  on a planar object where a thin needle, placed on  $P$  will balance it perfectly.  $P$  is called the **center of mass**.

A **moment** is  $m_i x_i$ , where  $m$  is a mass and  $x$  is its corresponding distance from a reference point. To find the center of mass of a system of disparate massed points, we sum all of the moments and divide by total mass, i.e

$$\bar{x} = \frac{\sum_{i=1}^n m_i x_i}{\sum_{i=1}^n m_i}$$

This works for point masses and 2d problems, but what about continuous distributions of mass?

We use the fact that the center of a mass of a rectangle is exactly on its center. Let us take a section  $x = [a, b]$  of a function  $f(x)$ , the region bounded by it  $\mathcal{R}$ . We split it up into a bunch of rectangles. Let's say we have a rectangle from  $[x_1, x_2]$  with the middle of this interval at  $x = \bar{x}_i$ . The center of mass of this rectangle will be exactly in its center  $(\bar{x}_i, \frac{1}{2} f(\bar{x}_i))$ . The mass of the region of the rectangle is its density times its area, so  $m_i = \rho f(\bar{x}_i) \Delta x$ .

The moment (about the y-axis) of the region  $R_i$  is its mass times the distance to the reference point, so

$$M_y(R_i) = (\rho \cdot f(\bar{x}_i) \cdot \Delta x) \cdot \bar{x}_i$$

The moment of the whole region  $\mathcal{R}$  will be the sum of every moment of each rectangle as the number of rectangles goes to infinity.

$$M_y = \lim_{n \rightarrow \infty} \sum_{i=1}^n \rho \bar{x}_i f(\bar{x}_i) \Delta x = \rho \int_a^b x f(x) dx$$

By changing the distance used to  $\frac{1}{2} f(\bar{x}_i)$  we can find the formula for the moment about the x-axis:

$$M_x = \lim_{n \rightarrow \infty} \sum_{i=1}^n \rho \cdot \frac{1}{2} (f(\bar{x}_i))^2 \cdot \Delta x = \rho \int_a^b \frac{1}{2} (f(x))^2 dx$$

Now the center of mass of all of  $\mathcal{R}$  (for both  $x$  and  $y$ ) using  $\bar{x} = \frac{M_y}{m}$ :

$$\bar{x} = \frac{1}{A} \int_a^b x f(x) dx \quad \bar{y} = \frac{1}{A} \int_a^b \frac{1}{2} (f(x))^2 dx$$

This all assumes that density  $\rho$  is constant. If it isn't, these will not work.

If the region  $\mathcal{R}$  is bounded by another function  $g(x)$  instead of the x-axis, then we adapt the formulas:

$$\bar{x} = \frac{1}{A} \int_a^b x [f(x) - g(x)] dx \quad \bar{y} = \frac{1}{A} \int_a^b \frac{1}{2} [(f(x))^2 - (g(x))^2] dx$$

### 8.3.3: Theorem of Pappus

#### ⚠ Attention

This subchapter will not be tested in this course but is here for posterity.

Let  $\mathcal{R}$  be a plane region that rests entirely on one side of a line  $l$  in the plane. If  $\mathcal{R}$  is rotated about  $l$ , then the entire volume of the resulting solid is the product of the area  $A$  of  $\mathcal{R}$  and the distance  $d$  traveled by the centroid of  $\mathcal{R}$ .

## 9.1: Modeling with Differential Equations

A **differential equation** is an equation that contains both an unknown function and one or more of its derivatives.

### 9.1.1: Models for Population Growth

Population growth is based not on a fixed rate, but on what the population at the time is. If  $P$  is the population at time  $t$ , then we can say that the rate of growth of the population is proportional to the population at  $t$ , or

$$\frac{dP}{dt} = kP$$

where  $k$  is a constant.

Of course, populations cannot keep growing forever, so if there is some sort of 'carrying capacity'  $M$ , the formula may look more like this:

$$\frac{dP}{dt} = kP \left( 1 - \frac{P}{M} \right)$$

This is covered in depth in chapter 9.4, but since we're not doing that chapter, we don't care.

Obvious solutions to this equation like  $P(t) = M$  or  $P(t) = 0$  are called **equilibrium solutions**.

### 9.1.2: General Differential Equations

The **order** of a differential equation is the highest power present in the equation. To solve a differential equation, we try to find an equation that when substituted into the equation, satisfies it. There will often be multiple solutions. These solutions will usually be of a family, like solutions of type  $f(x) = 2kx$  for an arbitrary constant  $k$ . Finding a family of solutions like this is called the **general solution**.

Most of the time, you will need to narrow down your solutions by satisfying a condition, called an **initial-value problem**. Basically, your solution must pass through a point  $(x_0, y_0)$ .

## 9.3: Separable Equations

### 9.3.1: Separable Differential Equations

A **separable equation** is a first-order differential equation where you can factor it into a function of  $x$  and a function of  $y$ , i.e

$$\frac{dy}{dx} = f(x)g(y)$$

Then, you can separate terms of  $x$  and terms of  $y$ , and then integrate to implicitly describe  $g(y)$  in terms of  $x$ .

$$\int f(x) dx = \int g(y) dy$$

$$f(x) = g(y) + C$$

Solving for  $y$ , we get the solution(s) to the differential equation. This leaves a constant term, which will be eliminated given an initial-value constraint.

The constant is important, don't forget it. Technically both integrals will create a constant term, but since  $C$  is arbitrary, we can combine them into one.

### 9.3.2: Orthogonal Trajectories

An **orthogonal trajectory** is a family of curves that intersects a given line perpendicularly. We say the given line and the family of curves are orthogonal trajectories of one another.

To solve an orthogonal trajectory problem, turn the given curve(s) into a differential equation by differentiating it and then solving it for a general solution.

### 9.3.3: Mixing Problems

This kind of problem is as follows: if a fixed-capacity container is filled with some amount of a substance, something is added at a fixed rate, and the mixture leaves at a fixed rate.

The main function is the amount of substance in the container as a function of time, and its derivative the rate of which it is being added, minus the rate of which it is being removed.

The shtick here is remembering to take into account both entry and removal rates in the derivative.

## 3.8: Exponential Decay and Growth

**Exponential growth** is an application of differential equations. As a reminder, the equation

$$\frac{dy}{dx} = ky$$

means that the rate of change of  $y$  (which is a function of  $x$ ) is equal to  $y$  itself multiplied by a constant. At  $y(0) = c$ ,  $c$  is the **initial value** of the equation.

The only solutions of the differential equation  $dy/dt = ky$  are the exponential functions

$$y(t) = y(0)e^{kt}$$

(I recommend trying to derive this yourself. Hint: separate the equations.)

### 3.8.1: Population Growth

Manipulating

$$\frac{dP}{dt} = kP$$

$$\frac{dP}{dt} \frac{1}{P} = k$$

The quantity  $\frac{dP}{dt} \frac{1}{P}$  is the **relative growth rate**. If the growth rate is proportional to the population size, we can say that the relative growth rate is constant.

For simple differential equations of this form, you can use the general solution provided like 10 seconds ago

$$P(t) = P(0)e^{kt}$$

to find the population at time  $t$ .

**Common mistake:** if you are provided a doubling rate or something, **do not use** the relative growth rate.  $\frac{dP}{dt} \frac{1}{P}$  compares infinitesimal changes in both variables and  $k$  is only for that case. In other words,  $\frac{dP}{dt} \neq \frac{\Delta P}{\Delta T}$ . The one on the left is the small change in  $P$  over the small interval  $t$ , and the other is the average rate of change over a finite time period. If you want to use finite values of growth, use the equation in the box instead.

### 3.8.2: Radioactive Decay

We do the same thing for radioactive decay. We can express

$$\begin{aligned}\frac{dm}{dt} &= km \\ \frac{dm \cdot m}{dt} &= k\end{aligned}$$

since radioactive materials decay faster the less material there is. Note that  $k$  in this case will be negative. We adapt the same general solution as well:

$$m(t) = m(0)e^{kt}$$

**Common mistake:** The 'half-life' is NOT  $k$ . To find the half life, you need to solve the equation  $m(t) = \frac{1}{2}m(0)$ , by finding  $k$  somehow.

### 3.8.3: Newton's Law of Cooling

This is a bit more complex. Newton's Law of Cooling is that the rate of cooling/warming of an object is proportional to the temperature difference between the object and the surroundings. Let  $T(t)$  be the temperature of the object and  $T_s(t)$  the temperature of the surroundings.

$$\begin{aligned}\frac{dT}{dt} &= k(T - T_s) \\ \frac{dT \cdot (T - T_s)}{dt} &= k\end{aligned}$$

You can reduce the term  $(T - T_s)$  to something like  $y(t)$  to make things simpler. If we do, the better version to use for problems will be

$$y(t) = y(0)e^{kt}$$

In these kinds of questions, you will probably have to take the limit of  $T$  or  $T_s$  at some point since things don't get cooler/warmer forever.

### 3.8.4: Continuously Compounded Interest

If an account is continuously granted interest per year, interest credited per year will increase as money in the account increases.

You may be familiar with the equation

$$A = A_0 \left(1 + \frac{r}{n}\right)^{nt}$$

from high school, which works, but if we take  $\lim_{n \rightarrow \infty}$ , we can use the differential equation

$$\begin{aligned}\frac{dA}{dt} &= rA \\ \frac{dA \cdot A}{dt} &= r\end{aligned}$$

and use

$$A(t) = A(0)e^{rt}$$

where  $A(t)$  is the amount of money in the account at time  $t$ ,  $r$  is the interest rate per year as a decimal, and  $t$  is the number of years.

This is called **continuous compounding**, where instead of compounding our interest at months or years, we compound at infinitesimally small blocks of time (since we took  $n \rightarrow \infty$ ) which lets us get a lot more precise.

## 9.5: Linear Equations

### 9.5.1: Linear Differential Equations

A first-order **linear differential equation** is an equation that can be put into the form

$$\frac{dy}{dx} + P(x)y = Q(x) \quad \text{or} \quad y' + P(x)y = Q(x)$$

where  $P(x)$  and  $Q(x)$  are functions continuous on a given interval.

These are not separable because there are terms containing both  $x$  and  $y$ , but we can still factor it.

All first-order linear differential equations can be factored by multiplying both sides of the formula by an **integrating factor**  $I(x)$ . The goal is to find an integrating factor such that when multiplied, the equation becomes

$$I(x)(y' + P(x)y) = I(x)Q(x)$$

Using the product rule on the left hand side with  $\frac{d}{dx}(I(x)) = I(x)P(x)$  and  $\frac{dy}{dx}y = y'$ :

$$(I(x)y)' = I(x)(y' + P(x)y) \quad (\text{LHS})$$

$$\therefore (I(x)y)' = I(x)Q(x)$$

Integrating both sides, we get

$$I(x)y = \int I(x)Q(x) dx + C$$

And we can solve for  $y$ :

$$y(x) = \frac{1}{I(x)} \int I(x)Q(x) dx + C$$

Challenge: try to derive a formula for  $I(x)$ . You will need to separate the equation at some point. Or just look below for the general formula for  $I(x)$  for a first-order linear differential equation.

To solve the linear differential equation  $dy/dx + P(x)y = Q(x)$ , multiply each side by the integrating factor  $I(x) = e^{\int P(x) dx}$  and then integrate both sides.

### 9.5.2: Application to Electrical Circuits

#### Attention

This subchapter will not be tested in this course but is here for posterity.

Ohm's Law says that the voltage drop due to a resistor is equal to  $RI$ . Faraday's Law says the voltage drop due to an inductor is equal to  $L \frac{dI}{dt}$ . If we have a circuit with both an inductor and resistor, the sum of the voltage drops must equal to sum of the voltage gain:

$$L \frac{dI}{dt} + RI = V(t)$$

(If you didn't already learn this in PHYSICS 1E03, you will soon. Enjoy!)

This is a first-order linear equation. The solution gives  $I(t)$ , current at time  $t$ .

## 10.1: Curves Defined by Parametric Equations

### 10.1.1: Parametric Equations

Functions are defined as  $y = f(x)$ , where the  $y$ -coordinate is a function of another variable  $x$ . **Parametric equations** are plotted such that both their  $y$  and  $x$  coordinates are functions of a third variable, usually  $t$ , such that  $g(t) = f(t)$  and  $(x, y) = (f(t), g(t))$ .

Since most parametric equations are not functions, they fail the vertical line test and can subsequently not be defined by a single function of form  $y = f(x)$ . However, if you are able to eliminate  $t$  and express a portion of the curve in terms of  $x$  and  $y$ , that is called **eliminating the parameter**. For example, a circle  $y^2 = 1 - x^2$  cannot be expressed as a function, but you can get close by plotting only one half with  $y = \pm\sqrt{1 - x^2}$ .

Parametric curves can also plot the direction of movement. For example, as  $t$  increases, the curve will be drawn in any specific direction, unlike a function where the function will always be plotted to the right as  $x$  increases.

### 10.1.2: The Cycloid

#### Attention

This subchapter will not be tested in this course but is here for posterity.

The curve traced by a point on the circumference of a circle as it rolls is called a **cycloid**.

I don't actually know if you need to know how it is derived, but essentially, we find the  $x$ -value by finding the total distance rolled by the center of the circle using the expression  $r\theta$ , then subtract the  $x$ -component of the radius,  $r \sin \theta$ , to get the  $x$  of a point on the edge.

To find the  $y$ -value, we do the same thing: the center of the circle is always  $r$ , so we subtract the  $y$ -component  $r \cos \theta$  to find the  $y$ -value of a point on the edge.

All of this uses the variable  $\theta$  as the third variable instead of  $t$ , representing the degrees rotated by the circle.

$$x = r(\theta - \sin \theta) \quad y = r(1 - \cos \theta) \quad \theta \in \mathbb{R}$$

You can eliminate the parameter, but you'll just get a complicated equation for a small part of the full cycloid.

### 10.1.2: Families of Parametric Curves

Some parametric curves can be in families. They can be categorized by their shared features. Often, they will share similarities in their formulae as well, such as with the linear curve family  $x = at + b, y = ct + d$ . Especially, they may look like progressions or have similar appearances as you change the constants  $a, b, c, d, \dots$

## 10.2: Calculus with Parametric Curves

### 10.2.1: Tangents

If we want to find the tangent at a point on a parametric curve, we can use the chain rule. The goal is to find  $\frac{dy}{dx}$ . The chain rule defines  $\frac{dy}{dx} = \left(\frac{dy}{du}\right) \cdot \left(\frac{du}{dx}\right)$ . Notice how the  $du$  term appears to cancel out. We do the same thing with  $x = f(t), y = g(t)$ ,

$$\frac{dy}{dx} = \left(\frac{dy}{dt}\right) \cdot \left(\frac{dt}{dx}\right) = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} \quad \text{if } \frac{dx}{dt} \neq 0$$

Intuitively,  $\frac{dy}{dx}$  represents the ratio between a small change in  $y$  and a small change in  $x$ . Since we can only represent  $x$  and  $y$  through the intermediary variable  $t$ , this formula says "for a small change in  $t$ : how much does  $y$ 's change compare to  $x$ 's



change?" It is the same thing as  $\frac{dy}{dx}$ , but this formula just has the addition of  $t$  since they both depend on it. So, we don't need to eliminate the parameter to find  $\frac{dy}{dx}$ .

If you need it, here is the formula for the second derivative,  $\frac{d^2y}{dx^2}$ :

$$\frac{d^2y}{dx^2} = \frac{d}{dx} \left( \frac{dy}{dx} \right) = \frac{\frac{d}{dt} \left( \frac{dy}{dx} \right)}{\frac{dx}{dt}}$$

To find the tangent, do as you would with a function: find  $\frac{dy}{dx}$  with the formula, substitute in, find the slope, and solve for the rest.

To find the second derivative, find  $\frac{dy}{dx}$ , then take its derivative with respect to  $t$ . Then, divide by the derivative of  $x$  with respect to  $t$ .

Also, at certain points, you may get multiple results for tangent lines. This happens when the curve doubles back over a point with a new distinct behavior. This will be represented by a situation resulting in a  $\pm$  or a polynomial while solving for  $t$  when given a point  $(x, y)$ , such as  $t^2 = 1$ . Alternatively, when facing trig functions, be careful that the points given cannot be obtained through using different angles in the trig functions: you need to account for each different valid input. Be careful!

### 10.2.2: Areas

The area underneath a curve is

$$A = \int_a^b y \, dx$$

Substituting  $y = g(t)$ ,  $dx = f'(t) \, dt$  :

$$A = \int_{\alpha}^{\beta} g(t) f'(t) \, dt$$

### 10.2.3: Arc Length

The formula for the arc length of a segment of a function is

$$L = \int_a^b \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx$$

Suppose this segment can also be described by parametric equations. Inserting what we know  $\frac{dy}{dx}$  to be,

$$L = \int_{\alpha}^{\beta} \sqrt{1 + \left( \frac{\frac{dy}{dt}}{\frac{dx}{dt}} \right)^2} \frac{dx}{dt} \, dt = \int_{\alpha}^{\beta} \sqrt{\left( \frac{dx}{dt} \right)^2 + \left( \frac{dy}{dt} \right)^2} \, dt$$

The good thing about this formula is that it works on any curve, even curves that cannot be represented as  $y = f(x)$ . There's a big proof in the textbook using MVT, a Riemann sum, and such, but formally,

If a curve  $C$  is described by  $x = f(t)$ ,  $y = g(t)$ ,  $\alpha \leq t \leq \beta$ , where  $f'$  and  $g'$  are continuous on  $[\alpha, \beta]$  and  $C$  is traversed exactly once as  $t$  goes from  $\alpha \rightarrow \beta$ , then the length of  $C$  is

$$L = \int_{\alpha}^{\beta} \sqrt{\left( \frac{dx}{dt} \right)^2 + \left( \frac{dy}{dt} \right)^2} \, dt$$

While using this formula, be careful that  $C$  is only traversed once! If you define your bounds incorrectly, you may go over the same curve twice (like using  $2\pi$  instead of  $\pi$  with a circle of radius 1)

## 3. The Polar System, Partial Derivatives, and Multi-Integrals

This covers material after Test 2. Final stretch.

### 10.3: Polar Coordinates

### 10.3.1: The Polar System

While the Cartesian system uses an ordered pair of values to determine where on the plot a point is located, the **polar coordinate system** uses the format

$$(r, \theta)$$

where  $r$  is the **radius**, the distance between the **pole** (a point we choose as the origin) and the point, and  $\theta$  the **argument**, the angle from the pole the point is situated at in radians.

### 10.3.2: Relationship Between Polar and Cartesian Coordinates

The polar system uses a distance and angle approach, meaning we can convert it into Cartesian coordinates using trigonometry. Take the pole to be at  $(0, 0)$ .

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

To convert from polar to Cartesian, we do

$$r^2 = x^2 + y^2 \quad \tan \theta = \frac{y}{x}$$

Which can all be deduced from the Pythagorean theorem and trigonometry.

### 10.3.3: Polar Curves and Symmetry

The graph of a polar equation is all points  $(r, \theta)$  that satisfy the equation.

You will not be expected to sketch polar curves in this course, but you may be expected to match a curve to an equation or something of that like. For complex curves, graph convenient points like  $\frac{\pi}{6}, \frac{\pi}{4}, \dots$  and connect them with curves. Ask yourself questions like what happens where  $r =$  some value you can see on the graph? What happens at this location where you can see an oddity?

You can also take advantage of symmetry to sketch polar curves.

- If a polar equation is unchanged when  $-\theta$  replaces  $\theta$ , it is symmetric about the polar axis.
- If the equation is unchanged when  $-r$  replaces  $r$  or  $\theta + \pi$  is replaced by  $\theta$ , the curve is symmetric about the pole.
- If the equation is unchanged when  $\pi - \theta$  replaces  $\theta$ , the curve is symmetric about  $\theta = \frac{\pi}{2}$

## 10.4: Calculus in Polar Coordinates

### 10.4.1: Area

For a region bounded by a polar curve between two angles  $a, b$ , we split the area into a bunch of sectors of a circle, then sum them all using a Riemann sum.

$$A = \int_a^b \frac{1}{2} (f(\theta))^2 d\theta$$

If you wish to find area bounded by two curves,

- Solve simultaneously to find intersections
  - This can be complex, you may need to use symmetry or geometry to find points of intersection that you can't find by solving simultaneously.
  - Drawing or sketching the curves are highly recommended.
  - This is because there are many ways of representing the same point in polar geometry.
- Solve with the formula  $A = \frac{1}{2} \int_a^b (|f(\theta)|^2 - |g(\theta)|^2) d\theta$ 
  - You can eschew the absolute value and figure out which one is on top and which one is underneath or just do it this way. Your choice

### 10.4.2: Arc Length

To find arc length, we use the cartesian-to-polar equations to convert polar curves into parametric equations, then use the same method we did for those. For a polar curve with  $r = f(\theta)$ :

$$L = \int_a^b \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2} d\theta$$

### 10.4.3: Tangents

We convert polar curves to parametric equations again and use the same method. The formula for the slope of a tangent line at  $r = f(\theta)$  is

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{\frac{dr}{d\theta} \sin \theta + r \cos \theta}{\frac{dr}{d\theta} \cos \theta - r \sin \theta}$$

We can find horizontal tangents by finding points where  $\frac{dy}{d\theta} = 0$  and vertical tangents by finding points where  $\frac{dx}{d\theta} = 0$ .

## 14.1: Functions of Several Variables

### 14.1.1: Functions of Two Variables

When a function has two variables, we denote it  $f(x, y)$ . It is a rule that gives to every ordered pair of real numbers  $(x, y)$  a corresponding value  $f(x, y)$ . The set in which all  $(x, y)$  belong to,  $D$ , is the domain of the function, and its range is the set of values of  $f(x, y)$ ,  $\{f(x, y) \mid (x, y) \in D\}$ .

As in one-variable functions we set  $y = f(x)$ , we can set  $z = f(x, y)$ .

A graph of a function of two variables is the set of all points  $(x, y, z)$  in  $\mathbb{R}^3$  such that  $z = f(x, y)$ ,  $(x, y) \in D$ .

### 14.1.2: Level Curves and Contour Maps

A **contour map** is one of those map things where a 2D plot where points of constant elevation form lines. Google it, you know what they are.

Specifically, the **level curves** of a function  $f$  of two variables are the curves with equations  $f(x, y) = k$ , where  $k$  is a constant.

Each line in a contour map is a level curve, and contour maps are usually made up of several level curves, each for a value of  $k$  at equally-spaced intervals.

The following two restrictions must also be followed:

- Level sets must be in the domain of  $f(x, y)$ . If not, you will have made meaning where there is none, and your level curves will be invalid in that spot. So, make sure you check the domain of  $f$  before making your curves.
- Level sets of different  $k$  values cannot cross one another. Well, they can, it just means  $f$  isn't a function.

### 14.1.3: Functions of Three or More Variables

This is the same deal as two-variable functions but for three variables.  $f(x, y, z) = w$ .

## 14.2: Limits and Continuity

### 14.2.1: Limits of Functions of Two Variables

These work similarly to how you would expect: if you can take your value  $f(x, y)$  as close as you like to a number  $L$  by moving the pair  $(x, y)$  closer to another pair  $(a, b)$ , then we can say that the limit of the function as  $(x, y) \rightarrow (a, b)$  is  $L$ .

Formally:

Let  $f$  be a function of two variables whose domain  $D$  includes points arbitrarily close to  $(a, b)$ . Then we say that the **limit** of  $f(x, y)$  as  $(x, y)$  approaches  $(a, b)$  is  $L$  and we write

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$$

if for every number  $\varepsilon > 0$  there is a corresponding number  $\delta > 0$  such that if  $(x, y) \in D$  and  $0 < \sqrt{(x - a)^2 + (y - b)^2} < \delta$  then  $|f(x, y) - L| < \varepsilon$

But again, you really don't need to know this exact formal definition.

### 14.2.2: Showing That a Limit Does Not Exist

On single-variable functions, we can approach a limit from either the left or right. We know if the limit from the positive end does not equal the limit from the negative end, the limit does not exist. Since there can be nearly infinite ways to approach any point since our graph looks more like a plane than a line, we adapt this rule.

If  $f(x, y) \rightarrow L_1$  as  $(x, y) \rightarrow (a, b)$  along a path  $C_1$  and  $f(x, y) \rightarrow L_2$  as  $(x, y) \rightarrow (a, b)$  along a path  $C_2$ , where  $L_1 \neq L_2$ , then  $\lim_{(x,y) \rightarrow (a,b)} f(x, y)$  does not exist.

Usually, the most convenient way to accomplish this is by having one path approach the point along the  $x$ -axis, and the other approach from the  $y$ -axis. Also note the inverse is not true: even if both limits exist, that **does not guarantee** the limit exists. As long as there is one path that does not give you the same  $L$  as the others, it will not exist.

### 14.2.3: Properties of Limits

A lot of the same ones as normal limits apply here.

1.  $\lim_{(x,y) \rightarrow (a,b)} [f(x, y) + g(x, y)] = \lim_{(x,y) \rightarrow (a,b)} f(x, y) + \lim_{(x,y) \rightarrow (a,b)} g(x, y)$
2.  $\lim_{(x,y) \rightarrow (a,b)} [f(x, y) - g(x, y)] = \lim_{(x,y) \rightarrow (a,b)} f(x, y) - \lim_{(x,y) \rightarrow (a,b)} g(x, y)$
3.  $\lim_{(x,y) \rightarrow (a,b)} [cf(x, y)] = c \lim_{(x,y) \rightarrow (a,b)} f(x, y)$
4.  $\lim_{(x,y) \rightarrow (a,b)} [f(x, y)g(x, y)] = \lim_{(x,y) \rightarrow (a,b)} f(x, y) \cdot \lim_{(x,y) \rightarrow (a,b)} g(x, y)$
5.  $\lim_{(x,y) \rightarrow (a,b)} \frac{f(x, y)}{g(x, y)} = \frac{\lim_{(x,y) \rightarrow (a,b)} f(x, y)}{\lim_{(x,y) \rightarrow (a,b)} g(x, y)}$ , provided that  $\lim_{(x,y) \rightarrow (a,b)} g(x, y) \neq 0$

The squeeze theorem is also here and it works about how you would expect.

If  $h(x, y) \leq f(x, y) \leq g(x, y)$  for all  $(x, y)$  in a disk around  $(a, b)$  (except possibly at  $(a, b)$  itself) and if

$$\lim_{(x,y) \rightarrow (a,b)} h(x, y) = L$$

and

$$\lim_{(x,y) \rightarrow (a,b)} g(x, y) = L$$

then:

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = L$$

Also, remember that algebraic combinations of common functions (exponentials, trig functions, polynomials, rationals, etc) are all continuous on their domains. If two functions of different domain are combined, the function may only exist when **both domains** are satisfied. In other words, the new domain will be the **intersection** of the domains.

### 14.2.4: Continuity

We base continuity on the direct substitution property above.

A function  $f(x, y)$  is called continuous at  $(a, b)$  if

$$\lim_{(x,y) \rightarrow (a,b)} f(x, y) = f(a, b)$$

We say that  $f$  is continuous on  $D$  if  $f$  is continuous at every point in  $D$ .

## 14.2.5: Functions of Three or More Variables

We can do all of this with functions of three or more variables. The notation

$$\lim_{(x,y,z) \rightarrow (a,b,c)} f(x, y, z) = L$$

signifies that the value of the function will approach  $L$  as the point  $(x, y, z)$  approaches the point  $(a, b, c)$ . Of course, same as two variables, continuity is defined as

$$\lim_{(x,y,z) \rightarrow (a,b,c)} f(x, y, z) = f(a, b, c)$$

Here's the formal definition if you want it, but you really don't need it.

If  $f$  is defined on a subset  $D$  of  $\mathbb{R}^n$ , then

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = L$$

means that for every number  $\varepsilon > 0$  there is a corresponding number  $\delta > 0$  such that

$$\begin{array}{l} \text{if } \mathbf{x} \in D \text{ and } 0 < |\mathbf{x} - \mathbf{a}| < \delta \\ \text{then } |f(\mathbf{x}) - L| < \varepsilon \end{array}$$

## 14.3: Partial Derivatives

### 14.3.1: Partial Derivatives of Functions of Two Variables

**Partial derivatives** are derivatives where we take one variable to be constant. For example,  $f(x, y)$ , if we keep  $y$  fixed at a constant  $c$ , we can take the derivative of the new single-variable function  $f(x)$ . We call it **the partial derivative of  $f$  with respect to  $x$  at  $(a, b)$**  and denote it by  $f_x(a, b)$ .

From the definition of a derivative, we can say that

$$f_x(a, b) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h}$$

and ditto for the derivative with respect to  $y$ . Now, if we let  $(a, b)$  vary, we get the two new functions of two variables

$$\begin{array}{l} f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h} \\ f_y(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y + h) - f(x, y)}{h} \end{array}$$

and we can solve.

There are also a bunch of other fun ways to portray partial derivatives:

$$\begin{array}{l} f_x(x, y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x, y) = \frac{\partial z}{\partial x} = f_1 = D_1 f = D_x f \\ f_y(x, y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x, y) = \frac{\partial z}{\partial y} = f_2 = D_2 f = D_y f \end{array}$$

Deriving partial derivatives in practice is pretty simple, just treat one variable as a constant, and differentiate with respect to the other one.

### 14.3.2: Interpretations of Partial Derivatives

A function  $f(x, y)$  graphically is a surface in a 3D space. If we take the derivative at  $(a, b)$  while setting a variable to constant, what we get graphically is the slope of the tangent line at the point in question to the curve created by the

constant variable. It's kind of hard to explain. Go watch a YouTube video if you really want to know. Not technically tested but can make some specific types of problems easier.

### 14.3.3: Derivatives of Three or More Variables

#### ⚠ Attention

This subchapter will not be tested in this course but is here for posterity.

We can also do partial derivatives for functions of three or more variables, for example

$$f_x(x, y, z) = \lim_{h \rightarrow 0} \frac{f(x+h, y, z) - f(x, y, z)}{h}$$

Which works about the same you would expect: hold  $y$  and  $z$  constant and differentiate  $f(x)$ . Just don't try to represent it geometrically.

### 14.3.4: Higher Derivatives

Since partial derivatives of a function of two variables  $f$  are also functions of two variables, we can create partial derivatives of those. Here are some examples of what that looks like.

$$\begin{aligned} (f_x)_x &= f_{xx} = f_{11} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2} \\ (f_x)_y &= f_{xy} = f_{12} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 z}{\partial y \partial x} \\ (f_y)_x &= f_{yx} = f_{21} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 z}{\partial x \partial y} \\ (f_y)_y &= f_{yy} = f_{22} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2} \end{aligned}$$

These look hard, but all you need to do in practice is go in order left to right. For  $f_{xy}$  for example, differentiate with respect to  $x$ , then differentiate the result with respect to  $y$ .

Notice how  $f_{xy} = f_{yx}$ . We can extend this using **Clairaut's Theorem**:

Suppose  $f$  is defined on a disk  $D$  containing the point  $(a, b)$ . If the functions  $f_{xy}$  and  $f_{yx}$  are both continuous on  $D$ , then:

$$f_{xy}(a, b) = f_{yx}(a, b)$$

Which lets us do fun things like  $f_{xyy} = f_{yxy} = f_{yyx}$  with even higher-order partial derivatives.

### 14.3.5: Partial Differential Equations

#### ⚠ Attention

This subchapter will not be tested in this course but is here for posterity.

Partial derivatives can show up in differential equations. For example,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is known as Laplace's equation and its solutions are harmonic waves.

Another example is the wave equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$$

## 14.4: Tangent Planes and Linear Approximations

### 14.4.1: Tangent Planes

The **tangent plane** at a point  $(x, y, z)$  on a plane  $z = f(x, y)$  is defined as a plane where it just touches  $(x, y, z)$  and shares the same slope as  $z = f(x, y)$  in all directions.

Suppose  $f$  has continuous partial derivatives. An equation of the tangent plane to the surface  $z = f(x, y)$  at the point  $(x_0, y_0, z_0)$  is

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

### 14.4.2: Linear Approximations

The **linearization** of a function  $f$  at a point  $(a, b)$  is a **linear function**  $L(x, y)$  that is a good approximation to  $f$  at  $(a, b)$ .

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

In single-variable functions, this is the same thing as finding the tangent.

If  $z = f(x, y)$ , then  $f$  is **differentiable** at  $(a, b)$  if  $\Delta z$  can be expressed in the form

$$\Delta z = f_x(a, b)\Delta x + f_y(a, b)\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$$

where  $\varepsilon_1$  and  $\varepsilon_2$  are functions of  $\Delta x$  and  $\Delta y$  such that  $\varepsilon_1$  and  $\varepsilon_2 \rightarrow 0$  as  $(\Delta x, \Delta y) \rightarrow (0, 0)$ .

Essentially, what this says is that the error terms  $\varepsilon$  should vanish as you approach the point, meaning the approximation is good.

This is very impractical because ain't no one actually gonna do this, so we usually use

If the partial derivatives  $f_x$  and  $f_y$  exist near  $(a, b)$  and are continuous at  $(a, b)$ , then  $f$  is differentiable at  $(a, b)$ .

### 14.4.3: Differentials

$$dz = f_x(x, y) dx + f_y(x, y) dy = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

$dz$  is known as the **total differential**.

Note  $dz$  and  $\Delta z$  are not the same quantity.

$\Delta z$  refers to how much  $z$  changes when you move from  $(x, y)$  to  $(x_1, y_1)$ , which is basically the actual change in  $z$ .

$dz$  represents the approximate change: when we linearize to create an approximation (a tangent plane) at a point,  $dz$  refers to the estimate of how much  $z$  will change as you go from  $(x, y)$  to  $(x_1, y_1)$  based on the approximation you made. It's just a number after the calculations are done.

### 14.4.4: Functions of Three or More Variables

#### Attention

This subchapter will not be tested in this course but is here for posterity.

We can do all of this for functions with more than two variables. For example, the linear approximation of  $w = f(x, y, z)$  is

$$f(x, y, z) \approx f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) = L(x, y, z)$$

the **increment** of  $w$  ( $\Delta w$ ) is

$$\Delta w = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z)$$

and the **differential**  $dw$  is

$$dw = \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \frac{\partial w}{\partial z} dz$$

## 14.5: The Chain Rule

You better know the chain rule pretty damn well to get this chapter.

Let us revisit the 'normal' chain rule: If  $y = f(x)$  and  $x = g(t)$ , i.e  $y = f(g(t))$  where both  $f$  and  $g$  are differentiable, then  $y$  is differentiable in terms of  $t$  and

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

### 14.5.1: The Chain Rule - Case 1

Let us say we have a function  $z = f(x, y)$  and both  $x = g(t)$  and  $y = h(t)$  are differentiable functions of a fourth variable  $t$ . This is **Case 1**.

Suppose that  $z = f(x, y)$  is a differentiable function of  $x$  and  $y$ , where  $x = g(t)$  and  $y = h(t)$  are both differentiable functions of  $t$ . Then  $z$  is a differentiable function of  $t$  and

$$\begin{aligned} \frac{dz}{dt} &= \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \\ &= \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \end{aligned}$$

### 14.5.2: The Chain Rule - Case 2

Let us say we have a function  $z = f(x, y)$  and both  $x = g(s, t)$  and  $y = h(s, t)$  of two more variables  $s, t$ . This is **Case 2**.

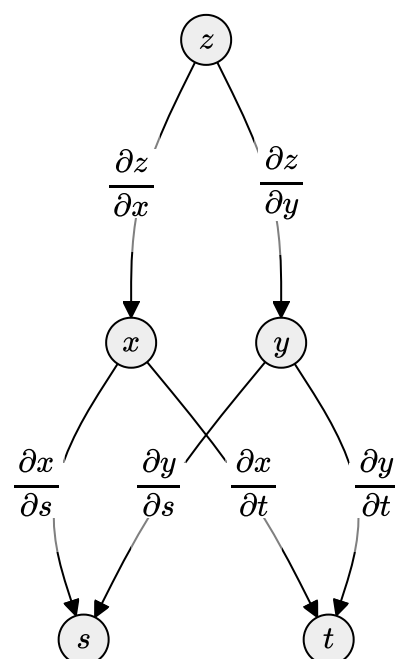
Since we have another variable now, in order to differentiate  $z = j(s, t)$ , we need to find two partial derivatives, each holding one of  $s, t$  constant. We apply the formula from Case 1 to both to get

Suppose that  $z = f(x, y)$  is a differentiable function of  $x$  and  $y$ , where  $x = g(s, t)$  and  $y = h(s, t)$  are both differentiable functions of  $s, t$ . Then:

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} \qquad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t}$$

We can call  $s$  and  $t$  **independent variables**,  $x$  and  $y$  **intermediate variables**, and  $z$  the **dependent variable**.

This flowchart helps explain it. (Yes, it's janky, I'm relying on an Obsidian plugin to use math formatting inside the Mermaid chart)



Each layer represents one of the three types of variables. If you want to find a specific quantity, for example  $\frac{dz}{ds}$ , go down the flowchart to link all occurrences of  $z$  and  $s$ , multiply each item on the same path, and add it all together.



### 14.5.3: The Chain Rule - General

The tree diagram last time was overkill since we only had to deal with two intermediate variables. Let us expand our thinking to a general variable  $u$  that is a function of  $n$  intermediate variables  $x_1, x_2, \dots, x_n$ , which are each functions of  $m$  independent variables  $t_1, t_2, \dots, t_m$ . Notice in each case, there are  $n$  total variables.

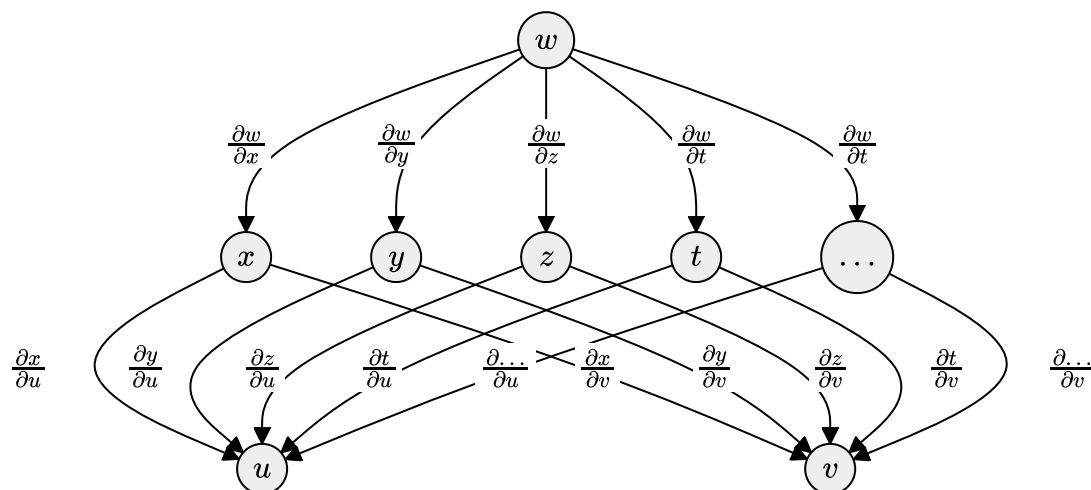
Suppose that  $u$  is a differentiable function of the  $n$  variables  $x_1, x_2, \dots, x_n$  and each  $x_j$  is a differentiable function of the  $m$  variables  $t_1, t_2, \dots, t_m$ .

Then  $u$  is a function of  $t_1, t_2, \dots, t_m$  and

$$\frac{\partial u}{\partial t_i} = \frac{\partial u}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2} \frac{\partial x_2}{\partial t_i} + \dots + \frac{\partial u}{\partial x_n} \frac{\partial x_n}{\partial t_i}$$

for each  $i = 1, 2, \dots, m$ .

Let's get another tree diagram in here.



Okay the label spacing is horrible but you see the point.

### 14.5.4: Implicit Differentiation

Suppose we have an equation  $F(x, y) = 0$  that describes  $y$  implicitly as a differentiable function of  $x$ . Remember that 'implicit' means there is a relation between  $y$  and  $x$ . If  $F$  is differentiable, we can differentiate it with the chain rule Case 1 and rearrange for  $\frac{dy}{dx}$ .

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{F_x}{F_y}$$

The reason this works when  $F(x, y)$  describes  $y$  implicitly is because we can rewrite it as  $F(x, y(x))$  and then we can use the chain rule.

Fun fact: In more advanced calculus, you will learn that solely having a format  $F(x, y) = 0$  does not necessarily describe  $y$  implicitly, and there are conditions that defines when it does: the **Implicit Function Theorem**. The following must be true:

- If  $F$  is defined on a disk containing  $(a, b)$  where:
  - $F(a, b) = 0$
  - $F_y(a, b) \neq 0$
  - $F_x$  and  $F_y$  are continuous on the disk

We can extend this to functions of more than one variable: if  $z$  is given implicitly by  $F(x, y, z) = 0$ , we can use the chain rule again:

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{F_x}{F_z}$$

$$\frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{F_y}{F_z}$$

The Implicit Function Theorem has conditions for this as well, the same thing as above but with spheres instead of disks.

## 14.6: Directional Derivatives and the Gradient Vector

### 14.6.1: Directional Derivatives

We already know the formulas for partial derivatives of a function  $z = f(x, y)$  in the  $x$  and  $y$  directions from 14.3. For reference:

$$f_x(x, y) = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h}$$

$$f_y(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y + h) - f(x, y)}{h}$$

But, if we want to go in any other direction that isn't  $\mathbf{i}$  or  $\mathbf{j}$ , we need something else. Let  $\mathbf{u}$  be an arbitrary unit vector in some direction. We have a function  $f$ , and we want to find its instantaneous rate of change in the direction from our starting point  $(x_0, y_0)$  to  $\mathbf{u}$ .

The directional derivative of  $f$  at  $(x_0, y_0)$  in direction of unit vector  $\mathbf{u} = \langle a, b \rangle$  is

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

What we did is move a small distance  $h$  from our starting point to our target, measure how much the function  $f$  changed with that small step, divide by the step's size  $h$  to get the average rate of change, then take the limit as the distance  $h$  approaches 0 to find the instantaneous rate of change.

Try plugging the unit vectors  $\mathbf{i}, \mathbf{j}$  in and seeing what happens ( $\langle 0, 1 \rangle, \langle 1, 0 \rangle$  respectively). Look familiar?

The formula above is the formal definition, but when doing problems, we usually use this one:

If  $f$  is a differentiable function of  $x$  and  $y$ , then  $f$  has a directional derivative in the direction of any unit vector  $\mathbf{u} = \langle a, b \rangle$  and

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b$$

**Warning:** Unit vectors **must** have length 1. If a problem requires you to calculate your own unit vector or derive a unit vector, be wary of this! You can check the length of a vector using the Pythagorean theorem. For example, using trigonometry with an angle will give you a vector, that while of the correct angle, is magnitudes too large and must be scaled down.

### 14.6.2: The Gradient Vector

This is pretty important for fields like machine learning.

If  $f$  is a function of two variables  $x$  and  $y$ , then the **gradient** of  $f$  is the vector function defined by

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j}$$

We get this from the last equation in the last subchapter: we can rewrite  $D_{\mathbf{u}}f(x, y) = f_x(x, y)a + f_y(x, y)b$  as a dot product.

$$\begin{aligned} D_{\mathbf{u}}f(x, y) &= f_x(x, y)a + f_y(x, y)b \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \langle a, b \rangle \\ &= \langle f_x(x, y), f_y(x, y) \rangle \cdot \mathbf{u} \end{aligned}$$

The first term comes up often enough we just gave it a special name.

We can therefore rewrite the aforementioned equation:

$$D_{\mathbf{u}}f(x, y) = \nabla f(x, y) \cdot \mathbf{u}$$

(I'm assuming you know what dot product is from physics)

### 14.6.3: Functions of Three Variables

#### Attention

This subchapter will not be tested in this course but is here for posterity.

The directional derivative of  $f$  at  $(x_0, y_0, z_0)$   
in the direction of a unit vector  $\mathbf{u} = \langle a, b, c \rangle$  is

$$D_{\mathbf{u}}f(x_0, y_0, z_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + ha, y_0 + hb, z_0 + hc) - f(x_0, y_0, z_0)}{h}$$

We can compact this by writing it in vector form:

$$D_{\mathbf{u}}f(\mathbf{x}_0) = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}$$

The gradient vector for three variables is:

$$\nabla f = \langle f_x, f_y, f_z \rangle = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$$

Then we rewrite the formula for the directional derivative:

$$D_{\mathbf{u}}f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u}$$

### 14.6.4: Maximizing the Directional Derivative

Sometimes, we want to find the direction from a point in which the rate of change is the greatest.

Suppose  $f$  is a differentiable function of two or three variables. The maximum value of the directional derivative  $D_{\mathbf{u}}f(\mathbf{x})$  is  $|\nabla f(\mathbf{x})|$  and it occurs when  $\mathbf{u}$  has the same direction as the gradient vector  $\nabla f(\mathbf{x})$ .

### 14.6.5: Tangent Planes to Level Surfaces

#### Attention

This subchapter will not be tested in this course but is here for posterity.

As a reminder, a **level plane** is a plane in which  $f(x, y, z) = c$  for some constant  $c$ . At any point  $P = (x_0, y_0, z_0)$ , the **tangent plane** is the plane that 'just touches' the surface at  $P$  and is flat in the neighborhood, serving as a good approximation to the surface at that point.

A tangent plane to the level surface  $F(x, y, z)$  at  $P(x_0, y_0, z_0)$  is defined as the plane that passes through  $P$  and has a normal vector (i.e the vector with  $\vec{n} \cdot \vec{v} = 0$  for normal vector  $\vec{n}$  and any vector in the plane  $\vec{v}$ )  $\nabla F(x_0, y_0, z_0)$ . Therefore, the

general equation of this tangent plane is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

## 14.6.6: Significance of the Gradient Vector

Here are some fun properties of the gradient vector.

Let  $f$  be a differentiable function of two or more variables and suppose that  $\nabla f(\mathbf{x}) \neq \mathbf{0}$ .

- The directional derivative of  $f$  at  $\mathbf{x}$  in the direction of a unit vector  $\mathbf{u}$  is given by  $D_{\mathbf{u}}f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u}$ .
- $\nabla f(\mathbf{x})$  points in the direction of the maximum rate of increase of  $f$  at  $\mathbf{x}$ , and that maximum rate is  $|\nabla f(\mathbf{x})|$ .
- $\nabla f(\mathbf{x})$  is perpendicular to the level curve or level surface of  $f$  through  $\mathbf{x}$ .

This lets do some cool things like plot the steepest rate of ascent up a contour map.

Also, if you're interested in machine learning or in compsci, check out [GRADIENT DESCENTS](#). You'll be learning that soon enough. Have fun!

## 15.1: Double Integrals over Rectangles

### 15.1.1: Volume and Double Integrals

Definite integrals in functions  $f(x)$  can be used to find areas. Similarly, **double integrals** can be used to find volumes.

Consider a function of two variables  $f$  defined on a rectangle  $R = [a, b] \times [c, d] = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, c \leq y \leq d\}$ . That  $\times$  is not a cross product, by the way. Geometrically, this means the function's domain is the set of all points  $(x, y)$  in a rectangular field bound by  $[a, b]$  and  $[c, d]$ .

Let the solid in that rectangular field bounded by the rectangle and the function  $z = f(x, y)$  be called  $S$ .

If we split up  $R$  into lots of smaller rectangles in a grid pattern, we notice we can approximate the volume of  $S$  by summing the volumes of all the smaller rectangles, ergo

$$V \approx \sum_{i=1}^m \sum_{j=1}^n f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$$

Intuitively, this means that for the  $i$ -th 'column' and  $j$ -th 'row' in the grid of rectangles, we evaluate the volume and add it to the total. Remember the volume of a rectangular prism is its area times its height. From the definition of a definite integral in 2D, we know that we should take the limit:

$$V = \lim_{m, n \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^n f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$$

and if we recall Riemann sums, we can condense this into a familiar form.

$$\iint_R f(x, y) dA = \lim_{m, n \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^n f(x_{ij}^*, y_{ij}^*) \cdot \Delta A$$

if this limit exists.

Putting it all together,

If  $f(x, y) \geq 0$ , then the volume  $V$  of the solid that lies above the rectangle  $R$  and below the solid  $z = f(x, y)$  is

$$V = \iint_R f(x, y) \, dA$$

The double-sum version (without the limit) is called a **double Riemann sum**.

### 15.1.2: The Midpoint Rule

Recall the **midpoint rule** for single integrals. If we have an integral  $\int_a^b f(x) \, dx$ , we can approximate it pretty well by dividing the interval into  $n$  subintervals and then using the value of the function at the midpoint of each subinterval to estimate the area.

$$\int_a^b f(x) \, dx \approx \Delta x \sum_{i=1}^n f(x_i^*)$$

where  $\Delta x = \frac{b-a}{n}$  and  $x_i^*$  is the height of the midpoint of the  $i$ th subinterval, or  $x_i^* = a + (i - \frac{1}{2})\Delta x$ .

We can do something similar for functions of two variables.

$$\iint_R f(x, y) \, dA \approx \sum_{i=1}^m \sum_{j=1}^n f(\bar{x}_i, \bar{y}_j) \Delta A$$

where  $\bar{x}_i$  is the midpoint of  $[x_{i-1}, x_i]$   
and  $\bar{y}_j$  is the midpoint of  $[y_{j-1}, y_j]$

You will be provided values for  $m$  and  $n$  which represent how many subintervals you will approximate with.

Note that we can only interpret a double integral as a volume if  $f$  is **positive**.

### 15.1.3: Iterated Integrals

This is a technique we can use to directly evaluate a double integral.

In our volume double integrals, we have had the differential term  $dA$  when  $f$  is a function of  $x$  and  $y$ . However, if we use the notation  $\int_c^d f(x, y) \, dy$ , it signifies that we are holding  $x$  constant and integrating with respect to  $y$ . This process is called **partial integration with respect to  $y$** . Note similarity to partial differentiation.

Now this result will be a function of  $x$ , so therefore we can write

$$A(x) = \int_c^d f(x, y) \, dy$$

And let us integrate both sides from  $a$  to  $b$ :

$$\begin{aligned} \int_a^b A(x) \, dx &= \int_a^b \left[ \int_c^d f(x, y) \, dy \right] dx \\ \int_a^b \int_c^d f(x, y) \, dy \, dx &= \int_a^b \left[ \int_c^d f(x, y) \, dy \right] dx \end{aligned}$$

Note we must work from the inside out. The following is a more practical theorem for the purpose of turning a double integral into iterated integrals named **Fubini's Theorem**:

If  $f$  is continuous on the rectangle  
 $R = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$   
 then

$$\iint_R f(x, y) \, dA = \int_a^b \int_c^d f(x, y) \, dy \, dx = \int_a^b \left[ \int_c^d f(x, y) \, dy \right] dx$$

More generally, Fubini's Theorem is valid if we assume  $f$  is bounded on  $R$ ,  $f$  is discontinuous only on a finite number of smooth curves, and that the iterated integrals exist. The next theorem follows from the special case where we assume that

$f(x, y)$  can be factored into the form  $g(x)h(y)$ .

$$\iint_R g(x)h(y) \, dA = \int_a^b g(x) \, dx \int_c^d h(y) \, dy$$

where  $R = [a, b] \times [c, d]$

### 15.1.4: Average Value

We know (at least you may know, since I don't think this is the syllabus for some math courses) that the **average value** of a function of one variable  $f(x)$  is

$$f_{\text{avg}} = \frac{1}{b-a} \int_a^b f(x) \, dx$$

Similarly, the average value of a function of two variables  $f(x, y)$  over a rectangle  $R$  is

$$f_{\text{avg}} = \frac{1}{A(R)} \iint_R f(x, y) \, dA$$

where  $A(R)$  is the area of the rectangle.

## 15.2: Double Integrals over General Regions

### 15.2.1: General Regions

Consider a general region  $D$  on a Cartesian plane.  $D$  is called **bounded** if it can be enclosed in a rectangular region  $R$ . Suppose  $D$  is bounded. We want to integrate a function  $f$  over  $D$ . To do so, let us define a piecewise function  $F(x, y)$  on domain  $R$ :

$$F(x, y) = \begin{cases} f(x, y) & \text{if } (x, y) \text{ is in } D \\ 0 & \text{if } (x, y) \text{ is in } R \text{ but not in } D \end{cases}$$

If  $F$  is integrable over  $R$ , we define the **double integral of  $f$  over  $D$**  by

$$\iint_D f(x, y) \, dA = \iint_R F(x, y) \, dA \quad \text{where } F \text{ is given above.}$$

This works is because when we take  $F(x, y)$  at a point out of  $D$ , we get 0. Thus, taking an integral including extra area not included in  $D$  is fine, as long as the integral includes all of  $D$ , which is true for  $R$ .

However, it is not uncommon for  $f$  to have discontinuities at the boundary points at  $D$ . However, if  $f$  is continuous on  $D$  and the boundary curve of  $D$  is well-behaved (you don't need to worry about that in this course), then  $\iint_D f(x, y) \, dA$  will exist.

We call a plane region  $D$  **Type 1** if it lies between the graphs of two continuous functions  $g_1(x), g_2(x)$  where  $g_1$  and  $g_2$  are continuous on  $[a, b]$ . Picture an area between two curves from Calc 1. That's a Type 1 plane region.

We put the rectangle  $R$  around it to contain  $D$ , and we let  $F$  be the piecewise function stated above. Through Fubini's Theorem, we can say that

$$\iint_R F(x, y) \, dA = \int_a^b \int_c^d F(x, y) \, dy \, dx$$

Recall that the rectangle  $R$  spans  $[a, b] \times [c, d]$ .  $c$  and  $d$  are the  $y$ -components, and  $D$  is bounded by the two functions  $g_1$

and  $g_2$ . Also note how  $F(x, y) = 0$  if  $y$  goes out of these bounds. Therefore, we can rewrite the above in another format.

If  $f$  is continuous on a Type 1 region  $D$  described by

$$D = \{(x, y) \mid a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

then

$$\iint_R f(x, y) dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) dy dx$$

Now we consider plane regions of **Type 2**. These are planar regions bounded between two continuous functions of  $y$ . Yada yada, it's the same thing as Type 1.

If  $f$  is continuous on a Type 2 region  $D$  described by

$$D = \{(x, y) \mid c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\}$$

then

$$\iint_R f(x, y) dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y) dx dy$$

Note the  $dx, dy$  have swapped places.

### 15.2.2: Changing the Order of Integration

We know we can state a double integral as an iterated integral in two different orders,  $dy dx$  and  $dx dy$ . But, sometimes, one order is way harder than the other or just impossible.

If you are provided an iterated integral, you can convert it back to a double integral and study its behavior to swap its order. There is no strict theorem to do this, as it is helpful to have a graph or sketch to be able to do it. However, the broad strokes are:

1. Turn your iterated integral back into a double integral.
  1. Don't forget your bounds:  $R = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$ .
2. Study the behavior of the graph. Your goal is to change the bounds of  $x$  and  $y$  such that it describes the same region  $R$ , but expresses it in the opposite order.
  1. Understand your inner variable and outer variable. If you want  $dx$  on the inside, you are looking for vertical strips that go bottom to top, and vice versa.
    1. Think about this: for a fixed value of the new outer variable, what are the smallest and largest values of the inner variable?
3. Write the new integral with the new bounds.
  1. If the region's lower bound for  $y$  is a single equation  $y = g_1(x)$  and the upper bound is  $y = g_2(x)$ , flipping gives limits on  $x$  from the left-most  $x$ -value to the right-most  $x$ -value of the *same* region, and inside limits on  $y$  that might now be functions of  $x$ . Works vice versa, of course.
4. Retry the integral in the new order.

### 15.2.3: Properties of Double Integrals

A lot of the properties of single integrals also apply to double integrals. Basically, everything works except for sign flipping:

$$\iint_D f(x) dx = - \iint_D f(x) dx$$

The single-integral property  $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$  has an analogue since we're in more dimensions now.

$$\iint_D f(x, y) dA = \iint_{D_1} f(x, y) dA + \iint_{D_2} f(x, y) dA$$

We can use this to evaluate regions not strictly of Types 1 or 2: For example, if a region crosses back over itself, like the

region between two circles, we can split it into multiple double integrals to solve it.

$$\iint_D 1 \, dA = A(D)$$

That's the area of  $D$ .

We can combine multiple properties to prove the following:

If  $m \leq f(x, y) \leq M$  for all  $(x, y)$  in  $D$ , then

$$m \cdot A(D) \leq \iint_D f(x, y) \, dA \leq M \cdot A(D)$$

Basically, if your plane in three dimensions can be bounded by two flat planes  $z = m, z = M$ , then the volume of the solid under your plane is always more than the volume of the solid bounded above by  $m$  and always less than the volume of the solid bounded above by  $M$ . This is pretty intuitive if you think about it.

Extra

Behaviors of Functions

Below is a table of various common functions/structures you may see. Note that this table counts convergence of **sequences**. It may be useful for series, i.e the Test for Divergence, but this mainly tracks what happens as  $x \rightarrow \infty$  in  $\lim_{x \rightarrow \infty} f(x)$ .

Function	$x \rightarrow \infty$	$x \rightarrow 0^+$	$x \rightarrow -\infty$
$c$	$c$	$c$	$c$
$x$	$\infty$	$0$	$-\infty$
$n^x$ (for $n > 1$ )	$\infty$	$1$	$0$
$n^{-x}$ (for $n > 1$ )	$0$	$1$	$\infty$
$\ln(x)$	$\infty$	$-\infty$	Undefined
$x!$	$\infty$	$1$	Undefined
$\frac{x^n}{n!}$	$\infty$	$0$	$\pm\infty$ (depends on $n$ )
$\frac{1}{x}$	$0$	Undefined	$0$
$r^x$ ( $r > 1$ )	$\infty$	$1$	$0$
$r^x$ ( $r = 1$ )	$1$	$1$	$1$
$r^x$ ( $0 < r < 1$ )	$0$	$1$	$\infty$
$r^x$ ( $r \leq -1$ )	Undefined (oscillates)	Undefined (oscillates)	Undefined (oscillates)
$\sin x$	Undefined ( $-1 \leftrightarrow 1$ )	$0$	Undefined( $-1 \leftrightarrow 1$ )
$\cos x$	Undefined ( $-1 \leftrightarrow 1$ )	$1$	Undefined ( $-1 \leftrightarrow 1$ )
$\tan x$	Undefined ( $-\infty \leftrightarrow \infty$ )	$0$	Undefined ( $-\infty \leftrightarrow \infty$ )
$\arcsin x$	Undefined ( $-1 \rightarrow 1$ )	$0$	Undefined ( $-1 \rightarrow 1$ )
$\arccos x$	Undefined ( $-1 \rightarrow 1$ )	$\frac{\pi}{2}$	Undefined ( $-1 \rightarrow 1$ )
$\arctan x$	$\frac{\pi}{2}$	$0$	$-\frac{\pi}{2}$
$\frac{\sin x}{x}$	$0$	$1$	$0$
Rational ( $n = d$ )	$\frac{\text{leading coeff of num}}{\text{leading coeff of den}}$	Constant	$\frac{\text{leading coeff of num}}{\text{leading coeff of den}}$
Rational ( $n > d$ )	$\pm\infty$	Constant	$\pm\infty$
Rational ( $n < d$ )	$0$	Constant	$0$

Scripts



Since Obsidian is a janky program especially when it comes to LaTeX, I had to use some regex scripts to format everything properly. For example, since Obsidian is an Electron app (you can even open the dev console like a webpage using **Ctrl - Shift - I**) and MathJax is also a janky mess sometimes, having lone double dollar signs  $$$$  (which escape blocks of LaTeX-formatted math) on lines will break the compiler.

Too bad that my Latex Suite shortcuts do that automatically.

So, I had to write this Python script to remove each of them using a neat little method and some regex, as well as some path patching so I can run it from the command line as part of an automated process.

```
import re
import os

def fix(text):
    pattern = re.compile(r'(?<!\)\)\$\\$(.*?)\\$\\$', re.DOTALL)

    def repl(match):
        content = match.group(1)

        lines = content.splitlines()

        while lines and lines[0].strip() == '':
            lines.pop(0)
        while lines and lines[-1].strip() == '':
            lines.pop()

        new_content = ' '.join(line.strip() for line in lines)

        return f'$$drLf{new_content}$$'

    fixed_text = pattern.sub(repl, text)

    return fixed_text

sdir = os.path.dirname(os.path.abspath(__file__))

npath = os.path.join(sdir, "note.txt")
fpath = os.path.join(sdir, "finish.txt")

with open(npath, "r") as f:
    a = f.read()

    with open(fpath, "w+") as f1:
        f1.write(fix(a))
```

I also wasn't happy with how the spacing between headings looked when I exported it to PDF, so I wrote a simple script to add extra lines.

```
import re

with open("output.txt", "r+") as f:
    text = f.read()
    # REMEMBER TO REMOVE THE HASHTAG FROM THE TITLE BEFORE RUNNING THIS!!!!
    newtext = re.sub("(\\r\\n|\\r\\n#)", "\\n\\n#", text)
    newertext = re.sub(r"(.*)", r"\1\n", newtext)
    ab = re.sub(r"\\r\\n|\\r|\\n(?<!\{\\}\n)(([a-z])\\n)(?!\\})", r"\\n$\\text{(\1)}$", text)
    f.truncate(0)
    f.seek(0)
    f.write(ab)
```

If you choose to make edits and notice your compiled PDF is broken or corrupted in some way, run it through these and see if it fixes things.

Also, I export this guide in particular with the settings **A3 paper, 18/16/12/12 mm margins top/bottom/left/right**. This is because some of the equations go off the page if I use A4 paper.

## Special Thanks

I want to express my thanks to everything that makes it possible for me to be making these guides.

Thank you to **Obsidian**, the best note-taking app ever. Being able to mess around with configs, add custom CSS snippets, and edit my files as raw `.md` files is instrumental to my process. Also, thank you to the creators of the theme **Typomagical**, my favorite Obsidian theme that makes writing these notes a pleasure to do.

I want to say that without the plugin **Quick Latex** (and now, **Latex Suite**), I would probably still be halfway through writing the Linear Algebra Study Guide. The inline expansion macros, as well as Latex Suite's complex regex/Javascript controllable macros, let me finish a giant block of equations in one minute rather than ten.

**Better Export PDF** also provides the perfect way to export a PDF in this program, adding headers/footers, custom margins, and smarter logic when it comes to rendering a Markdown-LaTeX hybrid note into a PDF. To a lesser extent, I want to wholeheartedly thank **Admonition** for cool callouts, **Advanced Tables** for making Markdown tables less of a drag, and **Mehrmaid** for adding the functionality of using MathJax inside Mermaid charts, which can't have been easy.

Thank you to **Syncthing** for automatically syncing my files from my laptop to PC: I tried to use Git for this. Don't. Use Syncthing.

## Donations

If you want to give a poor and starving university student who also has saved over thirty poor orphans (who were also starving by the way) from blazing homes a pittance, you may do so [HERE](#).

You have no obligation to do so, of course, it is just a token of your grat- oh hey hear that? The meowing? That's the fifty small kittens. That I also saved. From burning homes.

You still have no obligation to donate, by the way, just letting you know. :D

Thank you for using the guide!