# Calculus II (MTH 240)

Adam Szava

January 2021

## Introduction

This document is my compilation of notes from Calculus 2, mostly from my professor's lectures, but also from the OpenStax textbook: Calculus, Volume 2. Chapter headings match that textbook. The final chapter is from the OpenStax textbook: Calculus, Volume 3.

## 3.1 Integration by Parts

Integration by parts is used to integrate a product of functions. You treat the two functions being multiplied in the integrand both as functions of $x$ called $u(x)$ and $v(x)$. So in general:

$$\frac{d}{dx}uv = u'v + uv'$$

We can then integrate both sides with respect to $x$:

$$\int \frac{d}{dx}uv \ dx = \int u'v \ dx + \int uv' \ dx$$

On the left side the integration and differentiation cancel each other out.

$$uv \ = \int u'v \ dx + \int uv' \ dx$$

We then solve for either integral:

$$\int uv' \ dx = uv \ - \int u'v \ dx \tag{1}$$

Which can be simplified to:

$$\int u \ dv = uv - \int v \ du \tag{2}$$

Which is **not** to be interpreted as the integral of $u$ with respect to $v$, rather it is a way to remember the formula when actually solving problems. For a better understanding, consider equation (1).

If you were to be evaluating a definite integral, then the formula would not change, the only difference is that you have to evaluate the $uv$ term between the bounds, as in:

$$\int_a^b uv' \ dx = [uv]|_{x=a}^{b} - \int_a^b u'v \ dx \tag{3}$$

## 3.2 Trigonometric Integrals

When an integral is composed of the product of two trigonometric functions, there is occasionally an algorithmic method to solve them using identities.

Notable identities include:

$$sin^2(x) + cos^2(x) = 1 \tag{4}$$

$$cos(2x) = 2cos^2(x) - 1 \tag{5}$$

$$cos(2x) = 1 - 2sin^2(x) \tag{6}$$

$$sin(2x) = 2sin(x)cos(x) \tag{7}$$

$$sec^2(x) = 1 + tan^2(x) \tag{8}$$

$$sin(a)cos(b) = \frac{1}{2}(sin(a-b) + sin(a+b)) \tag{9}$$

$$cos(a)cos(b) = \frac{1}{2}(cos(a-b) + cos(a+b)) \tag{10}$$

$$sin(a)sin(b) = \frac{1}{2}(cos(a-b) - cos(a+b)) \tag{11}$$

There are two specific cases that we have a strategy for:

$$1) \int cos^j(x) \ sin^k(x) \ dx$$

$$2) \int sec^j(x) \ tan^k(x) \ dx$$

...where in both cases $j, k \in \mathbb{Z}^+$

**1)** In the first case there are 3 configurations that have a clear next step:

- If $j$ is odd $\rightarrow$ use (4) to eliminate all but one $cos(x)$, then u-sub $u = sin(x)$
- If $k$ is odd $\rightarrow$ use (4) to eliminate all but one $sin(x)$, then u-sub $u = cos(x)$
- If $j, k$ are even $\rightarrow$ use (5) and (6) to eliminate all $sin^2(x)$ and $cos^2(x)$

**2)** In the second case there are 3 configurations that have a clear next step:

- If $k$ is odd $\rightarrow$ use (8) to eliminate all but one $tan(x)$ then u-sub $u = sec(x)$
- If $j > 0$ is even $\rightarrow$ use (8) to eliminate all but two $tan(x)$ then u-sub $u = tan(x)$

- If $j = 0 \rightarrow$ use (8) to include one $sec^2(x)$ then split the integral.

In all these cases, the strategy will result in an integral that is easier to deal with.

Furthermore, there are two formulas for two other specific cases, called the **reduction formulas**. As the name suggests, these formulas equate the integral to another expression that contains the integral with a reduced exponent. They each pertain to one of the following integrals:

$$1)\int sec^n(x)\ dx$$

$$2)\int tan^n(x)\ dx$$

The following are the full reduction formulas:

$$\int sec^n(x)\ dx = \frac{1}{n-1}sec^{n-2}(x)\ tan(x) + \frac{n-2}{n-1}\int sec^{n-2}(x)\ dx$$

$$\int tan^n(x)\ dx = \frac{1}{n-1}tan^{n-1}(x) - \int tan^{n-2}(x)\ dx$$

There are more reduction formulas that are useful, but not mentioned in this chapter:

$$\int sin^n(x)\ dx = -\frac{1}{n}cos(x)\ sin^{n-1}(x) + \frac{n-1}{n}\int sin^{n-2}(x)dx$$

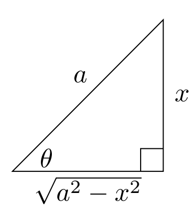
$$\int cos^n(x)\ dx = \frac{1}{n}cos^{n-1}(x)\ sin(x) + \frac{n-1}{n}\int cos^{n-2}(x)dx$$

## 3.3 Trigonometric Substitution

Trigonometric substitutions help us with integrals as follows:

| See this... | ...use this. |
| --- | --- |
| $\sqrt{a^2 - x^2}$ | $x = a\ sin(\theta)$ |
| $\sqrt{a^2 + x^2}$ | $x = a\ tan(\theta)$ |
| $\sqrt{x^2 - a^2}$ | $x = a\ sec(\theta)$ |

Essentially you are using the fact that each of those square roots are actually the solutions to a side of a right angled triangle using Pythagorean Theorem. This allows you to introduce a new variable $\theta$. Notice in the first case:



It is a reasonable thing to say given $\sqrt{a^2 - x^2}$ that:

$$sin(\theta) = \frac{x}{a}$$

... and that:

$$\theta = sin^{-1}\left(\frac{x}{a}\right)$$

Since $\theta$ is in the output of $sin^{-1}(\theta)$ it must be in it's range which is as follows:

$$-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

All of these facts combined are what let us do the trigonometric substitution, and make it useful. The rest of the substitutions have similar explanations. Here are the three main strategies used...

## $x = a\ sin(x)$ substitution for $\sqrt{a^2 - x^2}$

1. Sub $x = a\ sin(\theta),\ dx = a\ cos(\theta)\ d\theta$, for $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$
2. Simplify using $1 - sin^2(\theta) = cos^2(\theta)$
3. Simplify radical.
4. Integrate using strategies from section 3.2.
5. Sub $x$ back in by $\theta = sin^{-1}(\frac{x}{a})$ and simplify.

## $x = a\ tan(x)$ substitution for $\sqrt{a^2 + x^2}$

1. Sub $x = a\ tan(\theta), dx = a\ sec^2(\theta)\ d\theta$, for $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$
2. Simplify using $1 + tan^2(\theta) = sec^2(\theta)$
3. Simplify radical.
4. Integrate using strategies from section 3.2.
5. Sub $x$ back in by $\theta = tan^{-1}(\frac{x}{a})$ and simplify.

## $x = a\ sec(x)$ substitution for $\sqrt{x^2 - a^2}$

1. Sub $x = a\ sec(\theta),\ dx = a\ sec(\theta)tan(\theta)\ d\theta$, for $0 \leq \theta \leq \pi,\ \theta \neq \frac{\pi}{2}$
2. Simplify using $sec^2(\theta) - 1 = tan^2(\theta)$
3. Simplify radical.
4. Integrate using strategies from section 3.2.
5. Sub $x$ back in by $\theta = sec^{-1}(\frac{x}{a})$ and simplify.

## 3.4 Partial Fraction Decomposition

This integration technique is used to integrate *rational functions* in the form $\frac{P(x)}{Q(x)}$. Given the following integral:

$$\int \frac{3x}{x^2 - x - 2} dx$$

... it is too difficult to integrate by inspection, so we can convert it to an equivalent expression:

$$\int \left( \frac{1}{x+1} + \frac{2}{x-2} \right) dx$$

... which is relatively easy to integrate.

*Partial fraction decomposition* is the method by which we turn the first integral into the second.

In words, the process goes like this:

1. Divide the polynomials using polynomial long division if possible.
2. Factor the denominator into $n$ linear or $m$ quadratic factors.
3. Equate the original rational function to a certain sum of fractions.
4. Integrate the sum of fractions instead.

The complications come in that third step where we equate the rational function to the sum of other fractions. In general, to do this you take all the factors of the denominator and work from there. In each of the following fractions, the variables on the numerator are all unique. For each factor,

→ if it is linear, make a fraction in the form:

$$\frac{A}{(ax+b)}$$

→ if it is linear and appears twice, make to fractions in the form:

$$\frac{A}{(ax+b)} + \frac{B}{(ax+b)^2}$$

*Note: this extends to a linear factor appearing n times.*

→ if it is an irreducible quadratic factor, make a fraction in the form:

$$\frac{Ax+B}{(ax^2+bx+c)}$$

→ if it is an irreducible quadratic factor appearing twice, make a fraction in the form:

$$\frac{Ax+B}{(ax^2+bx+c)} + \frac{Cx+D}{(ax^2+bx+c)^2}$$

*Note: this extends to an irreducible quadratic factor appearing n times.*

This process is relatively difficult to understand in words, and so I will be including an example in this section:

## Example

$$\int \frac{x^2}{x^3 - x^2 + 4x - 4} dx$$

$$\frac{x^2}{x^3 - x^2 + 4x - 4} = \frac{x^2}{(x-1)(x^2+4)}$$

$$\frac{x^2}{(x-1)(x^2+4)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+4}$$

$$\frac{x^2(x-1)(x^2+4)}{(x-1)(x^2+4)} = \frac{A(x-1)(x^2+4)}{x-1} + \frac{(Bx+C)(x-1)(x^2+4)}{x^2+4}$$

$$x^2 = A(x^2+4) + (Bx+C)(x-1)$$

$$x^2 = Ax^2 + 4A + Bx^2 - Bx + Cx - C$$

$$x^2 = (A+B)x^2 + (C-B)x + (4A-C)$$

Which tells us that:

$$A + B = 1$$

$$C - B = 0$$

$$4A - C = 0$$

Solving using linear algebra techniques gives us:

$$A = \frac{1}{5}, B = \frac{4}{5}, C = \frac{4}{5}$$

Which means:

$$\frac{x^2}{(x-1)(x^2+4)} = \frac{1/5}{x-1} + \frac{(4/5)x + (4/5)}{x^2+4}$$

$$\int \frac{x^2}{(x-1)(x^2+4)} dx = \int \frac{1/5}{x-1} + \frac{(4/5)x+(4/5)}{x^2+4} dx$$

$$= \frac{1}{5}\ Ln|x-1| + \ \frac{4}{5}\int \frac{x+1}{x^2+4} dx$$

$$= \frac{1}{5}\ Ln|x-1| + \ \frac{4}{5}\int \frac{x}{x^2+4} + \frac{1}{x^2+4} dx$$

$$= \frac{1}{5}\ Ln|x-1| + \ \frac{4}{5}\int \frac{x}{x^2+4} dx + \frac{2}{5}\ tan^{-1}\left(\frac{x}{2}\right)$$

$$= \frac{1}{5}\ Ln|x-1| + \ \frac{2}{5}\ Ln\left|x^2+4\right| + \ \frac{2}{5}\ tan^{-1}\left(\frac{x}{2}\right) + C$$

## 3.7 Improper Integrals

An **improper integral** is an integral that either has $\infty$ or $-\infty$ as a limit of integration, or an integral that contains a discontinuity between its limits of integration.

Define:

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

... where $f(x)$ is continuous on $[a, \infty)$.

Similarly, define:

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

... where $f(x)$ is continuous on $(-\infty, a]$.

We can combine these two facts to evaluate a function like:

$$\int_{-\infty}^\infty f(x)\ dx = \int_{-\infty}^a f(x)\ dx + \int_a^\infty f(x)\ dx = \lim_{t\to-\infty} \int_t^a f(x)\ dx + \lim_{r\to\infty} \int_a^r f(x)\ dx$$

... where $a$ is just some real number, typically $a = 0$ unless there's some reason to choose another value. Essentially, in order to evaluate an integral with infinity as one or both of it's limits of integration, you replace the infinity with a variable, integrate in terms of that variable and then take the limit as that variable goes to infinity. **Given that all these limits exist**. If any one of these limits do not exist, the integral is said to **diverge**, otherwise, if the limits *do* exist, the integral is said to **converge**.

If your integrand has a discontinuity between the limit of integration, then you must integrate around that point, that is:

Let $f(x)$ be discontinuous at $c$, and continuous otherwise on $[a, b]$ where $c \in [a, b]$, then:

$$\int_a^c f(x)\ dx = \lim_{t\to c^-} \int_a^t f(x)\ dx$$

Similarly:

$$\int_c^b f(x)\ dx = \lim_{t\to c^+} \int_t^b f(x)\ dx$$

Together, we can conclude a formula for integrating a function with a discontinuity within it's limits of integration. Using the same $f(x)$ as above:

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx = \lim_{t\to c^-} \int_a^t f(x)\ dx + \lim_{t\to c^+} \int_t^b f(x)\ dx$$

**Comparison Theorem**

If:

$$0 \le f(x) \le g(x) \text{ for all } x \ge a$$

Then:

- $$\text{if } \int_a^\infty f(x)\ dx = \infty \text{ then } \int_a^\infty g(x)\ dx = \infty$$

- $$\text{if } \int_a^\infty g(x)\ dx = L \text{ then } \int_a^\infty f(x)\ dx \le L \text{ (and converges)}$$

This theorem is useful to show that a smaller function converges, or that a bigger function diverges. You typically choose the function to compare with, and that chosen function's divergence/convergence is ideally easy to evaluate.

## 4.3 Separable Differential Equations

This section begins with a lot of definitions to familiarize with terminology relating to differential equations.

A **Differential Equation** (DE) is an equation in variables $x, y, y', y'', y'''$, etc, where $y$ is a function of $x$.

A **solution to a DE** is a function $f(x)$ where subbing in $y = f(x)$ makes an equation which is always true.

The **order of a DE** is $n$ where the $n^{th}$ derivative of $y$ is the highest order term which appears.

The **general solution to a DE** is a parameterized family of functions.

An **initial value problem** consists of a DE with a specific condition. A solutions to this type of problem is a solution to the DE that satisfies the conditions.

A DE is **separable** if it can be rearranged into the form:

$$y' = f(x)\ g(y)$$

As in you can *separate* $x$ and $y$. In order to solve a separable differential equation, you follow this formula:

$$\int \frac{1}{g(y)}\ dy = \int f(x)\ dx$$

Do the integration, and rearrange to isolate $y$ if possible.

## 4.5 Linear Differential Equations

By definition, a **linear differential equation** (LDE) is a DE which can be written in the form:

$$y' + p(x)y = q(x)$$

Notably the coefficient on the $y'$ term is 1. In order to find a solution to this type of DE, we must first define an **Integrating Factor**.

In general (not just for LDEs) an *Integrating Factor* is a function usually denoted $I(x)$ which, when multiplied on both sides of a DE, allows both sides to be integrated. For LDEs, the *Integrating Factor* will always be in the form:

$$I(x) = e^{\int p(x)dx}$$

Where $p(x)$ is the same $p(x)$ in the LDE. This helps us solve the LDE because of the following:

$$y' + p(x)y = q(x)$$

$$e^{\int p(x)dx}y' + e^{\int p(x)dx}p(x)y = e^{\int p(x)dx}q(x)$$

$$\int e^{\int p(x)dx}y' + e^{\int p(x)dx}p(x)y\ dx = \int e^{\int p(x)dx}q(x)\ dx$$

Notice the the function in the integral on the left side will always be the result of a product rule differentiation, as in:

$$\int (e^{\int p(x)dx}y)'\ dx = \int e^{\int p(x)dx}q(x)\ dx$$

So:

$$e^{\int p(x)dx}y = \int e^{\int p(x)dx}q(x)\ dx$$

$$y = \frac{1}{e^{\int p(x)dx}} \int e^{\int p(x)dx}q(x)\ dx$$

And now we have the solution to the LDE.

## 5.1 Sequences

A sequence of numbers is an infinitely long list of numbers that has an initial element. An example of this is:

$$2, 4, 8, 16, 32, 64, 128, ...$$

This sequence of numbers begins with 2, and has infinitely many terms. One notation used commonly to describe sequences is as follows:

$$\{a_i\}_{i \geq a}$$

Where each $a_i$ is the $i^{th}$ term in the sequence. $i$ is called the index variable, and $a$ is the index lower bound, which can be any number (typically 1 or 0). Another form of the notation is as follows:

$$\{f(n)\}_{n\geq 1}$$

Where $f(n)$ is an explicit formula for calculating the $n^{th}$ term in the sequence. A **recurrence relation** is a formula that relates the $n^{th}$ term of the sequence to earlier terms. An example of this is the *Fibonacci Sequence* which is written as a recurrence relation as follows:

$$a_1 = 1 \quad a_2 = 1 \quad a_n = a_{n-1} + a_{n-2} \text{ for } n \geq 3$$

As you can see for a recurrence relation, you must explicitly define the starting position of the sequence.

An **Arithmetic Sequence** is a sequence where the difference between each term is some constant ($k$), and can be defined as:

$$a_1 = p$$

$$\text{As a recurrence relation: } a_n = a_{n-1} + k \text{ for } n \geq 2$$

$$\text{As an explicit formula: } a_n = (p - k) + nk$$

A **Geometric Sequence** is a sequence where the ratio between each term is some constant ($q$), and can be defined as:

$$a_0 = r$$

$$\text{As a recurrence relation: } a_n = a_{n-1}q \text{ for } n \geq 1$$

$$\text{As an explicit formula: } a_n = rq^n \text{ for } n \geq 0$$

If a question asks you to *solve* a recurrence relation, this means find an explicit formula for the sequence.

To solve a recurrence relation, you must either identify it as being an *arithmetic* or *geometric* sequence and then use the general forms as above. If the sequence does not match either of there structures, then write out the first few terms and try to find the patterns. In the case where you have to find the explicit formula yourself, you should then check the answer by subbing it in to the recurrence relation and showing $LHS = RHS$.

## Sequence Limits

Say $\{a_n\}_{n\geq 1}$ is a sequence, define its limit as follows:

$$\lim_{n\to\infty} a_n = L \text{ if as } n \text{ gets larger, } a_n \text{ gets as close as you want to L.}$$

If L exists, the sequence *converges*, if not it *diverges*.

In order to actually compute this limit, you take the explicit formula for the relation in the form $a_n = f(n)$ and then evaluate the following:

$$\lim_{n\to\infty} f(n) = \lim_{x\to\infty} f(x) = L$$

Let $\{a_n\}_{n\geq p}$ and $\{b_n\}_{n\geq q}$ be sequences, then if there is some $k$ where:

$$a_n = b_n \text{ for all } n \geq k$$

Then say that $\{a_n\}_{n\geq p}$ and $\{b_n\}_{n\geq q}$ are **eventually equal.** Essentially, if after some term, all the terms of the two sequences are equal, then you say that the two terms are *eventually equal*.
For two sequences $\{a_n\}_{n\geq p}$ and $\{b_n\}_{n\geq q}$ that are eventually equal, then:

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

## Limit Laws

- $$\text{if } a_n = c \text{ then } \lim_{n\to\infty} a_n = c$$
- $$\lim_{n\to\infty} ca_n = c \lim_{n\to\infty} a_n$$
- $$\lim_{n\to\infty} (a_n \pm b_n) = \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n$$
- $$\lim_{n\to\infty} (a_n b_n) = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n$$
- $$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n} \text{ given the denominator is not zero.}$$

Given that all these limits exist.

## Squeeze Theorem

Let $\{a_n\}, \{b_n\}, \{c_n\}$ be sequences where:

$$a_n \leq b_n \leq c_n \text{ eventually}$$

... then:

$$\text{if } \lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L \text{ then,}$$
$$\lim_{n\to\infty} b_n = L$$

### Terminology

The following is a list of definitions regarding sequences.
Say that the sequence $\{a_n\}$ is...

- **bounded above** if there is some $m$ where $a_n \leq m$ eventually.
  - show this by inspection.
- **bounded below** if there is some $m$ where $a_n \geq m$ eventually.
  - show this by inspection.
- **bounded** if it is both bounded above and below.
- **increasing** if $a_n \geq a_{n-1}$ eventually.
  - show that $a_n \geq a_{n-1}$ for all $n$, or for all $n \geq k$ for some $k$.
  - show that $f'(x)$ is always positive for all $n$, or for all $n \geq k$ for some $k$.
- **decreasing** if $a_n \leq a_{n-1}$ eventually.
  - show that $a_n \leq a_{n-1}$ for all $n$, or for all $n \geq k$ for some $k$.
  - show that $f'(x)$ is always negative for all $n$, or for all $n \geq k$ for some $k$.
- **monotone** if it is either *increasing* or *decreasing*.

**Monotone Convergence Theorem**: If $\{a_n\}$ converges, then it is bounded, and if $\{a_n\}$ is both monotone and bounded, then it converges.

## 5.2 Series

A series is the sum of terms up to some point in a sequence. An infinite series is the sum of all the terms in a sequence. One notation is:

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

The sequence of partial sums (sums up to some term $k$) is defined as:

$$\{\sum_{n=1}^{k} a_n\}_{k\geq 1} = \{S_1, S_2, S_3, S_4, ...\}$$

... where:

$$S_k = \sum_{n=1}^{k} a_n \text{ for } k \geq 1$$

Something to note is that sometimes you can define a formula for the partial sums of a series, $S_k = f(k)$, this is mentioned more later in this document. You can then define the sum of all the terms in an infinite sequence using the following:

$$\text{If } \lim_{k\to\infty} S_k = L, \text{ say that } \sum_{n=1}^{\infty} a_n = L$$

If this limit exists say that $\sum_{n=1}^{\infty} a_n$ converges, otherwise it diverges.

**A telescoping series** is a series where in each new term of the partial sum, some things cancel between it and the previous partial sum. In general the partial sum of this type of series is defined by:

$$S_k = a_0 - a_k$$

Important examples of series include:

- *The Harmonic Series:* $\sum_{n=1}^{\infty} \frac{1}{n}$ which diverges to $\infty$.
- *The Alternating Harmonic Series:* $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$ which converges.

**Properties of Convergent Series:**
Given that the infinite sums $\sum_n a_n$ and $\sum_n b_n$ are convergent, and $c \in \mathbb{R}$ then:

- $$\sum_n (a_n \pm b_n) = \sum_n a_n \pm \sum_n b_n$$
- $$\sum_n (ca_n) = c \sum_n a_n$$

## Arithmetic Series

For an arithmetic sequence with an explicit formula in the form:

$$dn + a$$

... where $a$ is the initial value, and $d$ is the common difference. The partial sum has a formula as follows:

$$\sum_{n=1}^{k} dn + a = S_k = \frac{k}{2}(a + a_k)$$

## Geometric Series

A geometric sequence is a sequence with an explicit formula in the form:

$$ar^n$$

... where $a$ is the initial value, and $r$ is the growth rate per term. Evaluating a geometric series depends on $a$ and $r$. Notably, if $a = 0$ then $\sum 0 = 0$ and so in all the following cases we are assuming that $a \neq 0$. Otherwise:

$$\sum_{n=1}^{k} ar^{n-1} = S_k = ka \text{ for r} = 1$$

$$\sum_{n=1}^{k} ar^{n-1} = S_k = \frac{a(1-r^k)}{1-r} \text{ for r} \neq 1$$

Evaluating this infinite series by taking the limit as $k \to \infty$ is as follows:

$$\lim_{k\to\infty} \frac{a(1-r^k)}{1-r} = \frac{a}{1-r} \text{ for } |r| < 1$$

In all other cases this series is divergent.

If you have the formula for the partial sums of a series, you can find $a_n$ by using the following formula:

$$a_n = S_n - S_{n-1}$$

## 5.3 Divergence and Integral Tests

*The divergence test* is the first test needed to decide whether a sequence converges or diverges.

$$\text{If } \lim_{n\to\infty} a_n \neq 0 \text{ or doesn't exist then } \sum_{n}^{\infty} a_n \text{ diverges.}$$

Note that if $\lim_{n\to\infty} a_n = 0$ then we learn nothing. The reason this is true is that for a series to converge, the absolute value of the amount you add to each partial sum must be getting smaller, and so the value of the sequence must approach 0 for the series to possibly converge.

*The integral test* is used for sequences with explicit formulas $f(x)$ where the function is positive, continuous, and decreasing for $[b, \infty]$ for some $b \in \mathbb{R}$, then:

$$\int_{b}^{\infty} f(x)\, dx \text{ converges if and only if } \sum_{n=b}^{\infty} a_n \text{ converges.}$$

Note that they do not necessarily converge to the same number. Also note that if you have a series that is all negative, you can still apply the integral test by multiplying all terms by (-1).

Another important series called a **p-series** is:

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

A p-series converges if $p > 1$ otherwise diverges.

### Approximating Series

It is usually very difficult if not impossible to find what a series converges to. Luckily we have a method of approximating the value a convergent series converges to, to any degree of accuracy. We can approximate:

$$S = \sum_{n=1}^{\infty} a_n \text{ by } S_k = \sum_{n=1}^{k} a_n$$

For any $S_k$, there is a section of the infinite sequence remaining that has not been taken into account with the partial sum, call this $R_k$.

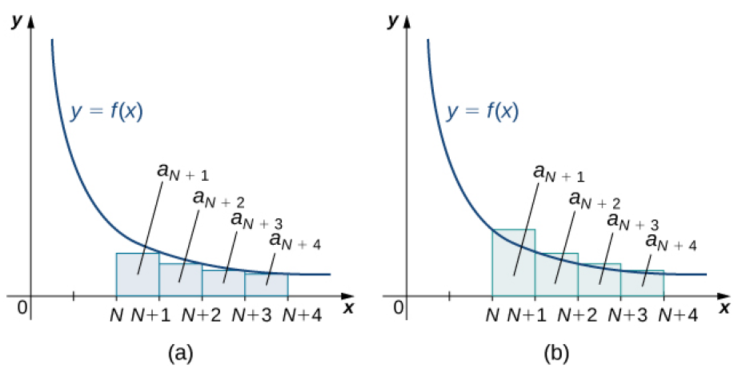
We can then say:

$$R_k = S - S_k$$

We want to understand how big $R_k$ is so that we know how accurate our approximation is. We can get a bound on it by using the following theorem:

$$S_k + \int_{k+1}^{\infty} f(x)dx < S < S_k + \int_{k}^{\infty} f(x)dx$$

We can see that the value of the infinite sum is bounded between an underestimate, and an overestimate. The following image helps illustrate this:




Subtracting $S_k$ from the inequality we get:

$$\int_{k+1}^{\infty} f(x)dx < R_k < \int_{k}^{\infty} f(x)dx$$

... and this is known as the remainder estimate.

## 5.4 Comparison Tests

This section introduces new tests you can use on a series to check for divergence/convergence. There are two types of comparison tests, *the comparison test*, and *the limit comparison test.* For these tests, you pick a $b_n$ that is easy to evaluate the convergence of and compare it to your $a_n$.

### Comparison Test

If $0 \leq a_n \leq b_n$ eventually then:

- $$\text{If } \sum_{n}^{\infty} b_n \text{ converges, then } \sum_{n}^{\infty} a_n \text{ also converges.}$$
- $$\text{If } \sum_{n}^{\infty} a_n \text{ diverges, then } \sum_{n}^{\infty} b_n \text{ also diverges.}$$

### Limit Comparison Test

**Case 1:** if $a_n, b_n > 0$ and:

$$\text{If } 0 < \lim_{n\to\infty} \frac{a_n}{b_n} < \infty \text{ and exists.}$$

Then either:

$$\text{...both } \sum_{n}^{\infty} a_n \text{ and } \sum_{n}^{\infty} b_n \text{ converge.}$$

$$\text{...or both } \sum_{n}^{\infty} a_n \text{ and } \sum_{n}^{\infty} b_n \text{ diverge.}$$

**Case 2:** if $a_n, b_n > 0$ and:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0 \text{ and } \sum_{n}^{\infty} b_n \text{ converges then } \sum_{n}^{\infty} a_n \text{ also converges.}$$

**Case 3:** if $a_n, b_n > 0$ and:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \infty \text{ and } \sum_{n}^{\infty} b_n \text{ diverges then} \sum_{n}^{\infty} a_n \text{ also diverges.}$$

## 5.5 Alternating Series

A series is alternating if it can be written in the following form:

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} (-1)^n b_n \text{ or } \sum_{n=0}^{\infty} (-1)^{n+1} b_n \text{ for } b_n > 0$$

Essentially, a series $a_n$ is alternating if it alternates between positive and negative terms, and the absolute value of each terms is its own sequence $b_n$. Whether the series is in the first form or the other depends on if the first term you are counting is positive or negative.

The **Alternating Series Test** is used to determine convergence and divergence of an alternating series, and is defined as follows:

### Alternating Series Test

$$\text{Let } a_n \text{ be an alternating series and } b_n = |a_n|$$

$$\text{If } \{b_n\} \text{ is decreasing and } \lim_{n\to\infty} b_n = 0, \text{ then } \sum_{n=0}^{\infty} a_n \text{ converges.}$$

### Alternating Series Estimation Theorem

This theorem is used to bound the remainder $R_k$ when approximating an alternating series. Say that:

$$\sum_{n=0}^{\infty} a_n = S \text{ and } S = S_k + R_k$$

If $b_n = |a_n|$, $\sum_{n=0}^{\infty} a_n$ is alternating, $b_n$ decreases and $\lim_{n\to\infty} b_n = 0$ then:

$$|R_k| \leq b_{k+1}$$

The following are some definitions relating to alternating series...
Let $\sum_{n=0}^{\infty} a_n$ be an alternating series:

- Say that $\sum_{n=0}^{\infty} a_n$ is absolutely convergent if $\sum_{n=0}^{\infty} |a_n|$ converges.
- Say that $\sum_{n=0}^{\infty} a_n$ is conditionally convergent if $\sum_{n=0}^{\infty} a_n$ converges but $\sum_{n=0}^{\infty} |a_n|$ diverges.

Importantly, if $\sum_{n=0}^{\infty} a_n$ is absolutely convergent, then $\sum_{n=0}^{\infty} a_n$ is convergent. These ideas are easily described in the following table:

| . | $\sum_{n=0}^{\infty} a_n$ converges | $\sum_{n=0}^{\infty} a_n$ diverges |
|---|---|---|
| $\sum_{n=0}^{\infty} b_n$ converges | Absolutely convergent | Not possible |
| $\sum_{n=0}^{\infty} b_n$ diverges | Conditionally convergent | Divergent |

The following two facts are true of alternating series:

1. If a series is absolutely convergent, then rearranging its terms will not change it's sum.
2. If a series is conditionally convergent to a positive sum, then whatever possible real number you want the series to converge to, there is a rearrangement for it.

## 5.6 Ratio and Root Test

This section will cover the final two tests of convergence and divergence, as well as an overview of a general strategy for using these tests.

## Ratio Test

Let $\sum_{n=0}^{\infty} a_n$ be a series, and:

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

If:

- $$0 \le L < 1 \text{ then } \sum_{n=0}^{\infty} a_n \text{ is absolutely convergent}$$
- $$L > 1 \text{ or } L = \infty \text{ then } \sum_{n=0}^{\infty} a_n \text{ is divergent.}$$
- $$L = 1 \text{ then the result is inconclusive.}$$

## Root Test

Let $\sum_{n=0}^{\infty} a_n$ be a series, and:

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

If:

- $$0 \le L < 1 \text{ then } \sum_{n=0}^{\infty} a_n \text{ is absolutely convergent}$$
- $$L > 1 \text{ or } L = \infty \text{ then } \sum_{n=0}^{\infty} a_n \text{ is divergent.}$$
- $$L = 1 \text{ then the result is inconclusive.}$$

Notice both of these tests have the same conditions on convergence/divergence when evaluating their results.

## General Strategy

This is a general strategy to determine convergence/divergence of an infinite series:

1. Is it a series we already know? (telescoping, p-series, geometric)
2. Is it alternating?

- alternating series test

3. Is it structurally similar to a p-series or a geometric series? (algebraic or geometric)
   - comparison test
   - limit comparison test
4. Does it involve a factorial or exponential?
   - ratio test
   - root test
5. Try the divergence test.
6. See if the integral test applies.
   - integral test

## 6.1 Power Series and Functions

A power series is essentially a power function of $x$ inside an infinite series with $n$ as the index variable. In general, a **power series** centred at $a \in \mathbb{R}$ is defined as follows:

$$\sum_{n=0}^{\infty} c_n (x-a)^n \text{ where each } c_n \in \mathbb{R}$$

Notice that if I center this power function at $a = 0$, and explicitly write out each term, the result is a familiar function type, namely a *polynomial*, with coefficients $c_n$, as in:

$$\sum_{n=0}^{\infty} c_n x^n = c_0 x^0 + c_1 x^1 + c_2 x^2 + c_3 x^3 \dots$$

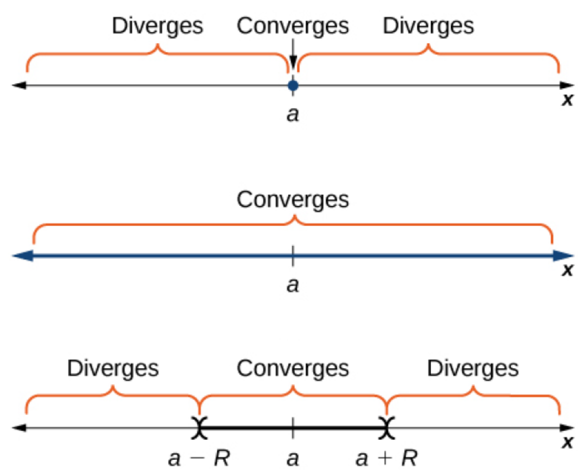
A natural question to ask about a power series is for what values of $x$ does this infinite series converge? This is where the concept of a *radius of convergence* comes in. There are three possibilities for convergence:

1. The series converges at $x = a$, and diverges at $x \neq a$ (radius of convergence of 0)
2. The series always converges (radius of convergence of $\infty$)
3. There is an $R \in \mathbb{R}$ such that the series either converges if:
   - $a - R < x < a + R$
   - $x < a - R$ or $x > a + R$

Each of these cases have their own respective *intervals of convergence*:

1. $[a, a]$
2. $(-\infty, \infty)$
3. $(a - R, a + R)$ or $[a - R, a + R)$ or $(a - R, a + R]$ or $[a - R, a + R]$

$R$ is always used to signify *radius of convergence*, while $I$ is always used to signify *interval of convergence*, which both relate to the same idea. The concept is called a *radius* of convergence because typically the values of $x$ that result in convergence are some distance $R$ to either side of $a$, the value the power function is centered at.




Power functions are most useful when turning a regular function $f(x)$ into an infinite polynomial, as in a power series. The first example we have is if you have an expression in the form:

$$\frac{a}{1-r}$$

Then you can express this as an infinite geometric series, essentially using the rule we know from earlier sections but *backwards*, as in:

$$\frac{a}{1-r} = \sum_{n=0}^{\infty} ar^n$$

We say that the infinite series converges to the function $\frac{a}{1-r}$.

Typically, when solving for $R$ and $I$, you *begin with the ratio test* to test for the conditions of convergence. Always note that $x$ is a constant to the sum.

## 6.2 Properties of Power Series

Recall that the sum operator ($\sum$) is preserved under addition and scalar multiplication. By this logic,

$$\text{Suppose } \sum_{n=0}^{\infty} c_n(x-a)^n \text{ converges to } f(x) \text{ on } I$$

$$\text{...and } \sum_{n=0}^{\infty} d_n(x-a)^n \text{ converges to } g(x) \text{ on } I$$

Then:

1. $$\sum_{n=0}^{\infty}(c_n \pm d_n)(x-a)^n \text{ converges to } f(x) \pm g(x) \text{ on } I$$

2. $$\sum_{n=0}^{\infty} bx^m c_n(x-a)^n \text{ converges to } bx^n f(x) \text{ on } I\text{, for any } m \in \mathbb{Z}^+ \text{ and } b \in \mathbb{R}$$

3. $$\sum_{n=0}^{\infty} c_n(b(x-a)^m)^n \text{ converges to } f(bx^m) \text{ for all } bx^m \in I\text{, for any } m \in \mathbb{Z}^+ \text{ and } b \in \mathbb{R}$$

The *first property* is most useful in reverse, meaning you can combine two functions $f(x)$ and $g(x)$ into a single power function centered at the same value $a$ if they are added or subtracted. Note that if the intervals of convergence do not match between the two functions, then the shared interval between them is the new interval of convergence. The *second property* shows that both scalar multiples $b$, and multiples of $x$ can be moved outside the sum. The *third property* shows that a power function can be evaluated at more complex arguments.

### Reindexing

Reindexing is a method used to change some property of the index variable, commonly the starting value. For example say we have:

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

You may be tempted to say that this infinite sum is equal to:

$$\frac{1}{1-(\frac{1}{x})} \text{ for } |x| < 1$$

This is untrue however, because the index variable starts at $n = 1$ whereas this rule only applies for infinite series starting at $n = 0$. To apply this rule we need to reindex. Focus on the index:

$$n = 1 \tag{1}$$

... we want the right side to be 0, so subtract 1 from both sides:

$$n - 1 = 0 \tag{2}$$

... now we can define a new index $k$:

$$k = n - 1 \tag{3}$$

... which we can sub (3) into (2) and get:

$$k = 0 \tag{4}$$

... meaning $k$ is an index variable which starts at 0. We can then rearrange (3) to get that:

$$n = k + 1 \tag{5}$$

... and from here we sub (5) into our original power series for $n$, as in:

$$\sum_{n=1}^{\infty} \left(\frac{1}{x}\right)^n = \sum_{k=0}^{\infty} \left(\frac{1}{x}\right)^{k+1}$$

... we can now use the original rule we tried to use to solve this question after a few more simplifications:

$$\sum_{k=0}^{\infty} \left(\frac{1}{x}\right)^{k+1} = \frac{1}{x} \cdot \sum_{k=0}^{\infty} \left(\frac{1}{x}\right)^k = \frac{1}{x} \cdot \frac{1}{1 - (\frac{1}{x})} \text{ for } |x| < 1$$

$$\sum_{n=1}^{\infty} \left(\frac{1}{x}\right)^n = \frac{1}{x - 1} \text{ for } |x| < 1$$

## Cauchy Product

The Cauchy product is essentially a way for two infinite series to be multiplied together. Mathematically it is defined by:

$$\text{Suppose } \sum_{n=0}^{\infty} c_n x^n \text{ converges to } f(x) \text{ on } I$$

$$\text{...and } \sum_{n=0}^{\infty} d_n x^n \text{ converges to } g(x) \text{ on } I$$

Then the Cauchy Product is defined as:

$$\left(\sum_{n=0}^{\infty} c_n x^n\right)\left(\sum_{n=0}^{\infty} d_n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} c_k d_{n-k}\right) x^n$$

... and that result converges to $f(x) \cdot g(x)$ on $I$.
*Note:* the $\sum$ within the $\sum$ is sometimes given the symbol $e_n$.

The best way to think of it is the inner sum $e_n$ is the coefficient on the $x^n$ term, as in:

$$\left(\sum_{n=0}^{\infty} c_n x^n\right)\left(\sum_{n=0}^{\infty} d_n x^n\right) = \sum_{n=0}^{\infty} e_n x^n$$

Notice that $e_n$ is not a power function, and is also not an infinite sum, it is a finite combination of terms in the two sequences $c_n$ and $d_n$. To illustrate this here are the first few expanded terms:

$$f(x) \cdot g(x) = (c_0 d_0)x^0 + (c_0 d_1 + c_1 d_0)x^1 + (c_0 d_2 + c_1 d_1 + c_2 d_0)x^2 \dots$$

Since $e_n$ is a finite sum, you can usually find an expression to represent $e_n$ that is not a sum, which would allow you to write a concise form of the product. The following finite sum formulas are useful to achieve this:

$$\sum_{k=0}^{n} a^k = \frac{1 - a^{n+1}}{1 - a}$$

$$\sum_{k=0}^{n} k a^k = \frac{a}{(1-a)^2} \cdot (1 - (n+1)a^n + n a^{n+1})$$

$$\sum_{k=0}^{n} k^2 a^k = \frac{a}{(1-a)^3} \cdot ((1+a) - (n+1)^2 a^n + (2n^2 + 2n - 1)a^{n+1} - n^2 a^{n+2})$$

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}$$

$$\sum_{k=0}^{n} k^4 = \frac{n}{30} \cdot (n+1)(2n+1)(3n^2 + 3n - 1)$$

### Differentiation and Integration of Power Series

Just like elementary functions, power functions can be differentiated and integrated.

$$\text{Suppose } \sum_{n=0}^{\infty} c_n x^n \text{ converges to } f(x) \text{ on } I$$

If $f(x)$ is differentiable on $I$ then:

$$\sum_{n=1}^{\infty} c_n n x^{n-1} \text{ converges to } f'(x) \text{ on } I$$

If $f(x)$ is integrable on $I$ then:

$$D + \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1} \text{ converges to } \int f(x)\, dx \text{ on } I$$

For some arbitrary $D \in \mathbb{R}$.

### Uniqueness of Power Series

If $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges to $f(x)$ on I, and $\sum_{n=0}^{\infty} d_n (x-a)^n$ also converges to $f(x)$ on I, then:

$$c_n = d_n \text{ for all } n \geq 0$$

## 6.3 Taylor and Maclaurin Series

This is the quintessential Calculus 2 section, and debatably the hardest and most useful section in the course.

Ideally, we'd like to be able to represent any general function $f(x)$ as a power function. This would give us a new way to express those functions, and compute their derivatives and integrals.

Suppose we have a function $f(x)$, and we know it *can* be represented by a power function, as in:

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

Can we make an explicit formula for the sequence $c_n$?
We can use fact that if the power function were to represent $f(x)$, then the power function evaluated at $a$ should equal $f(a)$, as in:

$$f(a) = \sum_{n=0}^{\infty} c_n (a-a)^n = \sum_{n=0}^{\infty} c_n (0)^n = c_0(1) + c_1(0) + c_2(0) \ldots$$

$$f(a) = c_0$$

Notice we just found the first value of the sequence $c_n$. You may also notice that we are essentially saying that $0^0 = 1$ and while this is not generally true, we assume it to be true when dealing with these power functions. Carrying on, by the exact same logic the derivative of the power function evaluated at $a$ should give us $f'(a)$, as in:

$$f'(a) = \sum_{n=0}^{\infty} c_n n(a-a)^{n-1} = c_0(0)(0) + c_1(1)(1) + c_2(2)(0)\ldots$$

$$f'(a) = c_1$$

Again, we now have a value for $c_1$. Let's now continue this pattern for the second and third derivatives:

$$f''(a) = \sum_{n=0}^{\infty} c_n n(n-1)(a-a)^{n-2} = c_0(0)(-1)(0) + c_1(1)(0)(0) + c_2(2)(1)(1)\ldots$$

$$f''(a) = c_2(2) \implies c_2 = \frac{f''(a)}{2}$$

The next terms in the series would all contain 0. Notice that our potential formula for $c_n$ just got more complicated because of the 2, but let's look at the next entry to see if there is a pattern, recall that so far we have:

$$c_0 = f(a)$$

$$c_1 = f'(a)$$

$$c_2 = \frac{f''(a)}{2}$$

In general, so far it looks like the $n^{th}$ term of $c_n$ will include the $n^{th}$ derivative of $f$ at $a$, written as:

$$f^{(n)}(a)$$

So a preliminary explicit formula for $c_n$ could be: $c_n = f^{(n)}(a) \cdot (\textit{something})$, and to find what that something is, it becomes more apparent after the next iteration using the third derivative:

$$f'''(a) = \sum_{n=0}^{\infty} c_n n(n-1)(n-2)(a-a)^{n-3} = c_0(0)+c_1(0)+c_2(0)+c_3(3)(2)(1)(1)\ldots$$

$$f'''(a) = c_3(3)(2)(1)(1) \implies c_3 = \frac{f'''(a)}{(3)(2)(1)} = \frac{f'''(a)}{3!}$$

By observing this result, we can deduce that for any $n \geq 0$:

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

This is part of the definition of a **Taylor Series**. More precisely:

## Definitions of the Taylor and Maclaurin Series

If $f$ has infinitely many derivatives at $x = a$ then the **Taylor Series** for the function $f$ at $a$ is:

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

... and the Taylor Series centered at $a = x = 0$ is called the **Maclaurin Series**.

There are many nuances to this definition. Notice that in order to produce a true infinite series definition for $f(x)$, the function must have infinite derivatives, also called being *smooth*. Another important note is that we do not know whether this power series *equals* $f(x)$, or if it does on some interval. What this means is we do not know whether this power function converges to $f(x)$ for all $x$, or at all. It is just a power series that relates to any general function $f(x)$. Exactly how it relates will be discussed later in this section.

Before we move on, an important note is also that if some function has a power series that converges to it on some interval, then that power series *is* the *Taylor Series*.

## Taylor Polynomials

The $k^{th}$ degree **Taylor Polynomial** denoted $T_k(x)$ is the $k^{th}$ partial sum of the **Taylor Series** of $f(x)$, as in:

$$T_k(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

Just as $f(x)$ is a finite (plottable) function, so is $T_k(x)$ for any set value of $k$. Take the following example:

$$f(x) = cos(x)$$

The $0^{th}$ to $8^{th}$ *Taylor Polynomials* for $f(x)$ centered at $x = 0$ are as follows (technically these are called Maclaurin Polynomials because the function is centered at $x = 0$):

$$T_0(x) = T_1(x) = \sum_{n=0}^{0} \frac{f^{(n)}(0)}{n!}x^n = 1$$

$$T_2(x) = T_3(x) = \sum_{n=0}^{2} \frac{f^{(n)}(0)}{n!}x^n = 1 - \frac{x^2}{2!}$$

$$T_4(x) = T_5(x) = \sum_{n=0}^{4} \frac{f^{(n)}(0)}{n!}x^n = 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

$$T_6(x) = T_7(x) = \sum_{n=0}^{6} \frac{f^{(n)}(0)}{n!}x^n = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!}$$

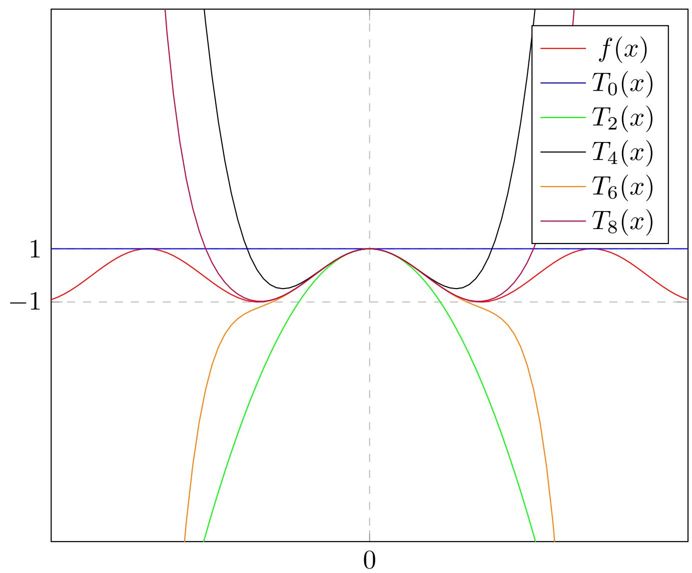
$$T_8(x) = \sum_{n=0}^{8} \frac{f^{(n)}(0)}{n!} x^n = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!}$$

Notice that there are pairs of $T_k(x)$ that are equal because of the fact that:

$$\left( \frac{d}{dx} cos(x) \right) (x = 0) = -sin(0) = 0$$

Now let's observe exactly how these functions act when plotted around our original function $f(x) = cos(x)$. The following plot will contain all those functions.

$cos(x)$ and Maclaurin Polynomials




There's a lot to observe in this plot. The main thing to note is that as the degree of the Maclaurin polynomial increases, it appears to wrap around the function more closely, and on larger and larger intervals. $T_0(x)$ converges to the function $f(x)$ only at $x = 0$, while $T_2(x)$ looks like it converges to $f(x)$ for some larger interval. In other words, the Taylor/Maclaurin polynomial converges to the function on a greater interval $I$ as the polynomial wraps tighter around the original function. This is very important as it describes the connection between the original function and it's Taylor/Maclaurin Series.

How do we quantify how much the Taylor/Maclaurin polynomial wraps around the initial function? How do we know that the $k^{th}$ degree Taylor polynomial is actually wrapping around the function $f(x)$? This brings us to our next topic of this section, **Taylor's Theorem with Remainder**.

## Taylor's Theorem with Remainder

To begin we write $f(x)$ as:

$$f(x) = T_k(x) + R_k(x)$$

Where $R_k(x)$ is some remainder that exists between $T_k(x)$ and the actual value of the function $f(x)$. If $T_k(x)$ truly converges to $f(x)$ over the whole real line then $\lim_{k\to\infty} R_k(x) = 0$. If this is true, then $T_\infty = f(x)$, in other words the infinite power series converges to $f(x)$ for all $x \in \mathbb{R}$. Formally:

The Taylor series for $f(x)$ at $x = a$ converges to $f(x)$ on $I$ if and only if

$$\lim_{n\to\infty} R_n(x) = 0 \text{ for all } x \in I$$

Usually however, we are not working with the entire infinite power series, and want to just work with a Taylor/Maclaurin polynomial approximation of the function $f(x)$. We then need to be able to bound the remainder function $R_n(x)$. Taylor's Theorem with Remainder states that (*Note:* I am switching from using $k$ as the index variable to $n$):

There exists some $c$ such that:

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

This equality is a really important fact. There is some value of $c$ that exists, and makes this true. It is however not easy to find that $c$ so typically we find the maximum value of $f^{(n+1)(x)}$ in the interval we care about and call it $M$, changing the equality to the following inequality:

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

$$\text{where } \left|f^{(n+1)}(x)\right| \leq M \text{ for all } x \in I$$

... and now we have a bound on $R_k(x)$. If you are evaluating for the maximum error on an interval, choose to evaluate the remainder function on the bound which is furthest from the center.

**Steps for Using Taylor's Remainder Theorem to bound $R_n(x)$ in an interval**:

1. Calculate $f^{(n+1)(x)}$ and $f^{(n+2)}(x)$.
2. Use $f^{(n+2)}(x)$ to determine if the maximum of $f^{(n+1)}(x)$ occurs on either end of the interval, or somewhere in between. Call that point $c$.
3. $$\left|R_n(x)\right| \leq \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

4. Choose $x$ to be a point in the interval which is furthest from the center of the power series ($a$), and would thus have the largest error, call that point $d$.

5. Evaluate:
$$\left|R_n(d)\right| \leq \frac{f^{(n+1)}(c)}{(n+1)!}(d-a)^{n+1}$$

## 6.4 Working with Taylor Series

There are many applications of Taylor Series within mathematics, that then stretch into the real world. We will explore those uses in this section.

Say you want to expand the following binomial:

$$(1+x)^2$$

... this wouldn't be hard, but what about:

$$(1+x)^{11}$$

... this would be extremely tedious, and what about:

$$(1+x)^{\frac{56}{13}}$$

... this would not be nice to try and expand.

For this purpose we use the following general formula called the **binomial series**:

$$(1+x)^r = \sum_{n=0}^{\infty} \binom{r}{n} x^n \quad \text{for } |x| < 1$$

... where:

$$\binom{r}{n} = rCn = \frac{r!}{n!(r-n)!}$$

... where $r$ and $n$ are integers. If $r$ is not an integer, then:

$$\binom{r}{n} = \frac{r(r-1)(r-2)\cdots(r-n-1)}{n!}$$

$\binom{r}{n}$ is called the combination operator, the exact use and reason for this particular combination of factorials will be omitted, you will learn it in a discrete math course. You can plug this into your calculator using the $nCr$ button (this text uses $rCn$ for some reason). Another very useful technique when using the binomial series is the substitution:

$$(b+x)^r = (b+a)^r \left(1 + \frac{x-a}{b+a}\right)^r$$

... this re-centres the binomial to $a$. The first part of the product is a constant, and the second is a new binomial centred at $a$.

The following is a **selection of commonly used Maclaurin Series** and their intervals of convergence:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \text{ for } x \in (-1,1)$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \text{ for } x \in \mathbb{R}$$

$$sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \text{ for } x \in \mathbb{R}$$

$$cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \text{ for } x \in \mathbb{R}$$

$$Ln(1+x) = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^n}{n} \text{ for } x \in (-1,1]$$

$$\tan^{-1}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \text{ for } x \in (-1,1]$$

$$(1+x)^r = \sum_{n=0}^{\infty} \binom{r}{n} x^n \text{ for } x \in (-1,1)$$

## Differential Equations and Integrals

Power series can be used to solve differential equations and integrals. To solve a differential equation using power series, just assume that:

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

... and substitute that into the differential equation, then solve for $c_n$ using properties of power series.

A power series can also be used to solve integrals, and they are particularly useful when trying to find a non-elementary antiderivative. The most famous example of this is the following:

$$f(x) = \int e^{-x^2} dx$$

Here, $f(x)$ is a **non-elementary function** meaning it cannot be obtained by combining algebraic functions, trigonometric functions, exponential functions, logarithmic functions, and/or inverse trig functions using addition, subtraction, multiplication, division, and/or composition.

$f(x)$ can only by expressed as a power series. This can be done by integrating the power series representation of $e^{-x^2}$.

## 4.1 Functions of Several Variables

A function of several variables, is one which takes in multiple input values, and has multiple output values. Specifically we will be focusing on functions with multiple input variables and a single output variable.

A function of two variables (meaning a function with two input variables and one output variable) can be written as:

$$z = f(x, y) = f(x_1, x_2)$$

... here, $z$ is the single output variable, and $x$ and $y$ are the two input variables.

An example of this in use for a function of 2 variables could be like the following:

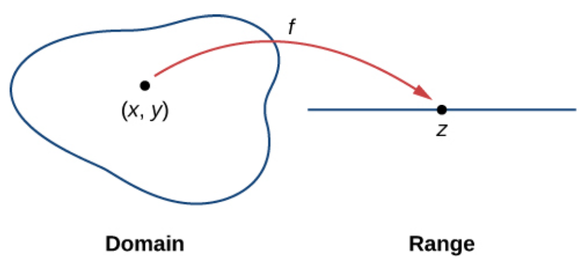
$$f(x, y) = \sqrt{9 - x^2 - y^2}$$

$$f(1, 2) = \sqrt{9 - 1^2 - 2^2} = 2$$

Now let's consider the **Domain** and **Range** of functions of multiple variables. Typically definitions in this section are done using functions of 2 variables, and they are either then defined in general or extend as you'd expect them to.

The **Domain** of a function of 2 variables is all the points $(x, y) \in \mathbb{R}^2$ such that $f(x, y)$ is defined. In general, the domain of a function of $n$ variables is all the points $(x_1, x_2, x_3, \ldots, x_n) \in \mathbb{R}^n$ such that $f(x_1, x_2, x_3, \ldots, x_n)$ is defined.

The **Range** of a function of 2 variables is all values $z \in \mathbb{R}$ such that there exists some $f(x, y) = z$ for some $(x, y) \in \mathbb{R}^2$. In general, the range of a function of $n$ variables is all values $z \in \mathbb{R}$ such that there exists some $f(x_1, x_2, x_3, \ldots, x_n) = z$ for some $(x_1, x_2, x_3, \ldots, x_n) \in \mathbb{R}^n$.




For example, the domain of the function used earlier:

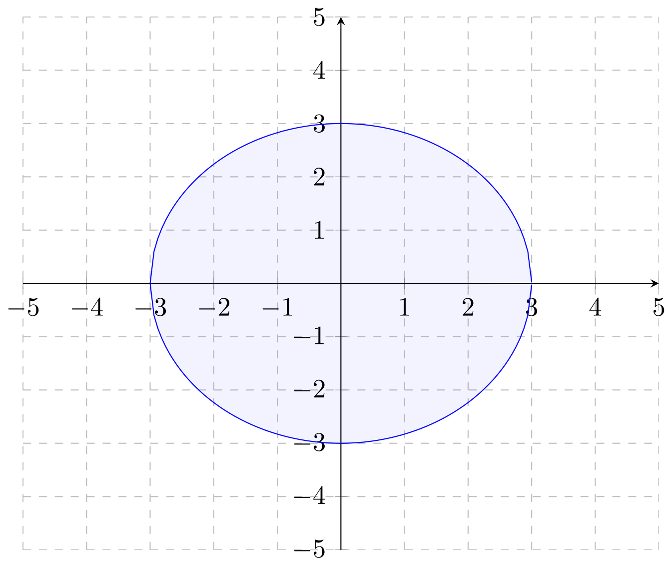
$$f(x, y) = \sqrt{9 - x^2 - y^2}$$

... is all $(x, y) \in \mathbb{R}^2$ such that:

$$x^2 + y^2 \leq 9$$

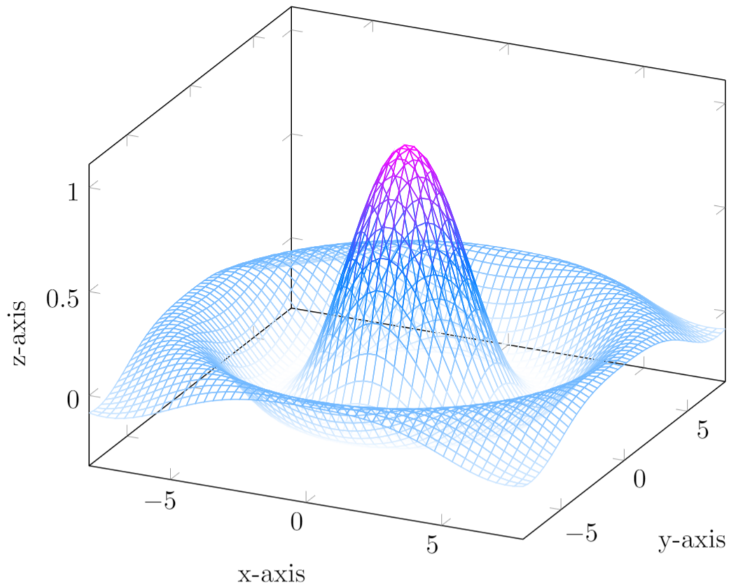
Note that this is not an interval, it is a set of points in $\mathbb{R}^2$ which can be visualized by the following graph:

Domain of $f(x, y) = \sqrt{9 - x^2 - y^2}$

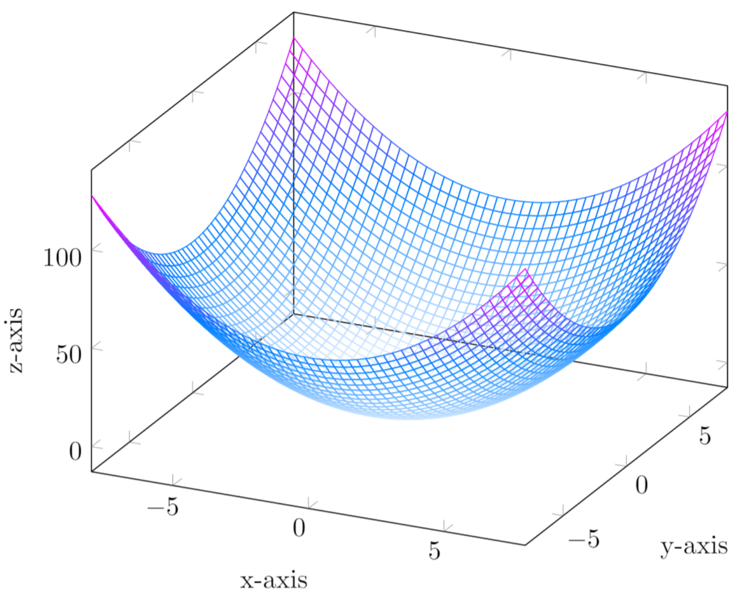


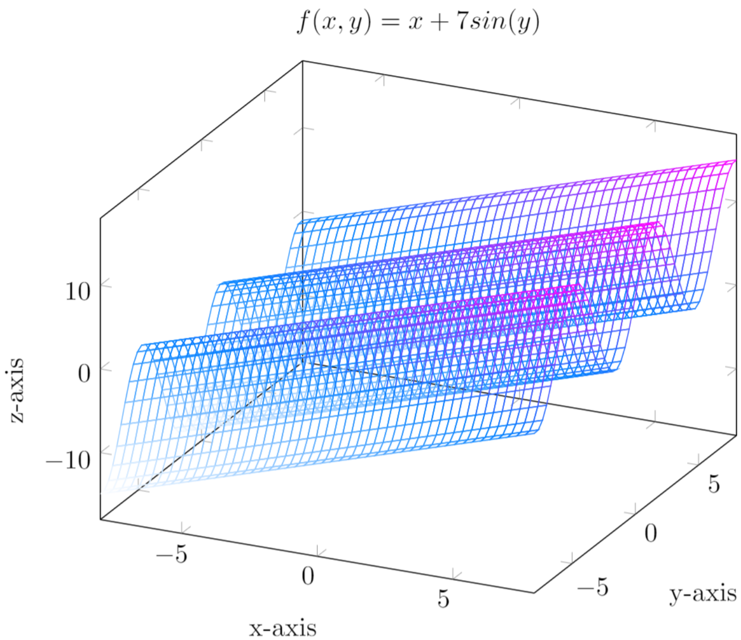

Any point contained within this set is in the domain of $f(x, y)$, including the border. Notice that a function of $n$ variables has a domain defined in $\mathbb{R}^n$. Then how do we graph multivariable functions? For a function of two variables, this is still relatively simple to do, however once you go into higher dimensions, visualization is impossible. Consider a function of 1 variable $y = f(x)$, here we have one input variable and 1 output variable, and so it's graph is a two dimensional plane. Now consider a function of two variables $z = f(x, y)$, here there is two input variables, and one output variable, so it's graph is a three dimensional space. This is typically visualized by considering the plane of $(x, y) \in \mathbb{R}^2$, and then the corresponding $z = f(x, y)$ is the height above the plane. This makes a three dimensional image called a **surface**. The following are a few examples:

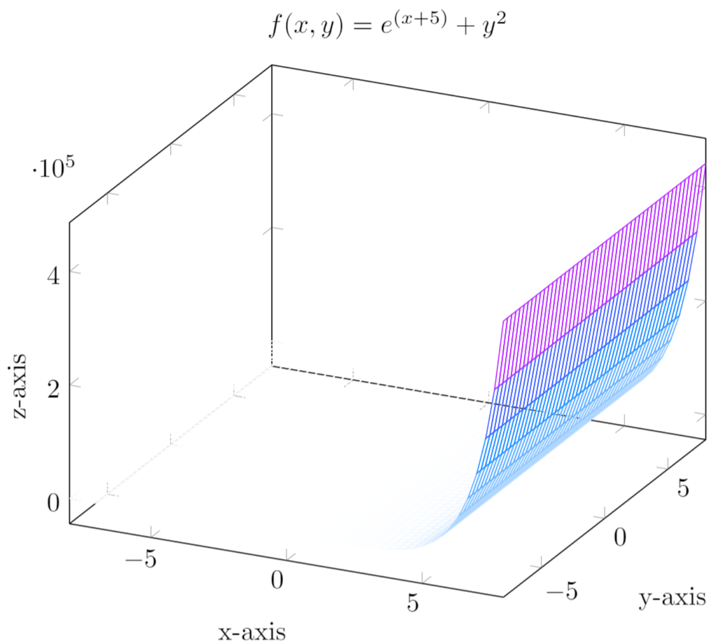
$$f(x,y) = \frac{sin(\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}$$



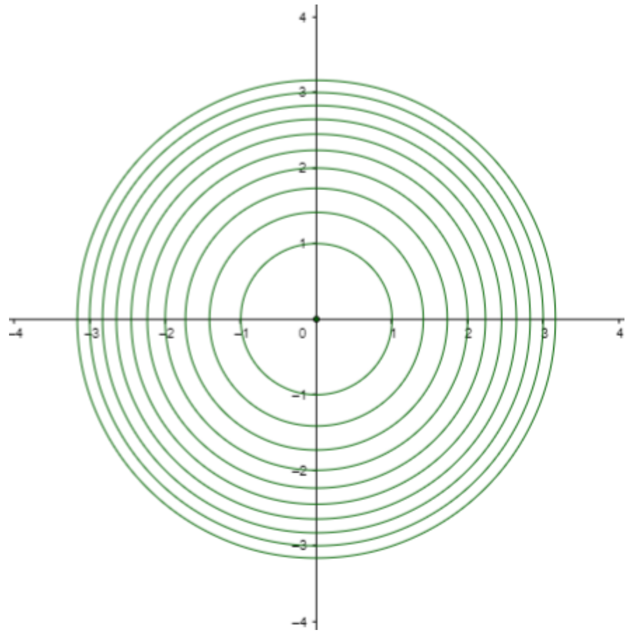

$$f(x,y) = x^2 + y^2$$





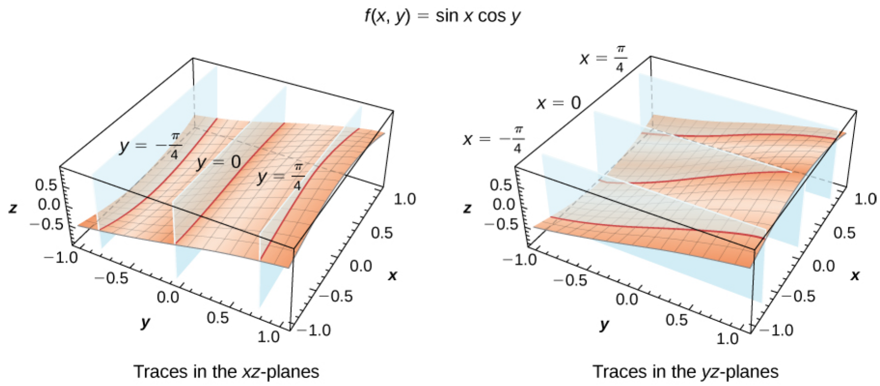

These graphs are almost impossible to draw without the help of a computer system. For this reason, we have another way of visualizing 3 dimensional plots which is called a **contour map**. Take for example the second 3D graph I showed, you can imagine taking horizontal slices of the graph an equal distances apart and then marking how the slice and the graph intersect. Each of these intersections will be some line/squiggle called a level curve in $\mathbb{R}^2$. Take each of those level curves and lay them flat on the $xy$-plane and you have a contour map such as the following (for $f(x, y) = x^2 + y^2$):



You can calculate a single **level curve** by setting $f(x, y) = k$ for some $k \in \mathbb{R}$ and plotting the resulting single variable function on the $xy$-plane. If $f$ is a function of three variables, then the you can take the equivalent of level curves (which are called **level surfaces**) for some fixed value of the output, as in the graph of $f(x, y, z) = k$ for some $k \in \mathbb{R}$.

Now imagine vertically slicing the graph with some plane, the intersection of the 3D plot and the plane will be some line/squiggle in $\mathbb{R}^2$ called the *vertical trace*. The **vertical trace** of a function $f(x, y)$ is either the graph of:

- $f(a, y) = z$ on the $yz$-plane for some $a \in \mathbb{R}$
- $f(x, b) = z$ on the $xz$-plane for some $b \in \mathbb{R}$




# 4.2 Limits and Continuity

The limit was defined in single variable functions using it's *functional definition.* This definition does not suffice for functions of multiple variables. The rigorous *epsilon-delta definition of the limit* is the true definition of the limit of a function of any number of variables including one. This definition will be described in this section, however you will not be asked in this course to apply the $\varepsilon$ - $\delta$ definition directly.

## $\varepsilon$ - $\delta$ Definition of the Limit

Let $f(x, y) = z$ be a function of two variables and one output. Say that:

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

... if for every $\varepsilon > 0$ there is a $\delta > 0$ such that:

$$\left|f(x, y) - L\right| < \varepsilon \text{ whenever } \sqrt{(x-a)^2 + (y-b)^2} < \delta$$

This definition is essentially saying that if $f(x, y)$ is within $\varepsilon$ distance from $L$, then there is a corresponding input value within a radius of $\delta$ from $(a, b)$ in the input plane that gets us within that length $\varepsilon$ from $L$, regardless of how small $\varepsilon$ is. If there is not restriction on how small $\varepsilon$ can be then the limit is truly equal to $L$. There is much nuance to this definition of the limit, however that will be omitted from this text.

## Limit Laws

Say that... $f(x, y)$ and $g(x, y)$ are two variables functions, $n \in \mathbb{Z}$ and $c \in \mathbb{R}$. Then:

1. $$\lim_{(x,y)\to(a,b)} c = c$$

2.

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

3.

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

4.

$$\lim_{(x,y)\to(a,b)} (f(x,y) \pm g(x,y)) = \lim_{(x,y)\to(a,b)} f(x,y) \pm \lim_{(x,y)\to(a,b)} g(x,y)$$

5.

$$\lim_{(x,y)\to(a,b)} cf(x,y) = c \lim_{(x,y)\to(a,b)} f(x,y)$$

6.

$$\lim_{(x,y)\to(a,b)} (f(x,y) \cdot g(x,y)) = \lim_{(x,y)\to(a,b)} f(x,y) \cdot \lim_{(x,y)\to(a,b)} g(x,y)$$

7.

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)}{g(x,y)} = \frac{\lim_{(x,y)\to(a,b)} f(x,y)}{\lim_{(x,y)\to(a,b)} g(x,y)}$$

8.

$$\lim_{(x,y)\to(a,b)} (f(x,y))^n = \left( \lim_{(x,y)\to(a,b)} f(x,y) \right)^n$$

... of course given that all exists, no denominators are 0, and all results are real.

There is a way to show a limit of a multivariable function **does not exist** which is unique to this type of limit, and that is as follows: Say $f(x,y) \to L$ as $(x,y) \to (a,b)$ along some single variable curve $y = g(x)$ where $\lim_{x\to a} g(x) = b$ if $\lim_{x\to a} f(x, g(x)) = L$. What this means is that if we were to approach our limit $(a,b)$ along some path in the input plane (similar to left and right sided limits) we should get the same limit result regardless of the path we choose. If we can find two path that result in different results for the limit, then we know that the limit does not exist.

Another useful technique to evaluate these limits is to convert a function to polar form, which is essentially just a change of variable in the form:

$$x = r\cos(\theta)$$

$$y = r\sin(\theta)$$

... which is useful if you see an expression containing $x^2+y^2$. There is essentially no other way to prove a limit is true except in certain situations where there is some algebraic trick, or using the $\varepsilon - \delta$ definition.

## Set Definitions and Theorems

Let $S$ be a set in $\mathbb{R}^2$. Say that $(x, y)$ is in the **interior** of $S$ if both:

- $$(x, y) \in S$$
- For some $\delta > 0$, every point in $\mathbb{R}^2$ with distance from $(x, y) < \delta$ is also in $S$.

What the second part means is that these is some radius away from the point that is non-zero, inside of which all the points are also part of $S$.

Say that $(x, y)$ is in the **boundary** of $S$ if both:

- $$(x, y) \in S$$
- No matter how small $\delta > 0$ is, there is some $(x_1, y_1) \in S$ and some $(x_2, y_2) \notin S$, where $(x_1, y_1) \neq (x, y) \neq (x_2, y_2)$, and the distance to $(x_1, y_1)$ and $(x_2, y_2) < \delta$.

The following is a list of shorter definitions:

- Say $S$ is **open** if it does not contain any of it's boundary points.
- Say that $S$ is **closed** if it contains all of it's boundary points.
- Say that an open set $S$ is **connected** if it is possible to draw a path in $S$ from one point in $S$ to every other point in $S$.
- Say that $S$ is a region if $S$ is:
  - Open
  - Connected
  - Non-empty

Let $(a, b)$ be a boundary point of the domain of $f(x, y)$. Say that:

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

... if for every $\varepsilon > 0$ there is a $\delta > 0$ such that:

$$\left|f(x, y) - L\right| < \varepsilon \text{ whenever } \sqrt{(x-a)^2 + (y-b)^2} < \delta$$

... and $(x, y)$ is in the domain of $f$.

This last sentence is the new addition, this means that a function can still have a limit even if the input is on the boundary because we just discount all the points which are not in the domain while calculating the limit.

Say that $f(x, y)$ is continuous at $(a, b)$ if:

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

If $f(x, y)$ and $g(x, y)$ are continuous at $(a, b)$ then:

- $f(x, y) \cdot g(x, y)$ is continuous at $(a, b)$
- $f(x, y) \pm g(x, y)$ is continuous at $(a, b)$
- $\frac{f(x,y)}{g(x,y)}$ is continuous at $(a, b)$, if $g(x, y) \neq 0$
- $h(f(x, y), g(x, y))$ is continuous at $(a, b)$ if $h(x, y)$ is continuous at $(f(a, b), g(a, b))$

*Note:* everything in this section applies as you'd expect for functions of more than two variables.

## 4.3 Partial Derivatives

A derivative is the measure of the sensitivity of the function to it's input variable(s). In single variable functions, we had a single input and thus had a single derivative. In a multivariable function of $n$ variables, there will be $n$ **partial derivatives** that each relate to the sensitivity of the functions output to one of each of the functions inputs. In each of these partial derivatives, the input variables not being differentiated is assumed to be constant. The following are it's formal definitions:

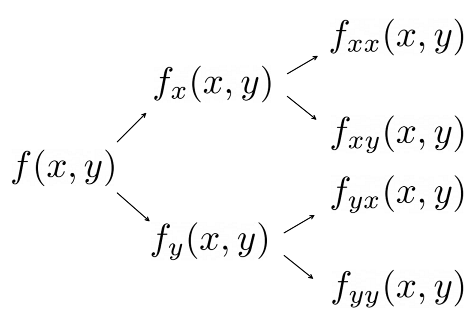
Say a function $f(x, y)$ is a function of two variables, then:

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

The main trick to taking partial derivatives is that you can treat it like a regular derivative, but consider the non-differentiating input variable to be a constant, and then differentiate normally with respect to the chosen input variable.

### Higher Order Partial Derivatives

Each partial derivative of a multivariable function is usually a new multivariable function. This means that each partial derivative of a function of $n$ variables has its own set of $n$ partial derivatives (given that the partial derivative doesn't lose any input variables through differentiation). In general, a function of $n$ variables has $n^k$, $k^{th}$ order partial derivatives.



The notation that is almost always used for higher order partial derivatives is:

$$f_{\text{some combination of input variables}}(x, y, z, ...)$$

**Clairaut's Theorem** states that if $f(x, y)$ is defined on some open disk $D$ in the input space, containing a point $(a, b)$ where $f_{xy}(x, y)$ and $f_{yx}(x, y)$ are continuous on the disk $D$, then:

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

Essentially, as long as you're in the interior of the domain, and using elementary functions then $f_{xy}(x, y)$ and $f_{yx}(x, y)$ will be the same.

A **partial differential equation** (PDE) is a differential equation that contains independent variables, and potentially partial derivatives of dependant variables. In this course the most you will be asked to do relating to PDEs is to verify a solution by subbing in a function and showing $LHS = RHS$.

## 4.5 The Chain Rule

Let the following be true:

$$f(x, y) = z, x = g(t), y = h(t)$$

... then what is $\frac{dz}{dt}$? This is effectively a single variable function ($t$ is the input). You can start by taking the partial derivatives of $z$ which would look like:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

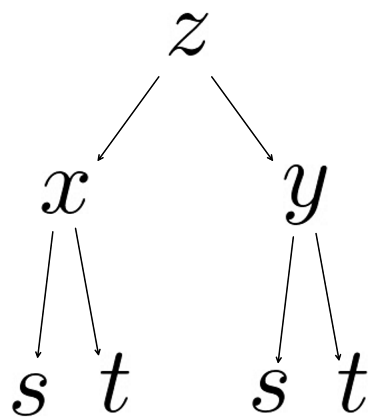
... we have to multiply by the standard derivative of $x$ with respect to $t$ because of the standard chain rule, but the new addition is that we take the sum of all the partial derivatives (in this case two). The functions that define $x$ and $y$ can also be multivariable functions, to deal with this we use the following definition. In full generality the **Chain Rule for multivariable functions** is:

Let $z = f(x_1, x_2, \ldots, x_m)$ where $x_i = x_i(t_1, t_2, \ldots, t_n)$ for each $i \leq m$, then:

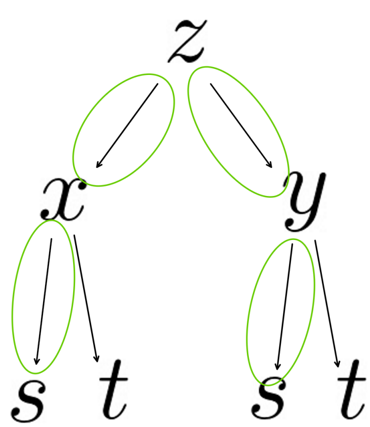
$$\frac{\partial z}{\partial t_j} = \sum_{i=1}^{m} \frac{\partial z}{\partial x_i} \cdot \frac{\partial x_i}{\partial t_j} \text{ for each } j \leq n$$

The best way to understand this is the use a dependence tree which looks like the following:

Let $z = f(x, y), x = g(s, t), y = h(s, t)$, then:




Then to take the derivative of $z$ with respect to either $s$ or $t$, take partial derivatives down the chain as highlighted in the following diagram:




... take the product of the partial derivatives of a chain, and then sum the chains together, as in:

$$\frac{dz}{ds} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s}$$

### Implicit Differentiation

Using multivariable functions, we can have a new perspective on implicit differentiation. Consider a general multivariable function in the form:

$$f(x, y) = z$$

If we choose $z = 0$, then we have a single variable function (just $x$ and $y$). We can use this to find $\frac{dy}{dx}$ by:

$$\frac{\partial}{\partial x} f(x, y) = 0$$

$$f_x(x,y)\frac{\partial x}{\partial x} + f_y(x,y)\frac{\partial y}{\partial x} = 0$$

... we can rearrange for $\frac{dy}{dx}$:

$$\frac{\partial y}{\partial x} = \boxed{\frac{dy}{dx} = -\frac{f_x(x,y)}{f_y(x,y)}}$$

## 4.7 Min/Max Problems in Multiple Variables

This section is concerned with optimization in multivariables, defining concepts like critical points and maxima/minima in multivariable inputs.

Say that $f(x,y)$ is a function and $(a,b)$ is in it's domain, then $(a,b)$ is a **critical point** of $f(x,y)$ if any of the following is true:

1. $$f_x(a,b) = 0 = f_y(a,b)$$

2. $$f_x(a,b) \;\; \text{DNE}$$

3. $$f_y(a,b) \;\; \text{DNE}$$

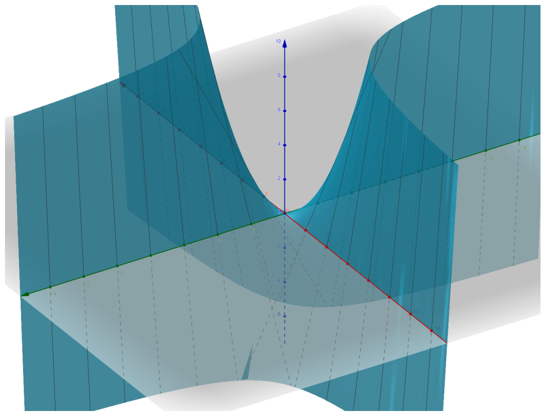
Say that $f(x,y)$ is a function and $(a,b)$ is in it's domain $D$, then:

- Say that $(a,b)$ is a local minimum if $f(x,y) \geq f(a,b)$ for all $(x,y)$ close enough to $(a,b)$.
- Say that $(a,b)$ is a local maximum if $f(x,y) \leq f(a,b)$ for all $(x,y)$ close enough to $(a,b)$.
- Say that $(a,b)$ is an absolute minimum if $f(x,y) \geq f(a,b)$ for all $(x,y) \in D$.
- Say that $(a,b)$ is an absolute maximum if $f(x,y) \leq f(a,b)$ for all $(x,y) \in D$.

If the open set $D$ contains a local extremum of $f(x,y)$ then the local extremum is a critical point of $f(x,y)$. A **Saddle Point** is a point $(x_0, y_0)$ where

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0$$

and $(x_0, y_0)$ is not a local extremum of $f(x,y)$. An example of this is the point at the origin of the following graph ($f(x,y) = xy$):



## The Second Derivative Test

This test is used to determine the nature of critical points, whether they are local extrema, the type of extrema, or if they are a saddle point.

Assume $f_x(a,b) = f_y(a,b) = 0$. Define:

$$D = f_{xx}(a,b) \cdot f_{yy}(a,b) - (f_{xy}(a,b))^2$$

... and evaluate the following:

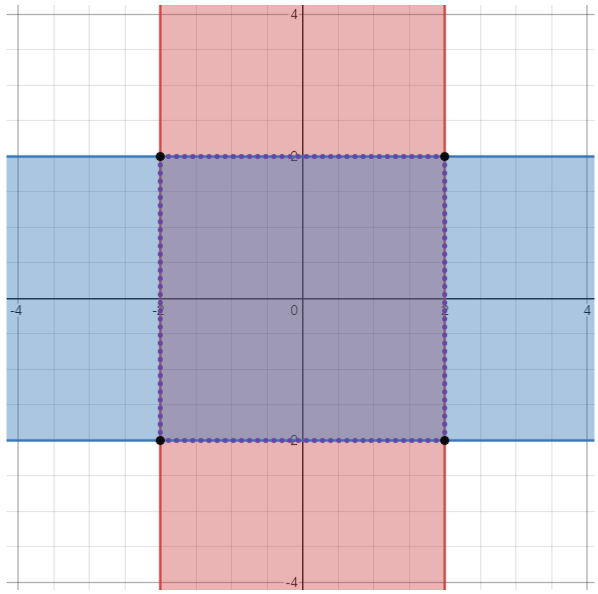
1. if $D > 0$ and $f_{xx}(a,b) > 0$ then $(a,b)$ is a local minimum.
2. if $D > 0$ and $f_{xx}(a,b) < 0$ then $(a,b)$ is a local maximum.
3. if $D < 0$ then $(a,b)$ is a saddle point.
4. if $D = 0$ then the test is inconclusive.

**Extreme Value Theorem** states that if $f(x,y)$ is continuous on a closed and bounded set $S$ then $f(x,y)$ will have an absolute minimum and an absolute maximum on $S$. Furthermore, that absolute min/max will occur either at a local extremum of $f(x,y)$ in $S$ or on the boundary of $S$.

The tricky part of this is that the boundary of $S$ could be (and usually is in multivariable functions) not just a point on a line. For example consider a function who's domain is:

$$S = \{(x,y) \in \mathbb{R}^2 \mid -2 \leq x \leq 2, -2 \leq y \leq 2\}$$

... which can be visualised with the following diagram:




Considering this domain, the max/min of the $f(x, y)$ can occur at either a critical point inside the domain, or on the boundary. To evaluate this you would need to find the maximum on each of the boundary lines (purple dotted lines in the diagram) which requires single variable optimization after defining an equation to describe the side (in this example the top boundary would be $y = 2$). Then *for each of those boundaries* evaluate their boundary points (this is required as part of single variable optimization). For this example after evaluating you should have a collection of points which looks like:

1. Critical points inside the domain.
2. 4 critical points on each of the boundary lines.
3. 4 corner points.

Disregard all points which are not in $S$ and consider the remaining points when determining min/max of the function $f(x, y)$ in the domain $S$.

## Conclusion

This concludes the course material for MTH 240. I hope that you found these notes helpful. Best of luck on the final exam!

- Adam Szava