# Calculus

Kevin

April 2021

# Contents

# 1   Two Standard Limits

The following two standard limits are important for differentiating trigonometric functions from the first principle, i.e., by definition.

$$\lim_{\theta \to 0} \frac{\sin(\theta)}{\theta} = 1 \tag{1}$$

$$\lim_{\theta \to 0} \frac{\cos(\theta) - 1}{\theta} = 0 \tag{2}$$

The proof is essentially about establishing an upper and lower bound for these two functions.

*Proof.* We prove the first one by squeeze lemma. We know that the following inequality is true on a sufficient small neighborhood $(-\epsilon, \epsilon)$ around the origin.

$$|\sin(\theta)| \leq |\theta| \leq |\tan(\theta)|$$

This is equivalent to

$$\left| \frac{\sin(\theta)}{\theta} \right| \leq 1 \leq \left| \frac{\sin(\theta)}{\theta \cos(\theta)} \right|$$

Rearranging the second inequality above gives

$$|\cos(\theta)| \leq \left| \frac{\sin(\theta)}{\theta} \right|$$

Hence,

$$|\cos(\theta)| \leq \left| \frac{\sin(\theta)}{\theta} \right| \leq 1 \tag{3}$$

By squeeze lemma,

$$\lim_{\theta \to 0} |\cos \theta| \leq \lim_{\theta \to 0} \left| \frac{\sin(\theta)}{\theta} \right| \leq \lim_{\theta \to 0} 1 \tag{4}$$

which implies that $\lim_{\theta \to 0} \frac{\sin(\theta)}{\theta} = 1$. However, to assert that the limit of this function exist, one needs to appeal to some $\epsilon$-$\delta$ argument to prove that the limits from both sides exist and agree.

There is a proof of the second limit in Haese (textbook). I'll present a different proof that still uses squeeze lemma. Note that

$$\left| \frac{1 - \cos(\theta)}{\theta} \right| \geq 0$$

holds trivially. This will be our lower bound. Now, consider the unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ and the following diagram (using Geogebra).
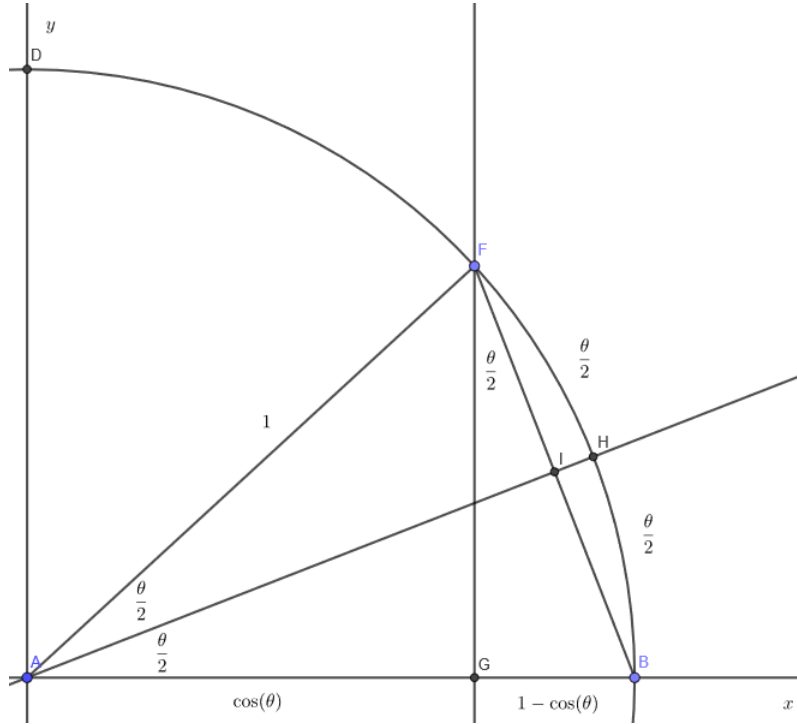


Figure 1: Geometric proof

It is easy to prove that $\triangle AIB \sim \triangle FGB$. $\because AH$ divides the angle $\theta$ into two equivalent parts, $\therefore$ $\angle GFB = \frac{\theta}{2}$. By the definition of $\sin(\theta)$, we have

$$\sin\left(\frac{\theta}{2}\right) = \frac{BG}{FB} = \frac{1 - \cos(\theta)}{FB} \geq \frac{1 - \cos(\theta)}{\theta}$$

Hence,

$$0 \leq \left| \frac{1 - \cos(\theta)}{\theta} \right| \leq \sin\left(\frac{\theta}{2}\right) \tag{5}$$

By squeeze lemma,

$$\lim_{\theta \to 0} 0 \leq \lim_{\theta \to 0} \left| \frac{1 - \cos(\theta)}{\theta} \right| \leq \lim_{\theta \to 0} \sin\left(\frac{\theta}{2}\right) \tag{6}$$

which implies that $\lim_{\theta \to 0} \frac{\cos(\theta) - 1}{\theta} = -\lim_{\theta \to 0} \frac{1 - \cos(\theta)}{\theta} = 0$. $\qquad \square$

# 2 The Derivatives of Trigonometric Functions (First Principle)

## 2.1 Sine and Cosecant

Let's start......

$$\frac{d}{dx}\sin(x) = \lim_{h \to 0} \frac{\sin(x+h) - \sin(x)}{h}$$
$$= \lim_{h \to 0} \frac{\sin(x)\cos(h) + \sin(h)\cos(x) - \sin(x)}{h}$$
$$= \lim_{h \to 0} \frac{\sin(x)(\cos(h) - 1) + \cos(x)\sin(h)}{h}$$
$$= \sin(x)\lim_{h \to 0} \frac{\cos(h) - 1}{h} + \cos(x)\lim_{h \to 0} \frac{\sin(h)}{h}$$
$$= \cos(x)$$

Two methods for driving the derivative of Cosecant.
**Method 1:**
csc is the reciprocal of sin, so let $u = \sin(x)$, then

$$\frac{d}{dx}\csc(x) = \frac{d}{du}\frac{1}{u}\frac{d}{dx}\sin(x) = -\frac{1}{u^2}\cos(x) = -\frac{\cos(x)}{\sin^2(x)}$$
$$= -\csc(x)\cot(x)$$

**Method 2 (First Principle):**
Start:

$$\frac{d}{dx}\csc(x) = \lim_{h \to 0}\left(\frac{1}{h}\left(\frac{1}{\sin(x+h)} - \frac{1}{\sin(x)}\right)\right)$$
$$= \lim_{h \to 0}\left(\frac{1}{h}\frac{\sin(x) - \sin(x+h)}{\sin(x+h)\sin(x)}\right)$$
$$= \lim_{h \to 0} \frac{2\sin((x - x - h)/2)\cos((2x+h)/2)}{h\sin(x+h)\sin(x)}$$
$$= \lim_{h \to 0} \frac{\sin((x - x - h)/2)\cos((2x+h)/2)}{\frac{h}{2}\sin(x+h)\sin(x)}$$
$$= \lim_{h \to 0} \frac{\sin(-h/2)}{h/2}\lim_{h \to 0}\frac{\cos(x+h/2)}{\sin(x+h)\sin(x)}$$
$$= -\frac{\cos(x)}{\sin^2(x)}$$
$$= -\csc(x)\cot(x)$$

## 2.2 Cosine and Secant

Can you complete these two similarly? (DO DO IT)

The details will not be as clear as the preceding subsection because I'm lazy, and I don't want to repeat similar argument.

$$\frac{d}{dx}\cos(x) = \lim_{h \to 0}\frac{\cos(x+h) - \cos(x)}{h}$$
$$= \lim_{h \to 0}\frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h}$$
$$= \cos(x)\lim_{h \to 0}\frac{\cos(h) - 1}{h} - \sin(x)\lim_{h \to 0}\frac{\sin(h)}{h}$$
$$= -\sin(x)$$

Now we turn to $\sec(x)$. I'll skip the first method because it's similar to Method 1 of differentiating $\csc(x)$.

$$\frac{d}{dx}\sec(x) = \lim_{h\to 0}\frac{\cos(x) - \cos(x+h)}{h\cos(x+h)\cos(x)}$$

$$= \lim_{h\to 0}\frac{2\sin((2x+h)/2)\sin((x+h-x)/2)}{h\cos(x+h)\cos(x)}$$

$$= \lim_{h\to 0}\frac{\sin(x+h/2)}{\cos(x+h)\cos(x)}\lim_{h\to 0}\frac{\sin(h/2)}{h/2}$$

$$= \sec(x)\tan(x)$$

## 2.3 Tangent and Cotangent

Use quotient rule to derive the derivative for these two if you are allowed to use it. Otherwise, use the following derivation.

$$\frac{d}{dx}\tan(x) = \lim_{h\to 0}\left(\frac{1}{h}\left(\frac{\sin(x+h)}{\cos(x+h)} - \frac{\sin(x)}{\cos(x)}\right)\right)$$

$$= \lim_{h\to 0}\frac{\sin(x+h)\cos(x) - \sin(x)\cos(x+h)}{h\cos(x+h)\cos(x)}$$

$$= \lim_{h\to 0}\frac{\sin(x+h-x)}{h\cos(x+h)\cos(x)}$$

$$= \lim_{h\to 0}\frac{\sin(h)}{h}\lim_{h\to 0}\frac{1}{\cos(x+h)\cos(x)} = \frac{1}{\cos^2(x)}$$

$$= \sec^2(x)$$

For Cotangent,

$$\frac{d}{dx}\cot(x) = \lim_{h\to 0}\left(\frac{1}{h}\left(\frac{\cos(x+h)}{\sin(x+h)} - \frac{\cos(x)}{\sin(x)}\right)\right)$$

$$= \lim_{h\to 0}\frac{\cos(x+h)\sin(x) - \cos(x)\sin(x+h)}{h\sin(x+h)\sin(x)}$$

$$= \lim_{h\to 0}\frac{\sin(x-x-h)}{h\sin(x+h)\sin(x)}$$

$$= \lim_{h\to 0}\frac{-\sin(h)}{h}\lim_{h\to 0}\frac{1}{\sin(x+h)\sin(x)}$$

$$= -\csc^2(x)$$

# 3 Proof of Rules of Differentiation (Incomplete)

The proof for rules of scalar multiplication and addition are trivial, so I'll not waste my time on those things.

**Theorem 3.1.** Suppose $f(x) = x^n$ $n \in \mathbb{Z}^+$, then $f'(x) = nx^{n-1}$.

*Proof.* Simple!

$$\frac{d}{dx}x^n = \lim_{h\to 0}\frac{(x+h)^n - x^n}{h}$$

$$= \lim_{h\to 0}\frac{x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \ldots + h^n - x^n}{h}$$

$$= \lim_{h\to 0}(\binom{n}{1}x^{n-1} + \binom{n}{2}x^{n-2}h + \ldots + h^{n-1}) = \binom{n}{1}x^{n-1} = nx^{n-1}$$

$\square$

**Theorem 3.2.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are two differentiable functions, then $(fg)' = f'g + fg'$.

*Proof.* We know the identity

$$\Delta(fg) = \Delta f \cdot g(x) + f(x + \Delta x) \cdot \Delta g$$

To visualize why that identity holds, consider the following diagram.
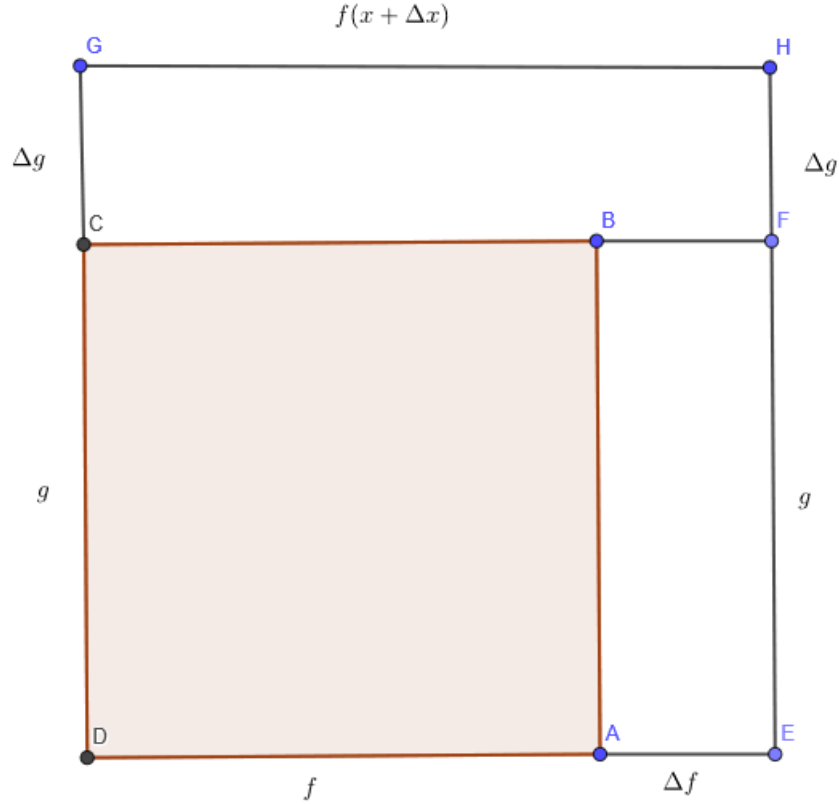


Figure 2: Illustration of $\Delta(fg)$

The shaded area is the value of the product $fg$. By increasing the variable by a small $\Delta x$, the change of the value is represented by the un-shaded (white) region.

Hence,

$$(fg)'(x) = \lim_{\Delta x \to 0} \left( \frac{\Delta f}{\Delta x} g(x) + f(x + \Delta x) \frac{\Delta g}{\Delta x} \right)$$
$$= f'g + fg'$$

$\square$

**Theorem 3.3.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are two differentiable functions such that $g \neq 0$. Then,

$$\left( \frac{f}{g} \right)' = \frac{f'g - fg'}{g^2}$$

*Proof.* The first proof that I'm going to present here uses the product rule that we just proved, but it requires some basic knowledge about implicit differentiation. Suppose $y(x) = \frac{f(x)}{g(x)}$, then $f(x) = y(x)g(x)$,

which means

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

$$y'(x) = \frac{f'(x) - y(x)g'(x)}{g(x)}$$

$$y'(x) = \frac{f'(x) - \frac{f(x)}{g(x)}g'(x)}{g(x)}$$

$$y'(x) = \frac{\frac{f(x)}{g(x)}g(x) - \frac{f(x)}{g(x)}g'(x)}{g(x)}$$

$$y'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$$

□

*Proof.* The quotient rule can also be proven by first principle. The proof is left to the readers as an exercise. HaHaHa! □

**Theorem 3.4.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are two differentiable functions such that range$(f) \subseteq$ dom$(g)$, then

$$(g \circ f)' = \frac{dg}{df}\frac{df}{dx}$$

*Proof.* The damn argument presented in most high school materials is NOT correct, those people know nothing about the failure of their fake proof... I'm gonna tell you why...

Consider the rate of change,

$$\frac{\Delta g}{\Delta x} = \frac{\Delta g}{\Delta f}\frac{\Delta f}{\Delta x}$$

so that $\Delta f$ cancels out, and taking the limit gives us the desired result. NO!!! Well, it's partially correct but incomplete, damn... because of the following situation might occur: **As $\Delta x$ goes to zero, $\Delta f$ might be equal to zero before $\Delta x$ really gets to zero**. It sounds a bit counterintuitive, but it's possible. So the first term needs to be adjusted. Here's what we should do:

We know that $g$ is differentiable at every $f(x)$ by assumption, meaning that $\frac{\Delta g}{\Delta x} = g'(f(x)) + E$, where $E$ is an error term that approaches to 0 as $\Delta x \to 0$. Thus, we should prove it by

$$\frac{\Delta g}{\Delta x} = (g'(f(x)) + E)\frac{\Delta f}{\Delta x}$$

Hence,

$$\lim_{\Delta x \to 0} \frac{\Delta g}{\Delta x} = g'f'$$

□

This adjustment is essential even if it seems to be trivial to our readers because simply writing a quotient there makes no sense.

# 4   The Derivative of Exponential and Logarithmic Function

Suppose we get $e^x$, then

$$\frac{d}{dx}e^x = \lim_{h \to 0} \frac{e^{x+h} - e^x}{h}$$

$$= \lim_{h \to 0} \frac{e^x(e^h - 1)}{h}$$

$$= e^x \lim_{h \to 0} \frac{e^h - 1}{h}$$

Set $y = e^h - 1$, then

$$\frac{d}{dx}e^x = e^x \lim_{y \to 0} \frac{y}{\ln(y+1)}$$

Now, substituting $k = \frac{1}{y}$ will give us

$$\frac{d}{dx}e^x = e^x \lim_{k \to \infty} \frac{1}{k \ln\left(\frac{1}{k} + 1\right)}$$
$$= e^x \lim_{k \to \infty} \frac{1}{\ln\left(1 + \frac{1}{k}\right)^k}$$
$$= e^x \frac{1}{\ln e}$$
$$= e^x$$

Therefore, if $f(x) = a^x$ where $a \in \mathbb{R}$, then

$$\frac{d}{dx}a^x = \frac{d}{dx}e^{x \ln a} = \ln(a)e^{x \ln(a)} = \ln(a)a^x$$

DONE!

Now, consider $\ln(x)$, then

$$\frac{d}{dx}\ln(x) = \lim_{h \to 0} \frac{\ln(x+h) - \ln(x)}{h}$$
$$= \lim_{h \to 0} \frac{\ln\left(\frac{x+h}{x}\right)}{h}$$
$$= \lim_{h \to 0} \frac{\ln\left(1 + \frac{h}{x}\right)}{h}$$
$$= \lim_{h \to 0} \frac{\frac{h}{x} \ln\left(1 + \frac{h}{x}\right)^{\frac{x}{h}}}{h}$$
$$= \lim_{h \to 0} \left(\frac{1}{x} \ln\left(1 + \frac{h}{x}\right)^{\frac{x}{h}}\right)$$
$$= \frac{1}{x} \ln(e)$$
$$= \frac{1}{x}$$

# 5 Differentiate Implicit Functions

Before discussing any differentiation, we need to understand what is an implicit function. For this section, we only care about those functions whose graphs are visible on $\mathbb{R}^2$. It's not going to cause any confusion unless we pass to higher dimensional cases, where the graphs become some sort of manifolds...

**Definition 1.** An implicit function is defined using a relation $f(x, y) = 0$

**Example 1.** The unit circle can be defined by an implicit function

$$x^2 + y^2 - 1 = 0$$

**Example 2.** Although the parabola $y = x^2$ is a function, we can also express that in implicit forms, i.e., $y - x^2 = 0$.

(Seems like an algebraic variety...) When you want to differentiate an implicit function (if it's differentiable) then there are mainly two ways for you to do:

- Convert that into different branches of explicitly defined functions. For instance we can write down the definition of a circle explicitly by $y = \pm\sqrt{1 - x^2}$.

- Simply assume that $y$ is a function of $x$ and differentiate both side of the equation with respect to $x$.

The second method is usually the most convenient but also the most confusing. Suppose you have an implicitly defined function $f(x, y) = 0$. Now, if you slightly change the value of $x$ by $h$, then to make sure that the pair $(x + h, y')$ still satisfies the equation, we need $y$ to change by a small step $h'$ also. So that $f(x + h, y + h') = 0$. If not, then $f(x + h, y) \neq 0$, which will not appear on the graph. Thus, what it really means to differentiate both side is to make sure that the net change of $f(x, y)$ is always 0. The meaning for this differentiation is to see how much $y$-coordinate needs to change when $x$-coordinate changed a really small step.

**Example 3.** We have $x^2 + y^2 = 1$, then differentiate both side with respect to $x$ gives $2x + 2y\frac{dy}{dx} = 0$, so that $\frac{dy}{dx} = -\frac{x}{y} = \frac{-x}{\pm\sqrt{1-x^2}}$. To see why this gives us the slope of the tangent line for the circle at each given point $(x, y)$, consider the radius joining $(0, 0)$ and $(x, y)$, its slope must be $y/x$. A simple geometry tells us that for another line $l$ to be tangent at $(x, y)$, then $l$ is perpendicular to the radius joining that point, and for this to be true, we require $m_l \frac{y}{x} = -1$, which agrees to the result of differentiation.

Practice more and read the following section about differentiating inverse trig functions.

# 6 The Derivatives of Inverse Trigonometric Functions

This chapter assumes that you know how to apply chain rule and the meaning of differentiating implicit functions.

The derivative of these inverse trig functions are elegant, as you'll see.

**Important note:** For the first section, you may read carefully to study the idea behind the derivation, but try to do the last two sections as exercises.

## 6.1 Arcsin

Suppose $y = \arcsin(x)$, then $x = \sin(y)$, meaning that

$$\frac{d}{dx}x = \frac{d}{dy}\sin(y)\frac{dy}{dx}$$
$$1 = \cos(y)\frac{dy}{dx}$$
$$\frac{dy}{dx} = \frac{1}{\cos(y)}$$

Now, substitute $\sqrt{1 - x^2} = \sqrt{1 - \sin^2(y)} = \cos(y)$, which yields

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - x^2}}$$

You can check that the domain of this function is $[-1, 1] \subset \mathbb{R}$.

## 6.2 Arccos

Suppose $y = \arccos(x)$, then $x = \cos(y)$.

$$1 = -\sin(y)\frac{dy}{dx}$$
$$\frac{dy}{dx} = \frac{-1}{\sin(y)}$$
$$\frac{dy}{dx} = -\frac{1}{\sqrt{1 - x^2}}$$

Do you see the relationship between the derivative of arcsin and arccos?

## 6.3  Arctan

Suppose $y = \arctan(x)$, then $x = \tan(y)$.

$$1 = \sec^2 \frac{dy}{dx}$$

$$\frac{dy}{dx} = \frac{1}{\sec^2(x)}$$

Use $1 + \tan^2(y) = \sec^2(y)$ to rewrite the equation above as

$$\frac{dy}{dx} = \frac{1}{1 + x^2}$$

**Question:** Can you guess answer to $\frac{d}{dx} \cot^{-1}(x)$? Justify your answer with rigorous argument.

## 6.4  Bonus Section!

We're going to differentiate $\sec^{-1}(x)$ and $\csc^{-1}(x)$!
   Suppose $y = \sec^{-1}(x)$, then $x = \sec(y)$.

$$1 = \sec(y) \tan(y) \frac{dy}{dx}$$

$$\frac{dy}{dx} = \frac{1}{\sec(y) \tan(y)}$$

Here it becomes a bit tricky. The domain of $\sec^{-1}(x)$ is $(-\infty, -1] \cup [1, \infty)$, so the range should be $[0, \pi/2) \cup (\pi/2, \pi]$. Hence,

$$\tan(y) = \begin{cases} \sqrt{x^2 - 1} & x \geq 1 \\ -\sqrt{x^2 - 1} & x \leq -1 \end{cases}$$

Therefore,

$$\frac{d}{dx} \sec^{-1}(x) = \frac{1}{|x|\sqrt{x^2 - 1}}$$

   We turn to $\csc^{-1}(x)$ now. Suppose $y = \csc^{-1}(x)$, then

$$\frac{dy}{dx} = -\frac{1}{\csc(y) \cot(y)}$$

Again,

$$\cot(y) = \begin{cases} -\sqrt{x^2 - 1} & x \geq 1 \\ \sqrt{x^2 - 1} & x \leq -1 \end{cases}$$

similarly for $\csc(y)$... Hence, one has

$$\frac{d}{dx} \csc^{-1}(x) = -\frac{1}{|x|\sqrt{x^2 - 1}}$$

# 7  Differentiability and Continuity

**Theorem 7.1.** Differentiability implies continuity.

*Proof.* Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable, then $\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = N < \infty$. Hence, if $\Delta f \not\to 0$, then the whole limit would blow up, so we know that $\Delta f \to 0$ as $\Delta x \to 0$. This is basically what the definition of continuity says. $\qquad \square$

**Theorem 7.2.** The converse of the preceding claim doesn't hold.

*Proof.* There is an simple but "philo" counterexample. Consider Thomae's Function, which is defined as

$$T(x) = \begin{cases} \dfrac{1}{q} & x = \dfrac{p}{q}, \gcd(p, q) = 1 \\ 0 & \text{Otherwise} \end{cases}$$

I claim without proof that $T(x)$ is continuous on $\mathbb{R} \setminus \mathbb{Q}$ (continuous at every irrational number) but it's nowhere differentiable. $\qquad\square$

Google "Thomae's Function" for more information.

# 8 L'Hopital's Rule

**Theorem 8.1.** Suppose $f, g : \mathbb{R} \to \mathbb{R}$ are two differentiable functions, then if **either**

$$\frac{f(x)}{g(x)} = \frac{0}{0} \text{ OR } \frac{f(x)}{g(x)} = \frac{\pm\infty}{\pm\infty}$$

then

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)}$$

provided that the limit of the second term exists.

Here are some examples of applications of L'Hopital's rule.

**Example 4.** The limit of ational functions using L'Hopital's rule:

$$\lim_{x \to \infty} \frac{ax + b}{cx + d} = \lim_{x \to \infty} \frac{a}{c} = \frac{a}{c}$$

Using standard method,

$$\lim_{x \to \infty} \frac{ax + b}{cx + d} = \lim_{x \to \infty} \frac{a + \frac{b}{x}}{c + \frac{d}{x}} = \frac{a}{c}$$

**Exercise 1.** The example above demonstrates how to manipulate the limit of the quotient of two linear function. Consider the quotient of two polynomial. Can you deduce a general result of the limit using L'Hopital's rule? (Remember! L'Hopital's rule can be applied repetitively.)

**Example 5.** Consider $\lim_{x \to -\infty} xe^x$. This function does not have the form of a quotient, and the limit looks like $-\infty \times 0$, which implies nothing. However, this does not imply that the limit is unsolvable. Our goal is to rewrite the expression in quotient form and apply L'Hopital's rule.

Let's try this

$$L = \lim_{x \to -\infty} \frac{e^x}{1/x}$$

Apply L'Hopital's rule,

$$L = \lim_{x \to -\infty} \frac{e^x}{-1/x^2} = \lim_{x \to -\infty} \frac{e^x}{2/x^3} = \dots$$

which gives us nothing, so let's try the other way.

$$L = \lim_{x \to -\infty} \frac{x}{e^{-x}} = \lim_{x \to -\infty} \frac{1}{-e^{-x}} = 0$$

**Example 6.** Consider

$$S = \lim_{x \to 0^+} (x^2 \ln^2(x))$$

then

$$S = \lim_{x \to 0^+} \frac{\ln^2(x)}{x^2} = \lim_{x \to 0^+} \frac{2\ln(x)\frac{1}{x}}{-2/x^2} = \lim_{x \to 0^+} \frac{\ln(x)}{-1/x} = \lim_{x \to 0^+} \frac{1/x}{1/x^2} = \lim_{x \to 0^+} x = 0$$

**Example 7.** When there is a product, you should attempt to rewrite it as a quotient in order to apply L'Hospital's rule. Remember that your intuition plays a very important role.

$$\lim_{x \to \infty} x \sin\left(\frac{7}{x}\right) = \lim_{x \to \infty} \frac{\sin(7/x)}{1/x} = \lim_{x \to \infty} \frac{\cos(7/x)\frac{-7}{x^2}}{-1/x^2} = \lim_{x \to \infty} 7\cos(7/x) = 7$$

These two examples show that practice is really important. The following two examples will demonstrate the techniques of using logarithmic function to evaluate the limit, which is usually applied to an exponential form.

**Example 8.** Consider

$$y = \lim_{t \to 0^+} (e^t + t)^{1/t}$$

This does not belong to any familiar form of expression, but since there is an exponent, we can use logarithmic function to "pull it down".

$$\ln(y) = \lim_{x \to 0^+} \frac{e^t + t}{t} = \lim_{x \to 0^+} \frac{e^t + 1}{1} = 2$$

So, $y = e^{\ln(y)} = e^2$, meaning that the limit is $e^2$.

**Example 9.** Suppose $z = \lim_{x \to \infty}(e^{-2t} + 3t)^{1/t}$, then

$$\ln(z) = \lim_{x \to \infty} \frac{e^{-2t} + 3t}{t} = \lim_{x \to \infty} \frac{-2e^{-2t} + 3}{1} = 3$$

Therefore, the limit is $z = e^3$.

We've already seen some tricks of manipulating expressions, but in some cases, you need to apply L'Hospital's rule multiple times (usually, 1 to 2 times is enough).

**Example 10.** Consider

$$\lim_{y \to \infty} \frac{y^2 - e^{6y}}{4y^2 + e^{7y}}$$
$$= \lim_{y \to \infty} \frac{2y - 6e^{6y}}{8y + 7e^{7y}}$$
$$= \lim_{y \to \infty} \frac{2 - 36e^{6y}}{8 + 49e^{7y}}$$
$$= \lim_{y \to \infty} \frac{-216e^{6y}}{343e^{7y}}$$
$$= \lim_{y \to \infty} \frac{-216}{343e^y}$$
$$= 0$$

There is an example in Haese textbook that shows how to find the limit of $\dfrac{e^x}{x^n}$. Take a look at it.

# 9 Taylor Series

This is a method to approach a function using polynomials. I'll first give the definition and discuss why it's like that.

**Definition 2.** Suppose $f : (a, b) \to \mathbb{R}$ is continuously differentiable at $x = c$ (the derivatives of all degree exists for $x = c$), then the Talor series of $f$ at $c$ is

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

There is an important quantity to be considered, which is the remainder $R_n$ defined as

$$R_n(x) = f(x) - \sum_{m=0}^{n-1} \frac{f^{(m)}(c)}{m!}(x-c)^m$$

The term $R_n$ is an error term that measure how close is the original function $f$ and the polynomial approximation. If $R_n$ tends to zero as $n \to \infty$, then we say that $f(x) = \sum_{m=0}^{\infty} \frac{f^{(m)}(c)}{m!}(x-c)^m$ (vice versa).

There's always a question at this point: Why is that power series approximate $f$? Here is the answer. Consider a power series

$$f(x) = \sum_{m=1}^{\infty} a_m x^m$$

such that $f$ converges for $|x| < R$, where $R$ is the radius of convergence, then because this power series is NICE when $|x| < R$, we can differentiate it term by term, i.e.,

$$\frac{d}{dx}f(x) = \frac{d}{dx}\sum_{m=1}^{\infty} a_m x^m = \sum_{m=1}^{\infty} \frac{d}{dx}a_m x^m$$

**Your task is to observe the pattern for $f'$, $f''$, $f'''$ and conclude what is $f^{(l)}$.** (This is your task because I'm lazy...)

Answer:

$$f^{(l)}(x) = \sum_{m=l}^{\infty} m(m-1)...(m-l+1)a_m x^{m-l}$$

Substitute $x = 0$ gives you a factorial. To offset the impact of differentiating (because it keeps changing the coefficient), we need to divide each term by $m!$. Hence,

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

Another potential problem: What is the difference between Taylor Series and Maclaurin Series? Well, the difference is subtle because Taylor series refers to the series expansion of a function at some point in its domain, whereas Maclaurin series refers to the Taylor series at $x = 0$.

## 9.1 Taylor Series of Trigonometric Functions (Incomplete)

We start by looking at the Maclaurin series of $\sin(x)$.

At $x = 0$, $\sin(x) = 0$ meaning that the constant term of its expansion is also zero. $(\sin(0))' = \cos(0) = 1$ shows that the degree 1 term of its expansion is 1. $(\sin(0))'' = -\sin(0) = 0$, so the quadratic term must be zero. In other words there's no proper quadratic approximation for $\sin(x)$ at $x = 0$. Then, we have $(\sin(0))''' = -\cos(0) = -1$, meaning that the coefficient for the cubic term is $\frac{-1}{3!}$. Repeat this process, you'll see that

$$\sin(x) = 0 + x + 0x^2 - \frac{1}{3!}x^3 + 0x^4 + \frac{1}{5!}x^5 + \dots$$

Hence,

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

This also shows that $\sin(x)$ is an odd function (connect to what Karbro said).

Then, we're going to look at $\cos(x)$.

The process is similar, so it will be left to the readers to complete. (I'm lazy...)

$$\cos(x) = 1 + 0x - \frac{1}{2!}x^2 + 0x^3 + \frac{1}{4!}x^4 + \dots$$

which means

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

### 9.1.1 Challenge

Can you derive the Maclaurin sereis of $\tan(x)$, $\sec(x)$, $\csc(x)$, and $\cot(x)$? I can't..... :(

### 9.1.2 Answer to the challenge

:)

Suppose $\tan(x) = \sum_{m=0}^{\infty} a_m x^m$, then

1. $\tan(0) = 0$ implies $a_0 = 0$;

2. $(\tan(0))' = \sec^2(0) = 1$ implies $a_1 = 1$;

3. $(\tan(0))'' = 0$ implies $a_2 = 0$;

4. $(\tan(0))''' = 6\sec^4(0) - 4\sec^2(0) = 2$ implies $a_3 = 2/6$;

5. ......

So,

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \dots$$

I'm not sure how to rewrite this using sigma notation...... :(

## 9.2 Taylor Series of Exponential and Logarithmic Function

A simple fact allows us to conclude the Maclaurin series of $e^x$ really quickly. Note that $\frac{d}{dx}e^x = e^x$, so

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

Now, for logarithmic function, we only care about the natural log because we can always change the base. However, $\ln(x)$ is undefined at $x = 0$ so we're going to consider $\ln(x+1)$ and derive its Maclaurin series. Suppose $\ln(1+x) = \sum_{m=0}^{\infty} b_m x^m$, then

1. $\ln(0+1) = 0$ implies $b_0 = 0$;

2. $(\ln(0+1))' = \frac{1}{0+1} = 1$ implies $b_1 = 1$;

3. $(\ln(0+1))'' = -\frac{1}{(0+1)^2} = -1$ implies $b_2 = -\frac{1}{2!}$;

4. $(\ln(0+1))''' = 2\frac{1}{(0+1)^3} = 2$ implies $b_3 = \frac{2}{3!}$;

5. ......

6. $(\ln(0+1))^{(k)} = (-1)^{k+1}(k-1)!$ implies $b_k = \frac{(-1)^{k+1}(k-1)!}{k!} = \frac{1}{k}$;

So,

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

## 9.3 Proof of De Moivre's Theorem

We learned in the chapter of complex number that $e^{ix} = \cos(x) + i\sin(x)$. We're going to give a proof using Maclaurin series.

*Proof.*

$$e^{ix} = \sum_{m=0}^{\infty} \frac{1}{m!}(ix)^m$$

$$= 1 + ix + \frac{1}{2}i^2x^2 + \frac{1}{3!}i^3x^3 + \dots$$

$$= 1 + ix - \frac{1}{2}x^2 - i\frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \dots$$

$$= \left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots\right) + i\left(x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \dots\right)$$

$$= \cos(x) + i\sin(x)$$

$\square$

## 9.4   Typical Questions

Given an aribitrary analytic function $f : \mathbb{R} \to \mathbb{R}$ and the $n$-th degree taylor polynomial of $f$ at $x = 0$ is given by $p(x)$. Then, you should know the value for $f^{(k)}(0)$ for $k \leq n$.

# 10   Integration

The author assumes that the readers know all basic anti-derivatives. To check whether you can adroitly apply them, consider the following lists of questions, give yourself 10 seconds to answer each question. If you can do it, then congrats, you may advance to the next paragraph. **State the indefinite integral of all the following functions**

1. $ax^b$

2. $\sin(x)$

3. $\cos(x)$

4. $a\sin(x)$

5. $a\cos(x)$

6. $\sec^2(x)$

7. $-\csc^2(x)$

8. $e^x$

9. $1/x$

However, when the function is more complicated, we need some more advanced techniques to figure out the answer.

## 10.1   Integration by substitution

The strategy is really simple. Suppose we have a continuous function $h : \mathbb{R} \to \mathbb{R}$ that is of the form $h(x) = f(x)f'(x)$, and we want to integrate with respect to $x$, meaning that

$$\int h(x)dx = \int f(x)f'(x)dx$$

If we substitute a new variable $u = f(x)$, then $du = f'(x)dx$, which turns this integral into

$$\int u du$$

(integrating with respect to $u$) Although after you you figure out what to substitute and how to substitute, the integral becomes fairly easy, you still need lots of practice in order to get a kind of intuition, i.e., you

need to perceive that there exists a function $f$ and its derivative $f'$ in the integrand. Whenever you see some "complicated structure" in the integrand, if it is a polynomial of degree $n$, then you may want to substitute it as $u$ if there is a polynomial of degree $n-1$ outside of that structure.

**Example 11.** We should substitute $u = x^2 + 1$ in the following case.

$$\int x\sqrt{x^2 + 1}dx$$

which then becomes

$$\int \frac{1}{2}\sqrt{u}du$$

because $du = 2xdx$, so we multiply a factor of $1/2$ to eliminate the effect of 2.

Also, make that you can sense the existence of the derivative of natural log.

**Example 12.** Consider

$$\int \frac{\ln(x)}{x}dx$$

By substituting $u = \ln(x)$, the integral actually becomes

$$\int udu$$

### 10.1.1  Trigonometric Substitution

What the author (I) found sometimes difficult is trigonometric substitution because it requires a quite adroit manipulation of trigonometric identities.

**Example 13.** Consider

$$I = \int \frac{dx}{\sqrt{1 - x^2}}$$

If you can write the answer down in two seconds, then this example is not for you because you already master the differentiation of inverse trigonometric functions. (Congratulations!) Suppose you don't the answer, but you know the strategy is substitution. The most obvious two are probably $\sqrt{1 - x^2}$ and $1 - x^2$, but in either case $du$ makes the expression really ugly, so don't do it. Instead, you should think about how to get rid of the square root sign. Here, the identity $\sin^2(\theta) + \cos^2(\theta) = 1$ does a great job. Let's see. Substitute $x = \sin(\theta)$, so that

$$I = \int \frac{dx}{\sqrt{1 - \sin^2(\theta)}}$$

We also need something to replace $dx$, but it's easy because $x = \sin(\theta) \implies dx = \cos(\theta)d\theta$. So,

$$I = \int \frac{\cos(\theta)d\theta}{\cos(\theta)} = \int d\theta = \theta + C = \arcsin(x) + C$$

**Exercise 2.** Find the following indefinite integrals by trigonometric substitution.

1. $\int \frac{dx}{1 + x^2}$

2. $\int \frac{-dx}{\sqrt{1 - x^2}}$

Now comes a challenge.

**Example 14.** We want to find

$$I = \int \frac{dx}{\sqrt{1 + x^2}}$$

It is tempting to substitute sine or cosine in to the expression because this is similar to the previous example,but no, we use $x = \tan(\theta)$ because $1 + \tan^2(\theta) = \sec^2(\theta)$ completes the square.After a series of manipulation,

$$I = \int \sec(\theta)d\theta = \ln|\sec(\theta) + \tan(\theta)| + C$$

we know that $\tan(\theta) = x$ by substitution. The value of $\sec\theta$ requires more work. Consider an right-angled triangle with opposite $x$ and adjacent 1, then the hypotenuse is $\sqrt{1 + x^2}$ by Pythagora's Theorem. This shows that $\sec\theta = \sqrt{1 + x^2}$, so the final answer is

$$\ln|x + \sqrt{1 + x^2}| + C$$

The integration of secant function is in the textbook (Pearson). Do check it out.

The strategy of substituting trigonometric terms doesn't only work for $x$ but actually in general. Whenever there is a square root sign or things like $1 + [f(x)]^2$ in the denominator, we always try to substitute trig terms to simplify it. Here is a sample question from Pearson's textbook.

**Example 15.** We want to find

$$\int \frac{e^{-2x}dx}{\sqrt{1 - e^{-4x}}}$$

Now, substitute $e^{-2x} = \sin(\phi)$, then $dx = \frac{\cos(\phi)d\phi}{-2e^{-2x}}$, which implies that $e^{-2x}dx = -\frac{1}{2}\cos(\phi)d\phi$. Thus, the integral turns out to be

$$-\frac{1}{2}\int \frac{\cos(\phi)d\phi}{\cos(\phi)} = -\frac{1}{2}\phi + C = -\frac{1}{2}\arcsin(e^{-2x}) + C$$

We want to generalize these trigonometric substitution as follows.

**Example 16.** Consider

$$I = \int \frac{dx}{\sqrt{a^2 - x^2}}$$

We know that $a^2\sin^2(\theta) + a^2\sin^2(\theta) = a^2$, so we'll substitute $x = a\sin\theta$ to obtain

$$\int \frac{a\cos\theta d\theta}{a\cos\theta} = \int d\theta = \arcsin\left(\frac{x}{a}\right) + C$$

If we remove the square root sign, then this substitution still works because the denominator becomes $a^2\cos^2(\theta)$, and this is a simple integral that we can solve. If there is a minus sign in the integrand, then we usually substitute $\cos(\theta)$ because its derivative contains a minus sign to cancel out the minus sign in the integrand.

**Example 17.** Now, we consider another class of integral

$$I = \int \frac{dx}{a^2 + x^2}$$

In this case, we usually consider substituting $x = a\tan(\theta)$ because the identity $a^2 + a^2\tan^2(x) = a^2\sec^2(x)$ will convert the sum in the denominator to a monomial. Thus,

$$I = \int \frac{a\sec^2\theta d\theta}{a\sec^2\theta} = \int d\theta = \theta + C = \frac{1}{a}\tan\left(\frac{x}{a}\right) + C$$

If we add a square root sign on the denominator, then we can also substitute tangent function. The only difference is that the denominator will become $a\sec(x)$.

You also need to anticipate the existence of constant in your substitution. If $I = \int x\sqrt{9 - x^2}$, then you can substitute $u = x^2$, but it still leaves you with the term $\sqrt{9 - u}$, which is not good, but because the derivative of a constant vanishes, you can always use $u = 9 - x^2$, and this helps you make progress. Also, you need to be able to identify the relationship between derivative and anti-derivative. For instance, test if you can identify what to substitute in $\int \frac{\csc(x)\cot(x)}{3 - \csc(x)}dx$. If you can't do it in two seconds, then you really need more practice (sry). Sometimes, the substitution is really complex. Don't be shocked or intimidated by the complex structure of the integrand.

**Example 18.** Consider the integral

$$I = \int (\sin(3x) - x^3)^5 (\cos(3x) - x^2) dx$$

The correct substitution is $u = \sin(3x) - x^3$.

**Example 19.** For this one,

$$4 \int \left( \frac{1}{x} - e^{-x} \right) \cos(e^{-x} + \ln(x)) dx$$

You use $u = e^{-x} + \ln(x)$.

**Example 20.** For

$$\int \frac{e^{\tan(x)}}{\cos^2(x)}$$

try substituting $u = \tan(x)$. You should be able to justify this substitution immediately.

### 10.1.2  Definite integration

Consider

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

we want to use substitution $u = g(x)$ so that $du = g'(x) dx$. Then,

$$I = \int_{x=a}^{x=b} f(g(x)) g'(x) dx = \int_{x=a}^{x=b} f(u) du = \int_{g(a)}^{g(b)} f(u) du$$

The reason for changing the parameter is because we want to integrate with respect to the variable $u$, which is determined by $x$.

## 10.2  Integration by part

We start by an example. Consider the indefinite integral

$$\int x e^x dx$$

In this case, there is no way to use substitution. For example, if substitute $u = x$, then it's basically useless. On the other hand, if you try substitute $u = e^x$, then $x = \ln(u)$, which is no good either. So, probably integration by substitution is not a good idea to do this one.

Think about the rules of differentiation and that integration is the reversed operation of differentiation. Substitution corresponds to undo the effect of chain rule. Can we reverse the product rule? Well, yes.

Suppose we have a product $f(x) = u(x) v(x)$ such that $f$ is differentiable, then by the product rule, we have

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$$

Now, we take the integral on both side

$$\int \frac{d}{dx}(uv) dx = \int u\frac{dv}{dx} dx + \int v\frac{du}{dx} dx = \int u\, dv + \int v\, du$$

Let's rearrange the equation, we found that

$$\int u\, dv = uv - \int v\, du$$

This gives the formula for integration by parts, which says that if the integrand has the form $u\, dv$, then you can find the indefinite integral by setting proper $u$ and proper $dv$ to convert the integral into something you know about. Now, let's return to the example $xe^x$.

Let's try setting $u = e^x$ and $dv = xdx$. Then, $du = e^x dx$ and $v = \frac{1}{2}x^2$, which means

$$\int xe^x dx = \frac{1}{2}x^2 e^x - \frac{1}{2}\int x^2 e^x dx$$

Ooops! We don't know the integral of $x^2 e^x$ yet, and that thing actually needs another integration by parts, which is not so convenient, so let's try the other way.

Let $u = x$ and $dv = e^x dx$, then $du = dx$ and $v = e^x$. Therefore, we have

$$\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x + C$$

DON'T FORGET THE ARBITRARY CONSTANT!

Sometimes, the two integrands are not obvious at all, so you might need to invoke $dv = dx$ in order to proceed integration by parts.

**Example 21.** Consider

$$I = \int \sin(\ln(x))dx$$

If you try any kind of substitution, you will fail. Meanwhile, this function is not a product, so it's time to invoke $u = \sin(\ln(x))$ and $dv = dx$! Using integration by parts, we get

$$I = x\sin(\ln(x)) - \int x\cos(\ln(x))\frac{dx}{x} = x\sin(\ln(x)) - \int \cos(\ln(x))dx$$

which is really nice. Now, use integration again (the same strategy), one can obtain

$$I = x\sin(\ln(x)) - x\cos(\ln(x)) + \int -x\sin(\ln(x))\frac{dx}{x} = x\sin(\ln(x)) - x\cos(\ln(x)) - \int x\sin(\ln(x))dx$$

Therefore,

$$I = \frac{x\sin(\ln(x)) - x\cos(\ln(x))}{2} + C$$

**Example 22.** We try to prove an identity in this example. Suppose $f : [a,b] \to \mathbb{R}$ is twice differentiable and its second derivative $f''$ is Riemann integrable, then we have

$$\int_a^b f(x)dx = \frac{(b-a)(f(a)+f(b))}{2} + \frac{1}{2}\int_a^b f''(x)(x-a)(x-b)dx$$

We want to use integration by parts because of the form of the right hand side. So, set $u = f(x)$, $dv = dx$, we get

$$\int_a^b f(x)dx = [xf(x)]_a^b - \int_a^b xf'(x)dx = bf(b) - af(a) - \int_a^b xf'(x)dx$$

Now, we approach this question from the right hand side. The integral at then end can be written as

$$\frac{1}{2}\int_a^b f''(x)(x-a)(x-b)dx = -\frac{1}{2}\int_a^b f'(x)(2x-a-b)dx$$

using integration by part with $u = (x-a)(x-b)$ and $dv = f''(x)dx$. Moreover, we have

$$-\frac{1}{2}\int_a^b f'(x)(2x-a-b)dx = -\frac{1}{2}\left(\int_a^b 2xf'(x)dx - (a+b)\int_a^b f'(x)dx\right)$$

$$= -\int_a^b xf'(x)dx + \frac{af(b) + af(a) + bf(b) - bf(a)}{2}$$

Therefore,

$$\frac{(b-a)(f(a)+f(b))}{2} + \frac{1}{2}\int_a^b f''(x)(x-a)(x-b)dx = \frac{(b-a)(f(a)+f(b))}{2} - \frac{1}{2}\int_a^b f'(x)(2x-a-b)dx$$

$$= \frac{bf(a)+bf(b)-af(a)-af(b)}{2} + \frac{af(b)-af(a)+bf(b)-bf(a)}{2} - \int_a^b xf'(x)dx$$

$$= bf(b) - af(a) - \int_a^b xf'(x)dx$$

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

which completes the proof

Note, even though integration by parts is a really strong tool for finding the anti-derivative, it is still limited. Functions like the following one cannot be evaluated using any method that is covered so far.

$$I = \int \frac{\cos(x)}{x}dx$$

This integral is often denoted $Ci(x)$, but you can write down the anti-derivative using power series, which will be introduced later.

### 10.2.1 Induction Method

There might exist some questions that ask you to prove an integral formula by induction, in which case integration by parts will be an essential tool. Here are some questions that I found online.

## 10.3 Reduction formula

Here, I'll derive the reduction formula for the powers of trigonometric functions.

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

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

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

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

The reduction formula for cosine was discussed in class, so I'll skip that. haha. . .

$$\int \csc^n(x) = \int \csc^{n-2}(x)\csc^2(x)dx$$

$$= -\cot(x)\csc^{n-2}(x) - (n-2)\int -\cot^2(x)\csc^{n-2}(x)dx$$

$$= -\cot(x)\csc^{n-2}(x) - (n-2)\int \csc^{n-2}(x)(\csc^2(x)-1)dx$$

$$= -\frac{1}{n-1}\cot(x)\csc^{n-2}(x) + \frac{n-2}{n-1}\csc^{n-2}(x)dx$$

Here, we also have a very interesting reduction formula that requires some tricks. It appears as an extension when computing the variance of the continuous random variable $Y$ with the probability density function (PDF) $e^{-1/2y^2}$.

**Example 23.** (The writer recomment reviewing the definition of improper integral first.)

Define $I_n = \int_{-\infty}^{\infty} y^n e^{-\frac{1}{2}y^2} dy$. We want to have a reduction formula of $I_n$ in terms of $I_{n-2}$, and the hint is to use integration by parts. People are told in class that they should first consider setting the polynomial term in the integrand as $u$. However, it is not practical in this case because you can't find a closed form to represent the anti-derivative of $e^{-1/2y^2}$.

Why not try the other way around?

Suppose $u = e^{-\frac{1}{2}y^2}$ and $dv = y^n dy$, then $du = -ye^{-\frac{1}{2}y^2} dy$ and $v = \frac{1}{n+1}y^{n+1}$, which gives us

$$\int y^n e^{-\frac{1}{2}y^2} dy = \frac{1}{n+1}y^{n+1}e^{-\frac{1}{2}y^2} - \frac{1}{n+1}\int y^{n+1}(-y)e^{-\frac{1}{2}y^2} dy$$

$$= \frac{1}{n+1}y^{n+1}e^{-\frac{1}{2}y^2} + \frac{1}{n+1}\int y^{n+2}e^{-\frac{1}{2}y^2} dy$$

This shows that

$$I_n = \lim_{b\to\infty}\left[\frac{1}{n+1}y^{n+1}e^{-\frac{1}{2}y^2}\right]_0^b + \lim_{b\to-\infty}\left[\frac{1}{n+1}y^{n+1}e^{-\frac{1}{2}y^2}\right]_0^b + I_{n+2}$$

Apply L'Hopital's rule (multiple times), we have

$$\lim_{b\to\infty}\frac{y^{n+1}}{(n+1)e^{\frac{1}{2}y^2}} = 0 \quad \lim_{b\to-\infty}\frac{y^{n+1}}{(n+1)e^{\frac{1}{2}y^2}} = 0$$

Hence,

$$I_n = \frac{1}{n+1}I_{n+2} \implies (n+1)I_n = I_{n+2}$$

Now, we just replace $n$ by $n-2$ to get

$$I_n = (n-1)I_{n-2}$$

**Can we work out the exact expression?** Well if take the facts $I_0 = 1$, $I_1 = 0$ for granted, then we can see that $I_{2n+1} = 0$ for all $n \in \mathbb{Z}^+$. This makes sense because such function is odd. Also, $I_{2n} = (2n-1)(2n-3)\cdots 3 \cdot 1$. In terms of factorial, we have

$$I_{2n} = \frac{(2n)!}{n!2^n}$$

## 10.4 Volume of Revolution

## 10.5 Foundation

Given a continuous function $f(x)$ over an interval $[a, b]$, we want to find the volume obtained by rotating the area bounded by $f(x)$, $x = a$, $x = b$, $y = 0$ with respect to the $x$-axis.

Suppose we have a partition $P = \{a = x_0 < x_1 < \ldots < x_n = b\}$ of the interval $[a, b]$. Clearly, we have

$$\sum_{i=1}^{n} \pi m_i^2 (x_i - x_{i-1}) \leq V \leq \sum_{i=1}^{n} M_i(x_i - x_{i-1})$$

Take the limit of upper sum and lower sum and apply squeeze lemma, we know that

$$V = \int_a^b \pi f(x)^2 dx$$

There are some tricky integrand related to this kind of integrals.

**Example 24.** Consider

$$\pi \int_1^e (\ln x)^2 dx$$

. We use integration by parts to get

$$\pi \int_1^e (\ln x)^2 dx = \pi([x \ln^2 x]_1^e - 2 \int_1^e x \ln x \frac{1}{x} dx)$$

$$= \pi([x \ln^2 x]_1^e - 2 \int_1^e \ln x dx)$$

$$= \pi(e - 2)$$

Remember that $\int \ln x dx = x \ln x - x + C$!

## 10.6   Improper Integral

Let $f : [a, \infty) \to \mathbb{R}$ be a real-valued function. If $\int_a^b f(x)dx$ exists for all $b > a$, and $\lim_{b \to \infty} \int_a^b f(x)dx$ exists, then we define the improper integral to be

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

**Remark 1.** Note that if $f : [a, \infty) \to \mathbb{R}$ is continuous, then $\int_a^b f(x)dx$ automatically exists for all $b > a$.

**Example 25.** An example of convergent improper integral.

$$\int_1^\infty \frac{dx}{x^2} = \lim_{b \to \infty} \frac{dx}{x^2} = \lim_{b \to \infty} \left[ \frac{-1}{x} \right]_1^b = 0 - (-1) = 1$$

**Example 26.** An example of divergent improper integral.

$$\int_1^\infty \frac{dx}{x} = \lim_{b \to \infty} \int_1^b \frac{dx}{x} = \lim_{b \to \infty} [\ln(x)]_1^b = \infty$$

**Example 27.** An example of undefined improper integral

$$\int_0^\infty \cos x dx$$

The reason is that $\lim_{x \to \infty} \sin(x)$ doesn't exist...

### 10.6.1   p-integrals

$p$-integrals refer to the following class of improper integrals.

$$\int_1^\infty \frac{dx}{x^p}$$

**Theorem 10.1.** $\int_1^\infty \frac{dx}{x^p}$ converges if $p > 1$ and diverges if $0 < p \leq 1$. For negative exponents, there is a similar result. The readers can state and prove by themselves

*Proof.* Suppose $p = 1$, then the integral diverges (one of our example). Suppose $0 < p \leq 1$, then $\frac{1}{x^p} \geq \frac{1}{x}$, which implies that the integral diverges. If $p > 1$, then we have

$$\int_1^\infty \frac{dx}{x^p} = \lim_{b \to \infty} \int_1^b \frac{dx}{x^p}$$

$$= \lim_{b \to \infty} \left[ \frac{1}{-p+1} x^{-p+1} \right]_1^b = \lim_{b \to \infty} \left( \frac{1}{-p+1} b^{-p} - \frac{1}{-p+1} \right)$$

$$= \frac{1}{1-p} < \infty$$

$\square$

**Theorem 10.2.** $\int_0^1 x^p dx$ converges if $p > -1$.

*Proof.* Let $p > -1$, we have

$$\int_0^1 x^p dx = \lim_{c \to 0} \int_c^1 x^p dx$$

$$= \lim_{c \to 0} \left[ \frac{1}{p+1} x^{p+1} \right]_c^1$$

$$= \lim_{c \to 0} \left( \frac{1}{p+1} - \frac{c^{p+1}}{p+1} \right)$$

$$= \frac{1}{p+1}$$

$\square$

The converse of the theorems above are left to the readers.

## 10.7 Tricky part

When evaluating integrals of the form $\int_{-\infty}^{\infty} f(x)dx$, we need to check the convergence of $\int_a^{\infty} f(x)dx$ and $\int_{-\infty}^a f(x)dx$. This is not equivalent to checking the existence of $\lim_{b \to \infty} \int_{-b}^b f(x)dx$ (This is often referred to as Cauchy Principal Value, or CPV).

## 10.8 Prepare for Normal Distribution

In this section, we will evaluate the integral $I = \int_{-\infty}^{\infty} e^{-x^2} dx$. It should be noted that there is no proper expression for the antiderivative of $f(x) = e^{-x^2}$, but we still have alternative method to access this improper integral. Consider the volume of revolution: Rotate $z = f(x)$, $x \geq 0$ with respect to the $z$-axis. The volume of revolution can be obtained By

$$\int_0^{\infty} 2\pi x e^{-x^2} dx = 2\pi \lim_{b \to \infty} \int_0^b x e^{-x^2} dx = 2\pi \lim_{b \to \infty} \left[ -\frac{1}{2} e^{-x^2} \right]_0^b = \pi$$

On the other hand, we can also integrate this using multivariable calculus. The surface formed by rotating the $z = f(x)$ can be considered as a function $g : \mathbb{R}^2 \to \mathbb{R}$ given by $g(x,y) = e^{-(x^2+y^2)}$. Hence, the volume can be expressed as

$$\iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2} dxdy = \int_{-\infty}^{\infty} e^{-y^2} \left( \int_{-\infty}^{\infty} e^{-x^2} dx \right) dy = \left( \int_{-\infty}^{\infty} e^{-x^2} dx \right) \left( \int_{-\infty}^{\infty} e^{-y^2} dy \right) = I^2$$

This shows that $I^2 = \pi$, so $I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

# 11 Differential Equation

A differential equation is an equation of the form

$$F\left( x, y, \frac{dy}{dx} \right) = 0$$

A solution to the differential equation is a function $\varphi(x) : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ is an open interval, such that $\varphi(x)$ satisfies the differential equation. An important note here is that the solution needs to be differentiable over the interval.

## 11.1 Variable Separation

If you have a differential equation that can be written in the form of

$$N(y)\frac{dy}{dx} = M(x)$$

then it is said to be **separable**. In this case, we may integrate on both sides w.r.t $x$ so that

$$\int N(y)\frac{dy}{dx}dx = \int M(x)dx$$

Clearly, $y$ is a function of $x$ (this is a integration substitution), so the left hand side can be simplified To

$$\int N(y)dy = \int M(x)dx$$

## 11.2 Homogeneous first order ODE

A function $f : V \to W$ between two vector spaces is homogeneous of degree $n$ if

$$f(t\vec{w}) = t^n f(\vec{w})$$

where $t$ is a scalar. In particular, if we consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ that is homogeneous of degree 0, then it only depends on the quotient $y/x$.

*Proof.* By homogenity, we have

$$f(tx, ty) = t^0 f(x, y) = f(x, y)$$

Substitute $t = 1/x$, we get

$$f(1, y/x) = (1/x)^0 f(x, y) = f(x, y)$$

Let $g(y/x) = f(1, y/x)$. The claim then follows directly. $\qquad\square$

## 11.3 Linear first order ODE

## 11.4 Non-homogeneous first order ODE

## 11.5 Seoncond Order ODE (bonus section)

# 12 Statistics

## 12.1 Basics

## 12.2 Geometric Distribution

Through out this section, we assume that the measurement is on the success of a variable. The expected value of a variable $Y$ that admits geometric distribution, usually denoted as $Y \sim \text{Geo}(p)$, can be computed by

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

To evaluate this sum, note that we can factor $p$ out, which leaves

$$E(Y) = p\sum_{n=1}^{\infty} nq^{n-1}$$

Now, consider the series $f(x) = \sum_{n=1}^{\infty} nx^{n-1}$. It looks like the derivative of some function. For any $x < 1$ (it is a probability), $f(x)$ is convergent, so we can integrate it term by term to find its anti-derivative to be $\sum_{n=1}^{\infty} x^n = \frac{1}{1-x}$ for $|x| < 1$. Hence, we have

$$E(Y) = p\frac{d}{dx}\left(\frac{1}{1-x}\right)\Bigg|_{x=q} = p\frac{1}{(1-q)^2} = p\frac{1}{p^2} = \frac{1}{p}$$

## 12.3 Binomial Distribution

The mean:

$$E(Y) = \sum_{i=1}^{n} i\binom{n}{i}p^n q^{n-i}$$

We use a similar trick (differentiation) to obtain the result of the summation above.

## 12.4 Probability Density Function (PDF)

*Throughout this section, $Y$ will denote a continuous random variable that takes values in $\mathbb{R}$.*
The writer assumes that the readers know what a probability density function (PDF) means. It is CERTAINLY different from Portable Document Format (PDF), and sometimes people need to distinguish between them. There are two important properties of a PDF $f(y)$:

- $\int_{-\infty}^{\infty} f(y)dy = 1$. Note that this is analogous to discrete random variable.

- $f(y) \geq 0, \forall y \in \mathbb{R}$. This is required because the probability can never take negative value.

The expectation $E(Y)$ and the variance $\text{Var}(Y)$ can all be defined analogously (compare to the case of discrete variable).

- $E(Y) = \int_{-\infty}^{\infty} yf(y)dy$. We say that the expected value exist if the integral converges to a finite value.

- $\text{Var}(Y) = \int_{-\infty}^{\infty} y^2 f(y)dy - \left(\int_{-\infty}^{\infty} yf(y)dy\right)^2$. We say that the variance exist if the integral converges to a finite value.

Here, the writer document some important examples presented in the class.

**Example 28.** Consider the function

$$f(y) = \begin{cases} k\sqrt{-x^2 + 6x - 5} & 1 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$$

We want to find $k \in \mathbb{R}$ s.t. $f(y)$ is a PDF. The lecturer gives us a hint: think geometrically.

$$\int_{-\infty}^{\infty} f(y)dy = k\int_{1}^{5} \sqrt{-x^2 + 6x - 5}$$

You will not be willing to compute this integral explicitly, so we will try to figure out the geometry of the pdf.

$$\frac{1}{k}f(y) = \sqrt{-x^2 + 6x - 5} = \sqrt{4 - (x-3)^2} = \sqrt{2^2 - (x-3)^2}$$

This is the equation of a half-circle centered at $(3, 0)$ with radius 2. Therefore,

$$k\int_{1}^{5} \sqrt{-x^2 + 6x - 5} = 2k\pi = 1$$

which implies that $k = \frac{1}{2\pi}$.

**Example 29.** will complete this part later...

## 12.5 Cumulative Density Function (CDF)

By definition, the CDF of a continuous random variable is

$$\int_{-\infty}^{y} f(t)dt$$

where $f(y)$ is the PDF of $Y$. Consequently,

$$f(y) = \frac{dF(y)}{dy}$$

where $F(y)$ is the cumulative density function. **It should be noted that $F(y)$ should always be continuous and monotonically increasing, so whenever a PDF is given as a piecewise function, we should be careful about the end points when deriving the formula of CDF.**

Some equalities:
$$P(Y \leq y_0) = F(y_0)$$
$$P(y_0 \leq Y \leq y_1) = F(y_1) - F(y_0)$$

# 13 Geometry

We are working in the Euclidean space!!!

## 13.1 Dot Product

Suppose $\vec{p}, \vec{q} \in \mathbb{R}^n$, then their dot product inside

$$\vec{p} \cdot \vec{q} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_n \end{bmatrix} \cdot \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_n \end{bmatrix} = \sum_{i=1}^{n} p_i q_i$$

A very important equation to be kept in mind is

$$\vec{p} \cdot \vec{q} = \|p\| \|q\| \cos \theta$$

where $\theta$ is the angle between the two vectors.

## 13.2 Cross Product

In $\mathbb{R}^3$, there is a reason for defining the cross product! We want to be able to find a third vector perpendicular to the given two! Here is the derivation, which I'll complete later.

## 13.3 Distance

It means what it means.

### 13.3.1 Lines

For two parallel lines that do not coincide, we use the following method. Suppose

$$l_1 : \vec{p_1} + \lambda_1 \vec{d_1}$$
$$l_2 : \vec{p_2} + \lambda_2 \vec{d_1}$$

Moreover, assume that the magnitude of the directional vector is 1, then we know that $D = \|\vec{p_2} - \vec{p_1}\| \sin \theta$, where $\theta$ is the angle between $\vec{p_1} - \vec{p_1}$ and the directional vector. This simply shows that

$$D = \|(\vec{p_2} - \vec{p_1}) \times \vec{d_1}\|$$

Note that if the magnitude of the directional vector is not equal to 1, then you need to take $\vec{d_1}/|\vec{d_1}|$, which normalizes the magnitude to 1. (Very Crucial!) Suppose $l_1$ and $l_2$ are skewed, i.e., $\vec{d_1} \times \vec{d_2} \neq 0$, then we can project $\vec{p_2} - \vec{p_1}$ onto the the vector $\vec{d_1} \times \vec{d_2}$.

$$D = \left\| \frac{\vec{d_1} \times \vec{d_2}}{\|\vec{d_1} \times \vec{d_2}\|} (\vec{p_2} - \vec{p_1}) \right\|$$

### 13.3.2 Line and Plane

Suppose

$$l : \vec{p} + \lambda \vec{d}$$
$$\Pi : ax + by + cz = d$$

Here is an important note. If the plane is defined using parametric equation, then you need to convert it into a cartesian equation before performing the following process! To avoid triviality, we assume that $l$ and $\Pi$ do not intersect. Let $\vec{n} = (a, b, c)$ be the normal vector and $\vec{q} = (x_0, y_0, z_0)$ be a point on the plane.

$$D = \frac{|(\vec{p} - \vec{q}) \cdot \vec{n}|}{\|\vec{n}\|}$$

### 13.3.3 Planes

Suppose

$$P : ax + by + cz = d$$
$$P' : a'x + b'y + c'z = d'$$

If two planes do not intersect, then they must be parallel, and their normal vectors should be co-linear, which implies that

$$(a', b', c') = \lambda(a, b, c)$$

for some $\lambda \in \mathbb{R}$. Let $\vec{p}, \vec{p'}$ be points on $P, P'$ respectively, then

$$D = \frac{|(\vec{p} - \vec{p'}) \cdot \vec{n}|}{\|\vec{n}\|}$$

By linearity of dot product, we can further simplify the the expression above.

$$D = \frac{|d - d'|}{\|\vec{n}\|}$$

### 13.3.4 Bonus

Supppose we have two $(n-1)$-dimensional "planes" inside an $n$-dimensional space $\mathbb{R}^n$, then if they are "parallel", then we can still apply the same formula for two planes but in a more general case (where the dot product are generalized).

# 14 Complex Numbers

It's denoted $\mathbb{C}$. Usually the set of complex numbers is $\{a + bi \mid a, b \in \mathbb{R}, \ i^2 = 1\}$. Since the arithmetics of $\mathbb{C}$ is too easy, I'll skip that part.

## 14.1 Different "forms" of Complex Numbers

- Cartesian form: This is just $z = a + bi$. We call $a$ the real part of $z$ and $b$ the imaginary part of $z$.

- Polar form: This usually refers to $z = re^{i\theta}$ for $\theta \in [0, 2\pi)$. In this way, any complex number can be uniquely expressed using a polar form. Note that polar form is especially useful for solving roots of unity.

Sometimes, we need to convert between the cartesian form and the polar form. It can be easily deduced using Pythagorean Theorem.

## 14.2 De Moivre's Theorem

I don't like this name. It's just a fancy term of a lemma. Here is the statement

**Theorem 14.1.** For any real number $x$ and a positive integer $n$, we have

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx$$

*Proof.* We use induction. First,

$$P(n) := (\cos x + i \sin x)^n = \cos nx + i \sin nx$$

By direct computation, the statement holds trivially for $n = 1$, so suppose it's true for some $k \geq 1$. then we have

$$(\cos x + i \sin x)^{k+1} = (\cos x + i \sin x)^k(\cos x + i \sin x) = (\cos kx + i \sin kx)(\cos x + i \sin x)$$
$$= \cos kx \cos x - \sin kx \sin x + i(\cos kx \sin x + \cos x \sin kx)$$
$$= \cos(kx + x) + i \sin(kx + x)$$
$$= \cos((k+1)x) + i \sin((k+1)x)$$

which shows that $P(k+1)$ is true. Hence, $P(n)$ is true for all $n \in \mathbb{Z}^+$. $\qquad\square$

The key of this proof is to use the sum formula of sine and cosine in the induction step.

### 14.3 Solve the roots of polynomials in complex numbers

Consider $x^n = re^{i\theta}$, where $\theta \in [0, 2\pi)$. We know that this equation has exactly $n$ solutions in $\mathbb{C}$, and they are

$$r^{1/n} e^{i\frac{\theta + 2k\pi}{n}}, k = 0, 1, ..., n-1$$

# 15 Some Problem Solving

## 15.1 The number of derrangement

Consider the set $S = \{1, ..., n\}$, which contains $n$ elements. We want to count the number of derrangements of this set, i.e., the number of bijective $f : S \to S$ such that $f(i) \neq i \forall i \in S$. The first approach is to derive an recursive formula:

Let $D_n$ denote the number of derrangements Pick the element $k \in S$ such that $k \mapsto 1$. (This is always possible because we're considering permutations). Then, there are two possibilities.

- If $1 \mapsto k$. Then, the permutation $f : S \to S$ must derrange the set $S \setminus \{1, k\}$, which corresponds to the number $D_{n-2}$. Since $k$ can be chosen arbitrarily except 1, we have $(n-1)D_{n-2}$ derrangements of this type.

- If $1 \not\mapsto k$, then 1 has exactly $(n-2)$ choices. For any other elements in $S \setminus \{k\}$, there are exactly 1 forbidden element, so there are $(n-1)D_{n-1}$ derrangements of this type.

Hence, $D_n = (n-1)D_{n-1} + (n-1)D_{n-2}$.

The second approach is "direct computation":

We count the number of permutations with non-trivial fixed point set.