LECTURE I INDETERMINATE FORMS

1. Indeterminate Forms

We note that

$$\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = \text{"}\frac{0}{0}\text{"}$$

where we call $\frac{0}{0}$ an **indeterminate** form, i.e., it cannot be assigned a value (real number) or infinity.

Why can't we determine the value of $\frac{0}{0}$? One may think two sequences of numbers

$$\left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\right\}, \quad \left\{\frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \ldots\right\}$$

Note clearly that each sequence converges to 0 eventually, so the limiting quotient is of the indeterminate form $\frac{0}{0}$. However, if we look at the quotient of each term, we have

$$\frac{1/2}{1/3} = 1.5$$
$$\frac{1/4}{1/9} = 2.25$$
$$\frac{1/8}{1/27} = 3.5$$
$$\vdots$$

that is, the sequence in the numerator relative to that in the denominator is getting bigger and bigger (eventually to ∞). This implies that the sequence in the numerator is going to zero more slowly than that in the denominator, even though a plausible answer to the limiting value of the ratio is $\frac{0}{0}$.

One may flip the roles of the two sequences, and find that the limiting value of the quotient is getting smaller and smaller (eventually to 0), i.e.,

$$\frac{1/3}{1/2} = 0.67...$$
$$\frac{1/9}{1/4} = 0.44...$$
$$\frac{1/27}{1/8} = 0.296...$$
$$\vdots$$

even though the limiting value is still plausibly $\frac{0}{0}$. With these two examples, we arrive at a completely absurd conclusion

$$\infty \approx \frac{0}{0} \approx 0.$$

This is pure garbage. Therefore, we cannot assign $\frac{0}{0}$ a concrete value or ∞, deeming it **indeterminate**.

Before we talk about indeterminate forms, we first list a few suspects that are actually "determined", though they may not seem so.

Example. (Non-examples of indeterminate forms)

(1)
$$\frac{0}{1} = 0$$
Consider the example
$$\lim_{x \to 0} \frac{x}{1+x}.$$

(2)
$$\frac{1}{0}$$
Consider the example
$$\lim_{x \to 0^+} \frac{1}{x} = +\infty, \quad \lim_{x \to 0^-} \frac{1}{x} = -\infty.$$
In general, $\frac{1}{0}$ is either infinity depending on which direction the asymptote is reached.

(3)
$$\frac{1}{\infty}$$
Consider the example
$$\lim_{x \to \infty} \frac{1}{x} = 0$$

(4)
$$a^\infty, \quad |a| \neq 1$$
If $|a| < 1$, then consider the example
$$\lim_{x \to \infty} \left(\frac{1}{2}\right)^\infty = 0.$$
If $|a| > 1$, then consider the example
$$\lim_{x \to \infty} 2^\infty = \infty.$$

(5) $\infty + \infty = \infty$.

Other indeterminate forms include $\frac{\infty}{\infty}$, $0 \times \infty$, $\infty - \infty$, 0^0, 1^∞ and ∞^0. Now, the question is, are these forms of equivalent or different "indeterminacies"? If equivalent, can one convert between each of these and the one indeterminate form $\frac{0}{0}$?

Example. Given the knowledge that $\frac{0}{0}$ is an indeterminate form, show that $\frac{\infty}{\infty}$ is also one.

- Shorter argument: consider
$$\frac{\infty}{\infty} = \frac{1/0}{1/0} = \frac{0}{0}.$$
- Longer but **better** argument:

  To show that $\frac{\infty}{\infty}$ is an indeterminate form, consider the two functions f and g such that $\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x)$, which then puts us in position,
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \text{"}\frac{\infty}{\infty}\text{"}.$$
  Via a simple algebraic manipulation, we note that
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{1/g(x)}{1/f(x)}.$$
  Now, since $\lim_{x\to a} f(x) = \infty = \lim_{x\to a} g(x)$, we have
$$\lim_{x\to a} \frac{1}{f(x)} = 0 = \lim_{x\to a} \frac{1}{g(x)}.$$
  Therefore,
$$\text{"}\frac{\infty}{\infty}\text{"} \overset{\mathbf{A}}{=} \lim_{x\to a} \frac{f(x)}{g(x)} \overset{\mathbf{B}}{=} \lim_{x\to a} \frac{1/g(x)}{1/f(x)} \overset{\mathbf{C}}{=} \text{"}\frac{0}{0}\text{"}$$
  where we draw an abstract equivalence between $\frac{\infty}{\infty}$ and $\frac{0}{0}$.

  The logical progression is as follows:

  At step **A**, we just used the definition of indeterminate forms to put us in position of the analysis – start with the item we want to analyze. At step **B**, we rewrite the original expression via an algebraic manipulation. At step **C**, we find that the original expression $\lim_{x\to a} \frac{f(x)}{g(x)}$, represented by the indeterminate form "$\frac{\infty}{\infty}$", is co-represented by "$\frac{0}{0}$". Since "$\frac{0}{0}$" is an indeterminate form, the representative "$\frac{\infty}{\infty}$" must also be an indeterminate form.

Exercise. (Bonus HW1) Argue that $0 \times \infty, \infty - \infty, 0^0, 1^\infty$ and ∞^0 are all indeterminate forms using a similar approach as above. One may utilize continuous functions such as $\log(x)$ and a^x, i.e.,
$$\lim_{x\to a} \log(f(x)) = \log\left(\lim_{x\to a} f(x)\right), \quad \lim_{x\to a} b^{f(x)} = b^{\lim_{x\to a} f(x)},$$
namely, the limit of a continuous function as $x \to a$ is equation to the function value at $x = a$. In the second case, $b \neq 0$.

LECTURE II L'HOPITAL'S RULE

1. L'Hopital's Rule: a Perspective from Linearization

L'Hopital's Rule: Statement. Given differentiable functions f and g such that $f(a) = g(a) = 0$, we have

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

If $f'(a) = g'(a) = 0$, and f and g are twice differentiable functions, then we repeat the application of this rule, and write

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

This can go on forever, as long as the derivatives have value 0 at $x = a$.

Justification using Linearization (optional). To determine the limit $\lim_{x\to a} \frac{f(x)}{g(x)}$, then direct evaluation yields an indeterminate form, we draw an inspiration from linearization.

Suppose f and g are both differentiable functions at $x = a$ and the derivatives f' and g' are continuous functions, with $g'(a) \neq 0$. Further suppose that $f(a) = g(a) = 0$, such that a direct evaluation of the limit $\lim_{x\to a} \frac{f(x)}{g(x)}$ yields the indeterminate form "$\frac{0}{0}$".

By differentiability, f and g each admit a linearization,

$$f(x) = \cancelto{0}{f(a)} + f'(a)(x-a) + err_f(x-a) = f'(a)(x-a) + err_f(x-a);$$
$$g(x) = \cancelto{0}{g(a)} + g'(a)(x-a) + err_g(x-a) = g'(a)(x-a) + err_g(x-a).$$

Suppose that the error term err_f and err_g goes to 0 as $x \to a$ (meaning that the approximation is exact at $x = a$). Thus,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(a)(x-a) + err_f(x-a)}{g'(a)(x-a) + err_g(x-a)} = \frac{f'(a)}{g'(a)}.$$

Since f and g are differentiable at $x = a$, their derivatives at $x = a$ are **continuous functions**, which means

$$f'(a) = \lim_{x\to a} f'(x),$$
$$0 \neq g'(a) = \lim_{x\to a} g'(x),$$

and thus so is their quotient (this is where the condition $g'(a) \neq 0$ comes in handy or otherwise you cannot perform the division)

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

We now arrive at the celebrated L'Hopital's Rule,

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

This equality is saying that the limiting value of a quotient of two functions is equal to that of the derivatives of the two functions. If both f and g are hitting zero as $x\to a$, the limiting ratio of function values depends on **the rate** at which $f(x)$ and $g(x)$ approach $f(a)$ and $g(a)$, respectively.

When we study Taylor series in Chapter 10, we will provide a proof of L'Hopital's rule, in general.

2. Examples

Now, we can soar. Always remember, you check first, after **every** limit evaluation, whether the expression gives an indeterminate form. If it doesn't, then you CANNOT use L'Hopital's Rule.

Example. (Sine identity)

$$\lim_{x\to 0}\frac{\sin x}{x}=\text{"}\frac{0}{0}\text{"}.$$

Thus, we apply L'Hopital's Rule,

$$\lim_{x\to 0}\frac{\sin x}{x}\overset{L'H}{=}\lim_{x\to 0}\frac{\cos x}{1}=\frac{1}{1}=1.$$

Remark. (optional) One can think of the linearization of $\sin x$ near $x=0$.

$$\sin x=\sin 0+\cos(0)(x-0)+error=x+error$$

where the error goes to zero as $x\to a$. Thus

$$\lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{x+error}{x}=1.$$

Example.

$$\lim_{x\to 0}\frac{\sin x}{x^2}=\text{"}\frac{0}{0}\text{"}.$$

Thus, we apply L'Hopital's Rule,

$$\lim_{x\to 0}\frac{\sin x}{x^2}\overset{L'H}{=}\lim_{x\to 0}\frac{\cos x}{2x}=\begin{cases}\infty & x\to 0^+\\ -\infty & x\to 0^-\end{cases}$$

This implies that $\lim_{x\to 0}\frac{\sin x}{x^2}$ does not exist.

Example.

$$\lim_{x\to 1}\frac{\arctan(x)-\frac{\pi}{4}}{x-1}=\text{"}\frac{0}{0}\text{"}.$$

Thus,

$$\lim_{x\to 1}\frac{\arctan(x)-\frac{\pi}{4}}{x-1}\overset{L'H}{=}\lim_{x\to 1}\frac{\frac{1}{1+x^2}}{1}=\frac{1}{2}.$$

Example.

$$\lim_{x\to 0}\frac{1-\cos x}{x+x^3}\overset{L'H,\frac{0}{0}}{=}\lim_{x\to 0}\frac{-\sin x}{1+3x^2}=0.$$

Note that there is no need to keep applying L'Hopital's rule. You only apply it when you have an indeterminate form.

Example. (Exponential growth defeats linear growth)

$$\lim_{x\to\infty}\frac{e^x}{x} = \text{"}\frac{\infty}{\infty}\text{"}$$

Thus, we apply L'Hopital's Rule,

$$\lim_{x\to\infty}\frac{e^x}{x} \overset{L'H}{=} \lim_{x\to\infty}\frac{e^x}{1} = \lim_{x\to\infty} e^x = \infty.$$

This implies that the exponential function blows up to infinity faster than the line $y = x$.

Example. (L'Hopital or not? That's never a question)

$$\lim_{x\to\infty}\frac{x^7+6x^2+3}{7x^7+3x^5+4x}.$$

Clearly, the top and bottom go to ∞ individually, which puts us in an indeterminate form of "$\frac{\infty}{\infty}$".

$$\lim_{x\to\infty}\frac{x^7+6x^2+3}{7x^7+3x^5+4x} \overset{L'H}{=} \lim_{x\to\infty}\frac{7x^6+12x}{49x^6+15x^4+4}.$$

It seems we still have "$\frac{\infty}{\infty}$". Alright, more L'Hopital's rule.

$$\lim_{x\to\infty}\frac{7x^6+12x}{49x^6+15x^4+4} \overset{L'H}{=} \lim_{x\to\infty}\frac{42x^5+12}{49\cdot 5x^5+60x^3}.$$

Arrrrghhh! Still "$\frac{\infty}{\infty}$"! More L'Hopital's rule? I think you probably see my point.

2.1. **Using Continuity.** When a function g is continuous, it has the following beautiful property such that

$$\lim_{x\to a} g(x) = g\left(\lim_{x\to a} x\right) = g(a).$$

This implies that we can replace x with a function that approaches the value a, i.e.,

$$\lim_{x\to a} f(g(x)) = f\left(\lim_{x\to a} g(x)\right) = f(g(a))$$

where f is also a continuous function (in particular at $x = g(a)$).

Here is a nice result using continuity:

If $\lim_{x\to a}\ln(g(x)) = L$, then

$$\lim_{x\to a} g(x) = \lim_{x\to a} e^{\ln(g(x))} = e^{\lim_{x\to a}\ln(g(x))} = e^L$$

where a may be either finite or infinite.

This result is because the function e^x is continuous (everywhere in its domain), and therefore, the limiting value of e^x as $x \to L$ is equal to the function value e^L.

Example. (Powers) $\lim_{x\to\infty} x^{\frac{1}{x}}$.

Here, direct evaluation yields the indeterminate form ∞^0, which is NOT the form to which we can apply L'Hopital's Rule (that requires a quotient). So, we make use the natural log. Note that

$$x^{\frac{1}{x}} = e^{\ln\left(x^{\frac{1}{x}}\right)} = e^{\frac{\ln x}{x}}.$$

Therefore,

$$\lim_{x\to\infty} x^{\frac{1}{x}} = \lim_{x\to\infty} e^{\ln\left(x^{\frac{1}{x}}\right)} = \lim_{x\to\infty} e^{\frac{\ln x}{x}} \overset{\text{continuity of } e^x}{=} e^{\lim_{x\to\infty}\frac{\ln x}{x}}.$$

We now focus on the limit

$$\lim_{x\to\infty}\frac{\ln x}{x}\overset{"\frac{\infty}{\infty}",L'H}{=}\lim_{x\to\infty}\frac{1/x}{1}=0.$$

Therefore,

$$\lim_{x\to\infty}x^{\frac{1}{x}}=e^{\lim_{x\to\infty}\frac{\ln x}{x}}=e^0=1.$$

2.2. **Indeterminate forms $0\times\infty$, $\infty-\infty$, 1^∞.**

Example. (**Create your own quotient**) In some situations, the function form in a limit evaluation is not necessarily a quotient, but direct evaluation still yields one of the indeterminate forms, such as $0\times\infty$. One needs to massage the expression into a quotient before applying L'Hopital's rule.

$$\lim_{x\to 0^+}\sqrt{x}\ln x="0\times(-\infty)".$$

So, we **make** a quotient out of this.

$$\lim_{x\to 0^+}\sqrt{x}\ln x=\lim_{x\to 0^+}\frac{\ln x}{1/\sqrt{x}}=-"\frac{\infty}{\infty}".$$

where $f(x)=\ln x$ and $g(x)=\frac{1}{\sqrt{x}}$. L'Hopital's rule then applies to f and g.

$$\lim_{x\to 0^+}\sqrt{x}\ln x=\lim_{x\to 0^+}\frac{\ln x}{1/\sqrt{x}}\overset{L'H}{=}\lim_{x\to 0^+}\frac{1/x}{-\frac{1}{2x^{3/2}}}=-\lim_{x\to 0^+}2\sqrt{x}=0$$

Example. (**Don't create your own quotient**)

$$\lim_{x\to\infty}\sqrt{x^2-x}-x="\infty-\infty".$$

Be mindful that not all problems **have** to use L'Hopital's Rule. The indeterminate form "$\infty-\infty$" is also not in position to apply L'Hopital's Rule.

We rationalize the expression by multiplying and dividing the conjugate

$$\begin{aligned}\lim_{x\to\infty}\frac{\left(\sqrt{x^2-x}-x\right)\left(\sqrt{x^2-x}+x\right)}{\sqrt{x^2-x}+x}&=\lim_{x\to\infty}\frac{x^2-x-x^2}{\sqrt{x^2-2}+x}\\&=-\lim_{x\to\infty}\frac{x}{\sqrt{x^2-2}+x}\\&=-\lim_{x\to\infty}\frac{\frac{x}{x}}{\frac{\sqrt{x^2-2}}{x}+\frac{x}{x}}\\&=-\lim_{x\to\infty}\frac{1}{\sqrt{\frac{x^2-2}{x^2}}+1}\\&=-\lim_{x\to\infty}\frac{1}{\sqrt{1-\frac{2}{x^2}}+1}\\&=-\frac{1}{2}.\end{aligned}$$

Example. (1^∞)

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x="1^\infty".$$

Again, we need to **make** a quotient. Note that

$$\left(1+\frac{1}{x}\right)^x = e^{\ln\left(1+\frac{1}{x}\right)^x} = e^{x\ln\left(1+\frac{1}{x}\right)} = e^{\frac{\ln\left(1+\frac{1}{x}\right)}{\frac{1}{x}}}.$$

Thus,

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

Now we focus on $\lim_{x\to\infty}\frac{\ln\left(1+\frac{1}{x}\right)}{\frac{1}{x}}$, which yields "$\frac{0}{0}$". So, we apply L'Hopital's rule,

$$\lim_{x\to\infty}\frac{\ln\left(1+\frac{1}{x}\right)}{\frac{1}{x}} = \lim_{y\to 0}\frac{\ln(1+y)}{y} \overset{L'H}{=} \lim_{y\to 0}\frac{\frac{1}{1+y}}{1} = 1.$$

Using this knowledge, we go back to the original expression and find

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

LECTURE III IMPROPER INTEGRALS OF TYPE I

Consider the innocent area: from $x = 1$ to ∞ of the function $\frac{1}{x}$, represented by the integral

$$\int_1^\infty \frac{1}{x} dx.$$

Do we get a finite area? Does $\frac{1}{x} \to 0$ as $x \to \infty$ help us here at all? Once we obtain an answer, can we make the same conclusion for

$$\int_1^\infty \frac{1}{x^2} dx$$

even though $\frac{1}{x^2} \to 0$ as $x \to \infty$ also?

We first develop some understanding of the case where the integrand is continuous over the interval of interest. One point is to extend a similar formula such as

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

to the infinite case, $\int_a^\infty f(x)\, dx$?

Definition. (Improper Integrals of Type I)

(1) If f is continous on $[a, \infty)$, then

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

(2) If f is continous on $[-\infty, b)$, then

$$\int_{-\infty}^b f(x)\, dx = \lim_{a \to -\infty} \int_a^b f(x)\, dx.$$

(3) (Piecing part 1 and 2 together) If f is continuous on $(-\infty, \infty)$, then

$$\int_{-\infty}^\infty f(x)\, dx = \int_{-\infty}^c f(x)\, dx + \int_c^\infty f(x)\, dx$$

where c is any real number.

In each case, if the limit exists and is finite, we say that the improper integral **converges** and that the limit is the **value** of the improper integral. If the limit fails to exist, the improper integral **diverges**.

One way to read statements such as "if ... then ..." is to know that the condition under "if" **guarantees** the conclusion after "then". However, even if the conditions fail to be met, the same conclusion may hold in specific situations – it's just no longer a guarantee.

Question: why is f being continuous important here?

HW: find an explicit example such that f is discontinuous on $[a, \infty)$ such that

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

no longer holds. You are free to choose any a, and the location of the discontinuity.

Example. Investigate the convergence of the improper integral

$$\int_1^\infty \frac{\ln x}{x^2} dx$$

First, note that the only bad point is $x = 0$ where both $\ln x$ and $\frac{1}{x^2}$ have an vertical asymptote. But our interval is away from it. So $\frac{\ln(x)}{x^2}$ is a continuous function on the interval of interest. It is an improper integral of Type I, in particular, of item 1 in the definition.

$$\int_1^\infty \frac{\ln x}{x^2} dx = \lim_{b\to\infty} \int_1^b \frac{\ln x}{x^2} dx.$$

It is thus enough to study

$$\begin{aligned}
A(b) &= \int_1^b \frac{\ln x}{x^2} dx \\
&\overset{IBP}{=} \ln(x)\left(-\frac{1}{x}\right)|_1^b - \int_1^b \left(-\frac{1}{x}\right)\left(\frac{1}{x}\right) dx \\
&= -\frac{\ln b}{b} - \left[\frac{1}{x}\right]|_1^b \\
&= -\frac{\ln b}{b} - \frac{1}{b} + 1.
\end{aligned}$$

Then, we take a limit.

$$\begin{aligned}
\int_1^\infty \frac{\ln x}{x^2} dx &= \lim_{b\to\infty} A(b) \\
&= \lim_{b\to\infty}\left(-\frac{\ln b}{b} - \frac{1}{b} + 1\right) \\
&= \lim_{b\to\infty} -\frac{\ln b}{b} - 0 + 1 \\
&\overset{L'H,\text{"}\frac{\infty}{\infty}\text{"}}{=} -\lim_{b\to\infty} \frac{1/b}{1} + 1 \\
&= 1
\end{aligned}$$

We say that the improper integral $\int_1^\infty \frac{\ln x}{x^2} dx$ converges and it has a finite area of 1.

Example.

$$\int_{-\infty}^\infty \frac{dx}{1+x^2}.$$

No vertical asymptotes. Behaviour at infinity is good –

$$\lim_{x\to\pm\infty} \frac{1}{1+x^2} = 0.$$

In fact, $\frac{1}{1+x^2}$ in continuous everywhere. Thus, we are in definition part 3.

Now, we note that $\frac{1}{1+x^2}$ is an even function (i.e. $f(-x) = f(x)$), which implies that an integral of it about a symmetric interval about $x = 0$ is just two times the integral on one side.

$$\int_{-\infty}^\infty \frac{dx}{1+x^2} = 2\int_0^\infty \frac{dx}{1+x^2}$$

which we treat as an improper integral.

$$\begin{aligned}
\int_0^\infty \frac{dx}{1+x^2} &= \lim_{b\to\infty} \int_0^b \frac{dx}{1+x^2} \\
&= \lim_{b\to\infty} \tan^{-1}(x) \,\Big|_0^b \\
&= \lim_{b\to\infty} \left[\tan^{-1}(b) - \cancelto{0}{\tan^{-1}(0)}\right] \\
&= \lim_{b\to\infty} \tan^{-1}(b) \\
&= \frac{\pi}{2}.
\end{aligned}$$

Thus,

$$\int_{-\infty}^\infty \frac{dx}{1+x^2} = 2\int_0^\infty \frac{dx}{1+x^2} = \pi$$

LECTURE IV IMPROPER INTEGRALS OF TYPE II

After understanding improper integrals of Type I, where the function is continuous over an infinite interval, we now study integral of functions with discontinuities. We begin with an innocent integral.

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

By first sight, it seems we can quickly identify the antiderivative. The rest is just evaluating the antiderivative at the endpoints.

$$\int_{-1}^{1} \frac{1}{x^2} dx = \left[-\frac{1}{x} \right]_{-1}^{1} = -\left[1 - (-1)\right] = -2.$$

Now, this finite area may seem plausible, and sometimes, due to its very finiteness, we believe it is the correct answer. Let's think about it for a little. First, note that $\frac{1}{x^2} \geq 0$ for every $x \in \mathbb{R}$, that is, the function **always** stays above the x-axis. But we find that its area over the interval $[-1, 1]$ is negative. This doesn't make any sense. So, what went wrong here?

We overlooked the fact that $\frac{1}{x^2}$ over the interval $[-1, 1]$ contains a vertical asymptote at $x = 0$, a (very bad) type of discontinuity. More generally put, when integrating a function over an interval that contains discontinuities, one must be more careful because the function, and thus its integral, may undergo singular behaviour.

The example above once again informs us that the continuity of the integrand is important, and should become part of the conditions that go in a formula for such integrals.

Definition. Integrals of functions that become infinite at a point within the interval of intergration are **improper integrals of Type II**.

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

$$\int_a^b f(x)\, dx = \lim_{c \to a^+} \int_c^b f(x)\, dx.$$

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

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

(3) If $f(x)$ is continuous at c where $a < c < b$ and continuous on $[a, c) \cup (c, b]$, then

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

In each case, if the limit exists and is finite, we say the improper integral **converges** and that the limit is the value of the improper integral. If the limit does not exist, the integral **diverges**.

Example. Investigate the convergence of

$$\int_0^1 \frac{1}{1-x} dx.$$

This improper integral of Type II part 2 because the integrand becomes discontinuous at the right end point.

$$\begin{aligned}
\int_0^1 \frac{1}{1-x} dx &= \lim_{b\to 1^-} \int_0^b \frac{1}{1-x} dx \\
&= -\lim_{b\to 1^-} \left[\ln|1-x|\right]_0^b \\
&= -\lim_{b\to 1^-} \left[\ln|1-b| - \ln 1\right] \\
&= -\lim_{b\to 1^-} \ln(1-b) \\
&= \infty
\end{aligned}$$

The reason why we lose the absolute value sign at the second last equality is that we know b is approaching 1 from the left, so $b < 1$ which implies $1-b$ is positive; thus $|1-b| = 1-b$ if $1-b > 0$.

The limit is infinite, so the integral diverges.

Example. Investigate the convergence of

$$\int_0^3 \frac{dx}{(x-1)^{2/3}}.$$

Noting that the integrand has a vertical asymptote at $x = 1$, which lies inside the interval, this improper integral is of Type II part 3. We separate the integral into two parts,

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = \int_0^1 \frac{dx}{(x-1)^{2/3}} + \int_1^3 \frac{dx}{(x-1)^{2/3}}$$

where each part now is of Type II part 1 and 2, respectively.

$$\begin{aligned}
\int_0^1 \frac{dx}{(x-1)^{2/3}} &= \lim_{b\to 1^-} \int_0^b \frac{dx}{(x-1)^{2/3}} \\
&= \lim_{b\to 1^-} \left[3(x-1)^{1/3}\right]_0^b \\
&= 3 \lim_{b\to 1^-} \left[(b-1)^{1/3} - (-1)^{1/3}\right] \\
&= 3.
\end{aligned}$$

Then,

$$\begin{aligned}
\int_1^3 \frac{dx}{(x-1)^{2/3}} &= \lim_{a\to 1^+} \int_a^3 \frac{dx}{(x-1)^{2/3}} \\
&= \lim_{a\to 1^+} \left[3(x-1)^{1/3}\right]_a^3 \\
&= 3 \lim_{a\to 1^+} \left[2^{1/3} - (a-1)^{1/3}\right] \\
&= 3 \times 2^{1/3}.
\end{aligned}$$

Altogether,

$$\int_0^3 \frac{dx}{(x-1)^{2/3}} = 3 + 3 \times 2^{1/3}.$$

LECTURE V COMPARISON TESTS FOR IMPROPER INTEGRALS

The method of investigating the convergence of improper integrals does not limit to just integration techniques and limits. Sometimes, antiderivatives are hard to spot. We must then rely on properties the function in the integrand and try to compare one integral to another. Even though we wouldn't be able to determine the exact value if the integral converges, at least we can show that it does converge (resp. diverge).

Direct Comparison Test

Let f and g be continuous functions on $[a, \infty)$ with $0 \leq f(x) \leq g(x)$ for all $x \geq a$. Then

(1) If $\int_a^\infty g(x)\, dx$ converges, then $\int_a^\infty f(x)\, dx$ also converges. (Sane big brother controls younger brother)
(2) If $\int_a^\infty f(x)\, dx$ diverges, then $\int_a^\infty g(x)\, dx$ also diverges. (Insane younger brother makes big brother go crazy)

Example. Does the integral $\int_1^\infty e^{-x^2} dx$ converge?

We note that this improper integral is of Type I part 1. However, even if we perform a limiting procedure, we can't find the antiderivative of e^{-x^2} in closed form. Thus, computing an explicit limit is not hopeful. We resort to answer whether the integral converges or not without determining the exact value.

Now, it is clear that

$$e^{-x^2} \leq e^{-x}, \quad x \geq 1$$

because e^{-x} is a monotone decreasing function (compare e^{-3} and e^{-3^2}). Therefore, by Direct Comparison Test, we have

$$\int_1^\infty e^{-x^2} dx \leq \int_1^\infty e^{-x} dx = \lim_{b\to\infty} \int_1^b e^{-x} dx = \lim_{b\to\infty} \left[-e^{-b} + e^{-1}\right] = \frac{1}{e} \approx 0.36788.$$

Example. (Quick checks)

(1) $\int_1^\infty \frac{\sin^2 x}{x^2} dx$ converges because
$$0 \leq \frac{\sin^2 x}{x^2} \leq \frac{1}{x^2}, \quad x \geq 1$$
and also
$$\int_1^\infty \frac{1}{x^2} dx$$
converges.

(2) $\int_1^\infty \frac{1}{\sqrt{x^2-1}} dx$ diverges because
$$\frac{1}{\sqrt{x^2-1}} \geq \frac{1}{\sqrt{x^2}} = \frac{1}{x}$$
and $\int_1^\infty \frac{1}{x} dx$ diverges.

(3) $\int_0^{\pi/2} \frac{\cos x}{\sqrt{x}} dx$ converges because

$$0 \leq \frac{\cos x}{\sqrt{x}} \leq \frac{1}{\sqrt{x}}, \quad x \in \left[0, \frac{\pi}{2}\right]$$

and

$$\begin{aligned}\int_0^{\pi/2} \frac{1}{\sqrt{x}} &= \lim_{a\to 0^+} \int_a^{\pi/2} \frac{1}{\sqrt{x}} dx \\ &= 2 \lim_{a\to 0^+} \left[\sqrt{x}\right]_a^{\pi/2} \\ &= 2 \lim_{a\to 0^+} \left[\sqrt{\frac{\pi}{2}} - \sqrt{a}\right] \\ &= \sqrt{2\pi} < \infty.\end{aligned}$$

(4) $\int_2^{\infty} \frac{1}{x^3+1} dx$ converges. Here, we make use of p-integral where $p = 2$.

$$f(x) = \frac{1}{x^3+1} \leq \frac{1}{x^3} = g(x)$$

and thus

$$\int_2^{\infty} \frac{1}{x^3+1} dx \leq \int_2^{\infty} \frac{1}{x^3} dx < \infty.$$

Limit Comparison Test

Now, following the last example, suppose it is changed a tiny bit:

$$\int_2^{\infty} \frac{1}{x^3-1} dx.$$

Does this integral still converge? You may feel the same way about it because when x is very large, the 1 is insignificant so the integral should look like $\int_2^{\infty} \frac{1}{x^3} dx$ which converges. However, you can't use the DCT because

$$\frac{1}{x^3-1} \geq \frac{1}{x^3}$$

which is the opposite inequality needed to prove convergence (we need smaller or equal to). Thus, what can we do? Should we give up on the useful information that the function is very similar to $\frac{1}{x^3}$ altogether?

Instead, we resort to something called the **Limit Comparison Test**. This procedure involves in identifying a function $g(x)$ such that

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

where the limit satisfies $0 < L < \infty$. Then, $\int_a^{\infty} f(x)\, dx$ converges **if and only if** $\int_a^{\infty} g(x)\, dx$ converges for some $a > 0$ (and f and g are continuous on $(a, \infty]$).

The intuition is as follows. If the limit is satisfied, it roughly suggests that $f(x) = Lg(x)$ for x very large. The convergence of any integral

$$\int_a^{\infty} f(x)\, dx$$

involves the behaviour of $f(x)$ near infinity; in fact, the convergence of this integral does NOT depend on $\int_a^b f(x)\, dx$ for any b.

Now, the rough estimate such that $f(x) = Lg(x)$ suggests that the integral

$$\int_a^\infty f(x)\,dx \approx L\int_a^\infty g(x)\,dx$$

which implies that the convergece or divergence of either integral controls the other.

Back to the example. We let $f(x) = \frac{1}{x^3-1}$ and choose the function $g(x) = \frac{1}{x^3}$. Clearly,

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty}\frac{\frac{1}{x^3-1}}{\frac{1}{x^3}} = \lim_{x\to\infty}\frac{x^3}{x^3-1} = \lim_{x\to\infty}\frac{1}{1-\frac{1}{x^3}} = 1 > 0.$$

Therefore, by LCT, $\int_2^\infty \frac{1}{x^3-1}dx$ converges.

Algorithm

To study the convergence of

$$\int_a^\infty f(x)\,dx$$

where $f(x) \geq 0$.

(1) Find a function $g(x) \geq 0$ such that its limiting behaviour as $x \to \infty$ is very similar to the function of interest $f(x)$. This g should be simple enough so that you can show its convergence/divergence, e.g., a p-integral, or $\int_a^\infty e^{-cx}dx$ for $c > 0$.

(2) Confirm that

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

where $0 < L < \infty$ (this is verifying your intuition from part (1)).

(3) Then the fate of original integral aligns with that of $\int_a^\infty g(x)\,dx$.

One More Example

Example. Investigate $\int_1^\infty \frac{1-e^{-x}}{x}dx$.

Now, if we were to perform a Direct Comparison, we need an educated guess of whether this integral converges or not. At first sight, the $1/x$ factor seems to imply that the integral does NOT converge. Thus, it is useful to find a function $g(x) \leq \frac{1-e^{-x}}{x}$ where $\int_1^\infty g(x)\,dx$ diverges (acting as the insane younger brother). This is not easy.

Back to the example, we find that $\frac{1-e^{-x}}{x}$ is almost like $\frac{1}{x}$ when x is large. Thus, a good candidate for comparison would be $g(x) = \frac{1}{x}$. Indeed,

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty}\frac{\frac{1-e^{-x}}{x}}{\frac{1}{x}} = \lim_{x\to\infty} 1 - e^{-x} = 1$$

which is nonzero and finite. Thus, the convergence of the integral of f depends on that of g.

LECTURE VI SEQUENCES AND SERIES

Sequences

Definition and Examples. Suppose you have P dollars and plan to deposit it in the bank at an annual interest rate of r (compounded yearly). At the end of year 1, you will receive rP dollars as your interest, totaling your assets to $P + rP = P(1+r)$. After two years, you will have

$$P(1+r) + rP(1+r) = P(1+r)^2$$

and similarly, after n years, you will have $P(1+r)^n$ dollars. Thus, we can form a table

Year	1	2	.	.	n	.	.
Asset value	$P(1+r)$	$P(1+r)^2$	.	.	$P(1+r)^n$	.	.

We call "Year" the **index** of the **sequence** "Asset value", where a **sequence** is an ordered list of numbers.

The formula

$$a_n = P(1+r)^n$$

is called the **rule** of the sequence. Sometimes, we also write

$$\{a_n\} = \{P(1+r)^n\}_{n=1}^{\infty}$$

where the $\{\}$ brackets represent a collection/set of numbers.

Example. Typical sequences:

(1) Even numbers, $a_n = 2n$, for $n = 0, 1, \ldots$.
(2) Odd numbers, $a_n = 2n + 1$, for $n = 0, 1, \ldots$.
(3) Alternating signs, $a_n = (-1)^{n+1}$.

Convergence and Divergence of Sequences. Consider $a_n = \frac{1}{n}$, i.e.,

$$\{a_n\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$$

whose terms approach 0 when n is large.

The sequence

$$\{b_n\} = \left\{1 - \frac{1}{n}\right\}_{n=1}^{\infty} = \left\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\right\}$$

approaches 1 as n increases.

The sequence

$$\{c_n\} = \{\sqrt{n}\}_{n=1}^{\infty} = \left\{1, \sqrt{2}, \sqrt{3}, \sqrt{4}, \ldots\right\}$$

gets larger than any number as n increases.

Lastly, the alternating sequence

$$\{d_n\} = \left\{(-1)^{n+1}\right\}_{n=1}^{\infty} = \{1, -1, 1, -1, 1, -1, \ldots\}$$

does not settle at a single value.

Optional: Rigorous Definition of the Limit of a Sequence

So, what is the proper definition of a convergent and divergent sequence?

If a sequence converges, let its limit be L. The sequence eventually becomes very close to L, but only when the sequence index is large enough.

Definition. (Optional) We say a sequence $\{a_n\}$ converges to the limit L if for every $\epsilon > 0$, there corresponds a positive integer N such that

$$|a_n - L| < \epsilon$$

whenever $n > N$.

Think of ϵ as a tolerance level (an error level you can accept), and N as some kind of a threshold index (that depends on your tolerance). This definition says the following: whichever tolerance you pick, I can go deep enough in the sequence to find a threshold such that any member of the sequence after this threshold satisfies your tolerance.

Remark. Per the previous rephrasing, we must note that $N = N(\epsilon)$, i.e., the threshold depends on the tolerance.

Example. Show that $\lim_{n\to\infty} \frac{1}{n} = 0$.

Proof. You pick a tolerance level ϵ. This number is fixed. The limit $L = 0$. Thus, we want to express the threshold past which the sequence element satisfies

$$|a_n - L| = \left|\frac{1}{n} - 0\right| < \epsilon \implies \frac{1}{n} < \epsilon \implies n > \frac{1}{\epsilon}.$$

This means, if you look at indices larger than $\frac{1}{\epsilon}$, the inequality $\frac{1}{n} < \epsilon$ holds. In other words, pick your threshold $N > \frac{1}{\epsilon}$ (say $N(\epsilon) = \left\lfloor \frac{1}{\epsilon} \right\rfloor + 1$, the smallest integer greater than $\frac{1}{\epsilon}$), then the inequality $\frac{1}{n} < \epsilon$ is satisfied for all $n > N$. □

Convergence means we expect the limit to be a finite value. Divergence is not exactly the converse. It is the case when either the limit is infinite or does not exist (such as oscillations). For example, the sequence

$$a_n = \sqrt{n}$$

diverges, but we say more specifically that it **diverges to positive infinity**. On the other hand, the alternating sequence

$$b_n = (-1)^{n+1}$$

simply diverges (but certainly not to any kind of infinity).

Definition. (Optional) The sequence $\{a_n\}$ **diverges to positive infinity** if for every number M, there is an integer N such that for all $n > N$, we have $a_n > M$. We write

$$\lim_{n\to\infty} a_n = +\infty, \text{or } a_n \to +\infty.$$

Similarly, we say a sequence $\{a_n\}$ **diverges to negative infinity** if for every number M, there is an integer N such that for all $n > N$, we have $a_n < M$. We write

$$\lim_{n\to\infty} a_n = -\infty, \text{or } a_n \to -\infty.$$

Remark. Note again, say, in the case of **divergence to positive infinity**, N is a threshold that depends on M. Tell me a number M, and I can look deep enough of the sequence so that I can find all numbers past the threshold N that beat M (or lose to M in the negative infinity case).

Example. (Optional) Show that $a_n = \sqrt{n}$ diverges to positive infinity.

Proof. So, pick your number M. I am going to find out at which n so that the resulting member of the sequence **begins** to beat M.

$$a_n = \sqrt{n} > M \implies n > M^2.$$

Therefore, as long as I go past the index $N > M^2$, a_n will beat M for any $n > N$ (consider $N = \lfloor M \rfloor + 1$).

If you still don't believe me, then specifically choose M, say $M = 5$. To find the member of the sequence that begins to beat M, I sweep through the terms and find

$$a_{25} = \sqrt{25} = 5 = M.$$

Now, if I look at a_{26}, I will find

$$a_{26} = \sqrt{26} > 5 = M.$$

Therefore, for $M = 5$, I find that when $N = 26$, $a_n > M$ for all $n > N = 26$. This is a special argument since we specified M, while the one above is for general M. □

Operations of Sequences

Taking limits of a sequence follows exactly the same rules as that of a function. Suppose $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$. Then

(1) (Sum/Difference) $\lim_{n\to\infty} (a_n \pm b_n) = \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n = A \pm B$.
(2) (constant multiple) $\lim_{n\to\infty} ka_n = kA$.
(3) (product) $\lim_{n\to\infty} a_n b_n = AB$.
(4) (quotient) $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}$ when $B \neq 0$.

Note that rule no.1 does not imply that the existence of LHS implies the individual limit exists. A simple example would be

$$a_n = n, \quad b_n = -n$$

where both diverge, but

$$a_n + b_n = 0$$

converges.

LECTURE VII

Introduction To Infinite Series

After infinite sequence, which is an infinite ordered list of numbers, we study the sum of such list, named **infinite series**.

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots .$$

Two intrinsic questions arise:

(1) Does the knowledge of the underlying sequence contribute to that of the series, i.e., properties of

$$\{a_1, a_2, a_3, \ldots, a_n, \ldots\}?$$

(2) Does the general knowledge of sequences help us understand series, i.e., convergence of sequences? If so, how do we translate between sequence and series?

The answer to 1 is yes, but it varies between cases. The answer to 2 is a resounding yes. In particular, when studying series, the principal object is the n^{th} **partial sum**, i.e.,

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

Clearly, studying s_n as $n \to \infty$, as a sequence, is equivalent to studying the infinite series.

Example. Consider the infinite series

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

Let's form the table of partial sums.

Partial sum	Value	Suggestive pattern
$s_1 = 1$	1	$2 - 1$
$s_2 = 1 + \frac{1}{2}$	$\frac{3}{2}$	$2 - \frac{1}{2}$
$s_3 = 1 + \frac{1}{2} + \frac{1}{4}$	$\frac{7}{4}$	$2 - \frac{1}{4}$
.	.	.
.	.	.
.	.	.
$s_n = 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{n-1}}$	$\frac{2^n - 1}{2^{n-1}}$	$2 - \frac{1}{2^{n-1}}$

From the pattern, we observe a rule for the n^{th} partial sum,

$$s_n = 2 - \frac{1}{2^{n-1}} \to 2, \quad as\, n \to \infty.$$

We say that the sum of the infinite series

$$1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{n-1}} + \cdots$$

is 2.

Definition. Given a sequence of numbers $\{a_n\}$, an expression of the form

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots$$

is an **infinite series**. The number a_n is the n^{th} term of the series. The sequence $\{s_n\}$ defined by

$$\begin{aligned} s_1 &= a_1 \\ s_2 &= a_1 + a_2 \\ s_3 &= a_1 + a_2 + a_3 \\ &.. \\ &.. \\ &.. \\ s_n &= a_1 + a_2 + a_3 + \cdots + a_n = \sum_{k=1}^{n} a_k \end{aligned}$$

is the **sequence of partial sums** of the series, the number s_n being the n^{th} **partial sum**. If the sequence of partial sums converges to a limit L, we say that the series **converges** and that its **sum** is L. In this case, we also write

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots = \sum_{n=1}^{\infty} a_n = L.$$

If the sequence of partial sums of the series does not converge, we say that the series **diverges**.

n^{th} term Divergence Test

Theorem. *If $\sum_{n=1}^{\infty} a_n$ converges, then $a_n \to 0$.*

The contrapositive statement of this theorem is:

If $\lim_{n\to\infty} a_n$ diverges or fails to exist, or converges to a number other than 0, then $\sum_{n=1}^{\infty} a_n$ diverges.

Thus, it suffices to check the divergence of the sequence to determine the divergence of the series.

Remark. If the sequence converges to 0, it does NOT tell us anything about the convergence or divergence of the series. In other words, do NOT use this divergence test to test for convergence.

Example. All of the following series are **divergent** due to the divergence test:

(1) $\sum_{n=1}^{\infty} n^2$.
(2) $\sum_{n=1}^{\infty} \frac{n+1}{n}$.
(3) $\sum_{n=1}^{\infty} (-1)^{n+1}$.
(4) $\sum_{n=1}^{\infty} \sin(n)$.
(5) $\sum_{n=1}^{\infty} -\frac{n}{2n+5}$.

The Geometric Series

A Motivation Example from Present Value.

Example. At the end of next year, you expect a payment of 100 dollars. Since this amount isn't at your own disposal right now, you can't invest it, i.e., deposit in the bank at an interest rate of I. Even though this amount is indeed "yours", you lost the interest-earning potential, which means it is currently worth less than 100 dollars.

In fact, the present value (PV) of this amount is

$$V_1 = \frac{100}{1+I}$$

where V_1 is to represent the present value for a payment to be received in **1** year (assuming that depositing the money in the bank is your only investment strategy).

Now, imagine certain life insurance policy pays you 100 dollars every year until eternity. Then, the present value of this policy is actually

$$V_1 + V_2 + V_3 + \cdots = \sum_{n=1}^{\infty} V_n = \sum_{n=1}^{\infty} \frac{100}{(1+I)^n},$$

that is, each year's payment has decreasing present value. This is an example of a **geometric series**. Here, we say $\frac{1}{1+I}$ is the **ratio** of the geometric series.

Definition.

$$a + ar + ar^2 + \cdots + ar^{n-1} + \cdots = \sum_{n=1}^{\infty} ar^{n-1}$$

in which a and r are fixed real numbers and $a \neq 0$.

With the standard form here, where we always extract **the first term** a and put the ratio to $n-1$ power, we can rewrite the previous example.

$$\sum_{n=1}^{\infty} \frac{100}{(1+I)^n} = \sum_{n=1}^{\infty} \left(\frac{100}{1+I}\right) \frac{1}{(1+I)^{n-1}}$$

where $a = \frac{100}{1+I}$ and $r = \frac{1}{1+I}$.

Geometric Series Formula. To study any infinite series, we focus on the n^{th} **partial sum**.

$$s_n = a + ar + ar^2 + \cdots + ar^{n-1}.$$

To find a rule for s_n may be completely hopeless, but here is a trick:

$$\begin{aligned} rs_n &= ar + ar^2 + ar^3 + \cdots + ar^n \\ \implies s_n - rs_n &= a - ar^n \\ \implies s_n &= \frac{a(1-r^n)}{1-r}, \quad r \neq 1. \end{aligned}$$

Woohoo! A general formula for s_n. We observe that when r satisfies certain conditions, the limit of s_n may be determined.

(1) If $|r| < 1$, then $r^n \to 0$ (common limits for sequences), and thus

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

which means the geometric series **converges**.

(2) If $|r| \geq 1$, then $r^n \to \infty$, which indicates that the geometric series **diverges**. The case where $r = 1$ is not deduced from the formula above, but by simply looking at the geometric series where

$$a + ar + ar^2 + \cdots = a + a + a + \cdots = \infty.$$

Example. Find the sum of the series where the underlying sequence is

(1) $\{a_n\} = \left\{\frac{1}{9}, \frac{1}{27}, \frac{1}{81}, \ldots\right\}$.

Solution. So,

$$\begin{aligned}\frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \cdots &= \frac{1}{9}\left(1 + \frac{1}{3} + \frac{1}{9} + \cdots\right). \\ &= \sum_{n=1}^{\infty} \frac{1}{9}\left(\frac{1}{3}\right)^{n-1} \\ &= \frac{1/9}{1 - (1/3)} \\ &= \frac{1}{6}.\end{aligned}$$

(2) $\{b_n\} = \left\{5, -\frac{5}{4}, \frac{5}{16}, -\frac{5}{64}, \ldots\right\}$.

Solution. So,

$$\begin{aligned}5 - \frac{5}{4} + \frac{5}{16} - \frac{5}{64} + \cdots &= 5\left(1 - \frac{1}{4} + \frac{1}{16} - \frac{1}{64} + \cdots\right). \\ &= \sum_{n=1}^{\infty} 5\left(-\frac{1}{4}\right)^{n-1} \\ &= \frac{5}{1 - (-1/4)} \\ &= 4\end{aligned}$$

Remark. **The geometric series formula only applies to geometric series, and nothing else. You must identify the first term and the ratio.** Do NOT abuse elsewhere as you simply can't.

Telescoping Series

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

The k^{th} partial sum is

$$\begin{aligned}s_k &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &= 1 - \frac{1}{k+1}.\end{aligned}$$

Note that $s_k \to 1$ as $k \to \infty$, the series converges, and its sum is 1.

Operations of Series

If $\sum_{n=1}^{\infty} a_n = A$ and $\sum_{n=1}^{\infty} b_n = B$, then

(1) (Sum/Difference) $\sum_{n=1}^{\infty} (a_n \pm b_n) = \sum_{n=1}^{\infty} a_n \pm \sum_{n=1}^{\infty} b_n = A \pm B$.

(2) (Constant multiple)

$$\sum_{n=1}^{\infty} ka_n = k \sum_{n=1}^{\infty} a_n = kA.$$

From 1, if one of the two series actually diverges, then the first item always yields a divergent series.

Example. $\sum_{n=1}^{\infty} \frac{3^{n-1}-1}{6^{n-1}}$.

We split into two series.

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{3^{n-1}-1}{6^{n-1}} &= \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{n-1} - \sum_{n=1}^{\infty} \left(\frac{1}{6}\right)^{n-1} \\
&= \frac{1}{1-\frac{1}{2}} - \frac{1}{1-\frac{1}{6}} \\
&= 2 - \frac{6}{5} \\
&= \frac{4}{5}.
\end{aligned}$$

Reindexing

Note that we have always used $n = 1$ and ∞ as the bounds. This may change according to your needs.

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

If you don't believe this, just write both sides out. More generally, to raise the starting value of the index by h units, we have

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

and to lower the starting value of the index by h units, we have

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

Example. We can write the geometric series

$$\sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = 1 + \frac{1}{2} + \frac{1}{4} + \cdots$$

as

$$\boxed{\sum_{n=0}^{\infty} \frac{1}{2^n}}, \quad \sum_{n=5}^{\infty} \frac{1}{2^{n-5}}, \quad \sum_{n=-4}^{\infty} \frac{1}{2^{n+4}}.$$

A trick to see this is: from the original series, look at the starting index, 1, and the starting power $n - 1$; note its sum is n. Thus, if you want to reduce the starting index without affecting the series, you must compensate by adding the same amount to the power. In the first example (boxed), we reduced the starting index by 1, thus we add 1 back to the power $n-1$ which gives n.

LECTURE VIII

A Monotone Theorem for Sequences

Theorem. *If a sequence a_n is monotone and bounded (from above AND below), then it converges.*

This fundamental result about sequences has major contributions to modern iterative numerical methods. However, its proof is also intuitive – if you can squeeze the **infinite** sequence between two numbers, namely, the upper and lower bound, while the sequence itself is either monotone increasing/decreasing, then eventually, it has to find its limit somewhere between the bounds.

We know already that a series can be viewed as the limit of its n^{th} partial sum, which is a sequence. Therefore, the theorem also shines light on the convergence of series.

Corollary. *A series $\sum_{n=1}^{\infty} a_n$ of nonnegative terms converges if and only if its partial sums are bounded from above.*

Now, a series of nonnegative terms is already bounded from below by 0. Furthermore, since you are always adding nonnegative numbers, the partial sum is non-decreasing – monotone increasing. Thus, the theorem applies if the sequence of partial sums is also bounded from above.

The Integral Test

So, from the Corollary, can we find such upper bound for a series of nonnegative terms? Anything we can borrow from all previous calculus knowledge?

Example. Consider the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Every series has a "box" interpretation. Consider boxes of various heights, but a fixed width of 1 unit. Then, the first term of the this series can be viewed as the area of a box with height

$$\frac{1}{1^2}$$

and width 1. Similarly, the second term is the area of a box with height

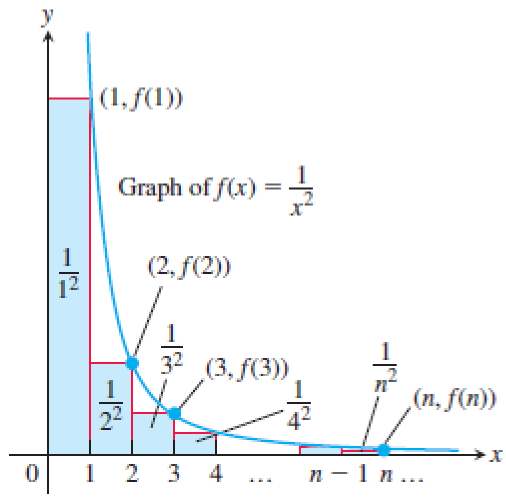
$$\frac{1}{2^2}$$

and width 1. So on and so forth.

So, consider the function $f(x) = \frac{1}{x^2}$. We see that the partial sum of the series is

$$\begin{aligned} s_n &= \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} \\ &= f(1) + f(2) + f(3) + \cdots f(n). \end{aligned}$$

Now, with the help of a sketch of the function f and the corresponding boxes



underneath the curve, we find that

$$
\begin{aligned}
s_n = f(1) + f(2) + f(3) + \cdots f(n) &< f(1) + \int_1^n \frac{1}{x^2} dx \\
&< 1 + \int_1^\infty \frac{1}{x^2} dx \\
&< 1 + 1 \\
&= 2.
\end{aligned}
$$

This is saying that the partial sum $s_n < 2$, that is, bounded from above. Since s_n is also monotone increasing, by the Corollary, we conclude that the series

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

converges.

This technique can be further improved in general to a theorem, called "the **Integral Test**".

Theorem. *Let a_n be a sequence of positive terms. Suppose that $a_n = f(n)$ where f is a continuous, positive and decreasing function for x for all $x \geq N$ (N a positive integer). Then the series $\sum_{n=N}^{\infty} a_n$ and the integral $\int_N^\infty f(x)\, dx$ both converge or both diverge.*

Let's apply it first and then see the proof. The idea is to converge again the underlying sequence into a function representation, and then investigate the improper integral of the function.

Example. Show that the p-series

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

converges if $p > 1$ and diverges if $p \leq 1$.

Solution. If $p > 1$, then $f(x) = 1/x^p$ is a positive decreasing function of x. What's more,

$$\int_1^\infty \frac{1}{x^p} dx = \frac{1}{p-1}$$

(see Lecture III on this). Therefore, the series converges by the Integral Test.

If $p \le 0$, the series diverges by the Divergence Test. How? Note that we can write $\frac{1}{n^p} = n^q$ for $q = -p > 0$. Obviously, $n^q \to \infty$ as $n \to \infty$.

If $0 < p < 1$, then $1 - p > 0$, and

$$\int_1^\infty \frac{1}{x^p} dx = \frac{1}{1-p} \lim_{b\to\infty} \left(b^{1-p} - 1\right) = \infty$$

(again, see Lecture III on this). Thus, the series diverges by the Integral Test.

Lastly, if $p = 1$, then we have the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots$$

which corresponds to the improper integral

$$\int_1^\infty \frac{1}{x} dx,$$

which we know diverges.

Example. The series $\sum_{n=1}^\infty \frac{1}{n^2+1}$ is NOT a p-series, but it converges by the Integral Test. To see this, let the function be

$$f(x) = \frac{1}{x^2+1}.$$

It is positive, continuous, and decreasing for $x \ge 1$ (you MUST verify that the chosen f satisfies these three criteria). Furthermore,

$$\begin{aligned}
\int_1^\infty \frac{1}{x^2+1} dx &= \lim_{b\to\infty} \left[\arctan x\right]_1^b \\
&= \lim_{b\to\infty} \left[\arctan b - \arctan 1\right] \\
&= \frac{\pi}{2} - \frac{\pi}{4} \\
&= \frac{\pi}{4}.
\end{aligned}$$

Thus, by the Integral test, $\sum_{n=1}^\infty \frac{1}{n^2+1}$ converges. However, we don't know what it converges to. The Integral Test does NOT say that $\pi/4$ or any other number is the sum of the series.

Remark. The Integral Test requires a good knowledge of integration techniques.

Example. Determine the convergence or divergence of the series.

(1) $\sum_{n=1}^\infty ne^{-n^2}$. (Technique: substitution)
(2) $\sum_{n=1}^\infty \frac{1}{2^{\ln n}}$. (Technique: substitution)

LECTURE IX COMPARISON TESTS FOR SERIES

Direct Comparison Test for Series.

Theorem. *Let $\sum a_n$ and $\sum b_n$ be two series with $0 \le a_n \le b_n$ for all n. Then*

(1) *If $\sum b_n$ converges, then $\sum a_n$ converges.*
(2) *If $\sum a_n$ diverges, then $\sum b_n$ diverges.*

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

Knowing the Direct Comparison Test, **we can skip the conversion to a function representation, and just compare the underlying sequence to another.** However, we still need to develop an intuition on whether the series converges or not.

The series here almost look like $\sum_{n=2}^{\infty} \frac{1}{n}$ which is a p-series for $p = 1$, which is known to diverge (due to the Integral Test). Thus, we should compare something that diverges. More precisely, we need to find a **smaller sequence** whose series diverges. Indeed

$$\frac{1}{n-1} > \frac{1}{n} \implies \sum_{n=2}^{\infty} \frac{1}{n-1} > \sum_{n=2}^{\infty} \frac{1}{n} = \infty.$$

Example. $\sum_{n=2}^{\infty} \frac{n-1}{n^4+2}$.

The underlying sequence behaves like $\frac{1}{n^3}$ when n is large, so we expect convergence due to p-series test. The goal now is to find a dominating sequence that beats the given one, whose series converges.

$$\frac{n-1}{n^4+2} \overset{\text{numerator made bigger}}{\le} \frac{n}{n^4+2} \overset{\text{denominator made smaller}}{\le} \frac{n}{n^4} = \frac{1}{n^3}.$$

Thus, by Direct Comparison Test for series, we have

$$\sum_{n=2}^{\infty} \frac{n-1}{n^4+2} \le \sum_{n=2}^{\infty} \frac{1}{n^3} < \infty.$$

Now, the inequalities here seem straightforward because the numerator and denominator give what we want easily.

Limit Comparison Test for Series. What if we have

$$\sum_{n=2}^{\infty} \frac{n+1}{n^4-2}$$

which also behaves like $\frac{1}{n^3}$ for large n, and thus implies convergence. How do we use the Direct Comparison Test now?

Solution. Note we still need a larger sequence whose series is convergent:

$$\frac{n+1}{n^4-2} \overset{\text{numerator made bigger}}{\le} \frac{n+n}{n^4-2} = \frac{2n}{n^4-2} \overset{\text{denominator made smaller}}{\le} \frac{2n}{n^4 - \frac{1}{2}n^4} = \frac{4}{n^3}.$$

The first inequality holds if $n \geq 1$ since we replaced the 1 with n. The second inequality holds iff

$$2 \leq \frac{1}{2}n^4 \iff n^4 \geq 4 \iff n \geq \sqrt{2} \approx 1.414....$$

But the series starts at $n = 2$ already, so we are safe. Thus, by Direct Comparison,

$$\sum_{n=2}^{\infty} \frac{n+1}{n^4-2} \leq \sum_{n=2}^{\infty} \frac{4}{n^3} = 4\sum_{n=2}^{\infty} \frac{1}{n^3} < \infty.$$

Remark. **This is so much pain. Can we use something simpler? Yes. We retain the valuable information that the underlying sequence looks like $\frac{1}{n^3}$ when n is large.**

Theorem. *Suppose that $a_n > 0$ and $b_n > 0$ for all $n \geq N$ (N an integer, problem dependent).*

(1) *If $\lim_{n\to\infty} \frac{a_n}{b_n} = c > 0$, then $\sum a_n$ and $\sum b_n$ both converge or both diverge.*
(2) *If $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ and $\sum b_n$ converges, then $\sum a_n$ also converges.*
(3) *If $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$ and $\sum b_n$ diverges, then $\sum a_n$ diverges.*

Few comments about how to understand each result:

(1) Similar to LCT for improper integrals.
(2) If $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, then b_n can be viewed as the **dominating sequence "big brother"**, and by DCT, the **"smaller brother"** $\sum a_n$ is sane.
(3) If $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$, then b_n can be viewed the **subordinating sequence "little brother"**, and by DCT, it is driving the **"bigger brother"** $\sum a_n$ insane.

The last example of DCT is quite some subtle work. **LCT helps us circumvent it.** Consider again the series

$$\sum_{n=2}^{\infty} \frac{n+1}{n^4-2}.$$

We already suspected that the underlying sequence looks like $\frac{1}{n^3}$, why don't we compare to it? Define

$$a_n = \frac{n+1}{n^4-2}, \quad b_n = \frac{1}{n^3}$$

and then we find

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\frac{n+1}{n^4-2}}{\frac{1}{n^3}} = \lim_{n\to\infty} \frac{n^4+n^3}{n^4-2} = 1.$$

We fall in the first case. Now, we already know by p-series that $\sum b_n = \sum \frac{1}{n^3}$ converges, and thus by Limit Comparison Test, $\sum a_n$ also converges.

Example. Does $\sum_{n=1}^{\infty} \frac{\ln n}{n^{3/2}}$ converge?

This relies on a bit of experience. We know that exponential functions beat any polynomials. Now, here is another fact – natural logarithm grows slower than any polynomials. You can prove it: for $a > 0$

$$\lim_{x\to\infty} \frac{\ln x}{x^a} \overset{L'H, \frac{\infty}{\infty}}{=} \lim_{x\to\infty} \frac{x^{-1}}{ax^{a-1}} = \frac{1}{a} \lim_{x\to\infty} \frac{1}{x^a} = 0.$$

A similar statement can be said for sequences
$$\lim_{n\to\infty} \frac{\ln n}{n^a} = 0, \quad a > 0.$$
Here, even though the above limit does NOT automatically tell us convergence of the series, its essence helps us select a sequence to compare. Let
$$a_n = \frac{\ln n}{n^{3/2}}, \quad b_n = \frac{1}{n^{5/4}}.$$
Then,
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\ln n}{n^{1/4}} = 0.$$
Further note that $\sum b_n$ is a p-series with $p = 5/4 > 1$ which converges. Together with the limit, these conditions put us in **Case 2** in LCT for series, and suggest that
$$\sum_{n=1}^{\infty} \frac{\ln n}{n^{3/2}} = \sum_{n=1}^{\infty} a_n \text{ converges.}$$

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

Solution. Compare to $b_n = \frac{1}{n}$.
$$\lim_{n\to\infty} \frac{\sin\left(\frac{1}{n}\right)}{\frac{1}{n}} = \lim_{m\to 0} \frac{\sin(m)}{m} = \lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$
Therefore, we are in Case 1 of LCT for series. We know already that $\sum_{n=1}^{\infty} b_n$ diverges. Thus, $\sum_{n=1}^{\infty} \sin\left(\frac{1}{n}\right)$ diverges.

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

Solution. One wishes to use the direct comparison
$$a_n = \frac{1}{3^{n-1}-1} < \frac{1}{3^{n-1}}$$
but this unfortunately is untrue.

However, we may use the RHS in the LCT for series. Let $b_n = \frac{1}{3^{n-1}}$. Then,
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\frac{1}{3^{n-1}-1}}{\frac{1}{3^{n-1}}} = \lim_{n\to\infty} \frac{3^{n-1}}{3^{n-1}-1} = \lim_{n\to\infty} \frac{1}{1-\left(\frac{1}{3}\right)^{n-1}} = 1.$$
This puts us in **Case 1** of LCT for series. We know already that
$$\sum_{n=2}^{\infty} b_n = \sum_{n=2}^{\infty} \frac{1}{3^{n-1}} = \frac{1/3}{1-1/3} = \frac{1}{2} < \infty, \quad \text{converges}$$
Therefore, $\sum_{n=2}^{\infty} a_n = \sum_{n=2}^{\infty} \frac{1}{3^{n-1}-1}$ also converges.

Example. $\sum_{n=2}^{\infty} \frac{(\ln n)^2}{n}$.

Solution. Intuition: $\ln n$ grows very slowly, so slow that it is almost a constant compared to any other polynomial growth. Thus we expect the series to diverge like $\sum \frac{1}{n}$.

Let $a_n = \frac{(\ln n)^2}{n}$ and $b_n = \frac{1}{n}$. Then we check
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\frac{(\ln n)^2}{n}}{\frac{1}{n}} = \lim_{n\to\infty} (\ln n)^2 = \infty.$$

We also know that $\sum \frac{1}{n}$ diverges. Thus, we are in **Case 3**. This implies that $\sum_{n=2}^{\infty} \frac{(\ln n)^2}{n}$ diverges.

LECTURE XI ABSOLUTE CONVERGENCE

Absolute Convergence

We know the series $\sum \frac{1}{n^2}$ converges by the Integral Test. What if now we have

$$1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \cdots$$

that is, the alternating series of inverse squares,

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

Does this series converge?

Definition. A series $\sum a_n$ converges absolutely (is absolutely convergent) if the corresponding series of absolute values, $\sum |a_n|$ converges.

Theorem. *(The Absolute Convergence Test) If $\sum_{n=1}^{\infty} |a_n|$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.*

Proof. The proof relies on the Direct Comparison Test.

For each n, we know that

$$-|a_n| \leq a_n \leq |a_n|.$$

Adding $|a_n|$ to all terms, we have

$$0 \leq a_n + |a_n| \leq 2|a_n|.$$

By the Direct Comparison Test,

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

converges since $\sum_{n=1}^{\infty} 2|a_n|$ converges.

Thus,

$$\begin{aligned}\sum_{n=1}^{\infty} a_n &= \sum_{n=1}^{\infty} (a_n + |a_n| - |a_n|)\\ &= \left(\sum_{n=1}^{\infty} a_n + |a_n|\right) - \left(\sum_{n=1}^{\infty} |a_n|\right)\end{aligned}$$

which is convergent because the two individual series are convergent. □

Then, to prove that $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2}$ converges or not, we check its absolute series,

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

as it is a p-series with $p = 2$. By Absolute Convergence Test, the original series $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n^2}$ converges.

Any kind of alternating series that may make you scratch your head about its convergence now can be checked by this test.

Example. $\sum_{n=1}^{\infty} \frac{\sin n}{n^2}$ is an alternating series because $\sin n$ can be positive or negative. However, taking an absolute series, we have

$$\sum_{n=1}^{\infty} \left| \frac{\sin n}{n^2} \right| \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

By the Absolute Convergence Test, $\sum_{n=1}^{\infty} \frac{\sin n}{n^2}$ also converges.

The Ratio Test

Let's study the convergence of $\sum_{n=0}^{\infty} \frac{(2n)!}{n!n!}$ (this shows up in probability theory and statistics) . Compare to, uh, ... we give up, temporarily.

Now, let's build some intuition from geometric series. Consider the geometric series $\sum ar^n$. We see that

$$\frac{a_{n+1}}{a_n} = \frac{ar^{n+1}}{ar^n} = r$$

which is true for every n, that is, the underlying sequence has this constant ratio r. We know that if $r < 1$, the series $\sum ar^n$ converges and if $r \geq 1$, we have divergence. So, it seems that checking the ratio of the underlying sequence may offer some information about the convergence of the series.

Theorem. *Let $\sum a_n$ be any series and suppose that*

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

Then

1. *The series converges absolutely if $\rho < 1$;*
2. *The series diverges if $\rho > 1$ or ρ is infinite;*
3. *The test is inconclusive if $\rho = 1$.*

Proof. Only a sketch of proof here. If the limit holds, then

$$a_{n+1} \approx \rho a_n$$

for every $n \geq N$ for some large N. Thus,

$$\sum_{n=N}^{\infty} a_{n+1} = \sum_{n=N}^{\infty} \rho a_n = \sum_{n=N}^{\infty} \rho^2 a_{n-1} = ... = \sum_{n=N}^{\infty} \rho^n a_1 = a_1 \sum_{n=N}^{\infty} \rho^n$$

where the last series is a geometric series. Certainly, the convergence criteria is $\rho < 1$ (and diverges if $\rho > 1$).

For $\rho = 1$, we just need to come up with two examples, one diverges and one converges but both yield a limiting ratio of 1. Consider $a_n = \frac{1}{n}$ and $b_n = \frac{1}{n^2}$.

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{1/\left(n+1\right)}{1/n} = \frac{n}{n+1} \to 1,$$

while

$$\left| \frac{b_{n+1}}{b_n} \right| = \frac{1/\left(n+1\right)^2}{1/n^2} = \frac{n^2}{n^2+2n+1} \to 1.$$

Yet $\sum a_n$ diverges but $\sum b_n$ converges. $\rho = 1$ is indeed inconclusive. $\square$

Example. $\sum_{n=0}^{\infty} \frac{(2n)!}{n!n!}$.

Solution. So, let $a_n = \frac{(2n)!}{n!n!}$. We have

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{(2(n+1))!}{(n+1)!(n+1)!}}{\frac{(2n)!}{n!n!}} \\ &= \frac{(2(n+1))!}{(2n)!} \frac{n!n!}{(n+1)!(n+1)!} \\ &= \frac{(2n+2)!}{(2n)!} \frac{1}{(n+1)(n+1)} \\ &= \frac{(2n+2)(2n+1)}{(n+1)(n+1)} \\ &\to 4. \end{aligned}$$

This implies that the series diverges since the limiting ratio is 4.

Example. $\sum_{n=0}^{\infty} \frac{2^n+5}{3^n}$.

Solution. We can solve this problem in three ways.

(1) Define $a_n = \frac{2^n+5}{3^n}$. Then, checking the ratio, we have

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{2^{n+1}+5}{3^{n+1}}}{\frac{2^n+5}{3^n}} \\ &= \frac{2^{n+1}+5}{2^n+5} \frac{3^n}{3^{n+1}} \\ &= \frac{1}{3} \frac{2^{n+1}+5}{2^n+5} \\ &= \frac{1}{3} \frac{2+5/2^n}{1+5/2^n} \\ &\to \frac{2}{3} < 1. \end{aligned}$$

By the Ratio Test, the series converges absolutely.

(2) We can prove this using DCT by comparing to $\sum_{n=0}^{\infty} \frac{2^n+2^n}{3^n} = \sum_{n=0}^{\infty} 2\left(\frac{2}{3}\right)^n < \infty$.

(3) We can do direct computation, and find

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{2^n+5}{3^n} &= \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n + 5\sum_{n=0}^{\infty} \left(\frac{1}{3}\right)^n \\ &= \frac{1}{1-\frac{2}{3}} + \frac{5}{1-\frac{1}{3}} \\ &= 3 + \frac{15}{2} \\ &= \frac{21}{2} < \infty. \end{aligned}$$

Remark. The Ratio Test should be applied to series who terms contain exponentials or factorials because expressions of these forms may cancel/divide nicely.

Optional: The Root Test

Another test derived from the geometric series convergence result is the Root Test.

Theorem. *Let $\sum a_n$ be any series and suppose that*

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

Then

(1) *The series converges absolutely if $\rho < 1$;*
(2) *The series diverges if $\rho > 1$ or ρ is infinite;*
(3) *The test is inconclusive if $\rho = 1$.*

Proof. Only a sketch here. Suppose the limit holds. Then,

$$\sqrt[n]{|a_n|} \approx \rho$$

for some very large n. This implies that

$$|a_n| \approx \rho^n \implies \sum |a_n| \approx \sum \rho^n$$

which is a geometric series. The convergence criteria of geometric series give us the same breakdown of ρ.

For $\rho = 1$, the same examples from the Ratio Test will also work – both yield a limiting root of 1 but one diverges and another converges. $\square$

Example. $\sum_{n=1}^{\infty} \frac{7}{(2n+5)^n}$.

Solution. Define $a_n = \frac{7}{(2n+5)^n}$. The root test yields

$$\begin{aligned}\lim_{n\to\infty} \sqrt[n]{|a_n|} &= \lim_{n\to\infty} \sqrt[n]{\frac{7}{(2n+5)^n}} \\ &= \lim_{n\to\infty} \frac{7^{1/n}}{2n+5} \\ &= 0 < 1\end{aligned}$$

which implies that the series converges absolutely. But you should think about why the limit is 0.

Example. $\sum_{n=1}^{\infty} \frac{e^n}{n^e}$.

Solution. Define $a_n = \frac{e^n}{n^e}$. Then,

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

which implies that the series diverges. But you should think about why the limit yields e.

Remark. The commonly occurring limits will come in handy here quite often for the Ratio Test or the Root Test.

LECTURE XI ALTERNATING SERIES AND POWER SERIES

Alternating Series

We know from p-series that the harmonic series ($p = 1$) is divergent. So, how about the alternating divergent series?

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

If you apply the Ratio Test to this,

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

which is inconclusive.

Theorem. *(The Alternating Series Test)*
The series

$$\sum_{n=1}^{\infty} (-1)^{n+1} u_n = u_1 - u_2 + u_3 - u_4 + \cdots$$

converges if the following conditions are satisfied:

(1) *The u_n's are all positive.*
(2) *The u_n's are eventually nonincreasing: $u_n \geq u_{n+1}$ for all $n \geq N$, for some integer N.*
(3) $u_n \to 0$.

Example. For the alternating harmonic series,

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

we identify that $u_n = \frac{1}{n} > 0$, u_n is decreasing and $u_n \to 0$.

Remark. The alternating series test would be useful when first the series is clearly alternating, while the Ratio Test yields an inconclusion.

Like the alternating harmonic series, there are many series out there that converges but not absolutely. We call this type of series "conditionally convergent".

Definition. A series that is convergent but not absolutely convergent is called **conditionally convergent**.
In series notation, we say a series $\sum a_n$ is **conditionally convergent** if $\sum |a_n|$ does not converge, but $\sum a_n$ does.

Remark. Since the definition of conditional convergence requires that you know the series does NOT converge absolutely, YOU MUST CHECK ABSOLUTE CONVERGENCE **FIRST** (by ACT or the Ratio Test).

(1) If ACT or the Ratio Test yields convergence, then the series converges absolutely, and we don't have to do more.

(2) **Only if ACT or the Ratio Test yields divergence**, then we may apply AST to the series if it is alternating, and **possibly** establish conditional convergence.

Example. Study whether the following series converges absolutely, conditionally or diverges.

(1) $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{\cos n}{n^2}$.
Solution: Absolute Convergence Test first.
$$\sum_{n=1}^{\infty} \frac{|\cos n|}{n^2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$
Thus, the original series converges absolutely.

(2) $\sum_{n=1}^{\infty} (-1)^n \ln\left(1 + \frac{1}{n}\right)$.
Solution: Absolute Convergence Test first. We study $\sum_{n=1}^{\infty} \ln\left(1 + \frac{1}{n}\right)$. I will first show the method that works, and then discuss other possibilities. We perform LCT using $b_n = 1/n$, and try to prove divergence. The reason to believe in divergence is that $a_n = \ln(1+1/n)$ does not decay fast enough.
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \frac{\ln\left(1+\frac{1}{n}\right)}{1/n} \overset{m=\frac{1}{n}}{=} \lim_{m\to 0} \frac{\ln(1+m)}{m} = \lim_{x\to 0} \frac{\ln(1+x)}{x} \overset{L'H}{=} \lim_{x\to 0} \frac{\frac{1}{1+x}}{1} = 1.$$
This puts us in Case 1 of LCT for series. $\sum a_n$ and $\sum b_n$ have the same fate, but we already know $\sum b_n$ diverges. Thus, $\sum a_n$ diverges, and the original series $\sum_{n=1}^{\infty} (-1)^n \ln\left(1 + \frac{1}{n}\right)$ does NOT converge absolutely.

————*Other approaches:*————

You may do the Ratio Test on the series of absolute values. But you find an inconclusive result:
$$\lim_{n\to\infty} \left| \frac{\ln\left(1+\frac{1}{n+1}\right)}{\ln\left(1+\frac{1}{n}\right)} \right| = \lim_{x\to\infty} \frac{\ln\left(1+\frac{1}{x+1}\right)}{\ln\left(1+\frac{1}{x}\right)} \overset{L'H}{=} \lim_{x\to\infty} \frac{\left(\frac{x+1}{x+2}\right)\cdot\left(-\frac{1}{(x+1)^2}\right)}{\left(\frac{x}{x+1}\right)\cdot\left(-\frac{1}{x^2}\right)} = \lim_{x\to\infty} \frac{x}{x+2} = 1.$$
DCT may work but you need to do a lot of work to find a divergent series smaller than the one at hand. Please don't do the Integral Test. It is applicable, but good luck finding the antiderivative of $\ln(1 + 1/x)$.

————*End of other approaches.*————

After not obtaining absolute convergence, we study whether the original series converges.

(a) $\ln\left(1 + \frac{1}{n}\right) > 0$ because $\ln(x) > 0, x \geq 1$
(b) $\ln(n)$ is a monotone increasing function in n, we see that $\ln\left(1+\frac{1}{n}\right)$ is then monotone decreasing in n
(c) $\ln\left(1 + \frac{1}{n}\right) \to 0$ as $n \to \infty$.

Thus, by AST, the original series converges. Altogether, the original series converges conditionally.

Summary of tests

(1) The n^{th}-term test for **Divergence**: unless $a_n \to 0$, the series diverges.
(2) **Geometric series**: $\sum ar^n$ converges if $|r| < 1$; otherwise it diverges.
(3) p-**series**: $\sum 1/n^p$ converges if $p > 1$ and diverges otherwise.
(4) **Series with nonnegative terms**:
  (a) The Integral Test;

   (b) Direct Comparison Test;
   (c) Limit Comparison Test.
(5) **Series with some negative terms**
   (a) The Ratio Test;
   (b) The Root Test;
   (c) Absolute Convergence Test.
(6) **Conditionally convergent?**
   (a) Check for Absolute Convergence First using ACT or the Ratio Test.
      (i) If through, conclude with absolute convergence and stop.
   (b) If not, then check if the series is alternating – if so, then apply the **Alternating Series Test**. If not, game over (for now), i.e., just say, "can't solve using current methods".

LECTURE XII POWER SERIES REPRESENTATION OF FUNCTIONS

Motivation of Power Series

The two main subjects in Calculus are Differentiation and Integration. We have learned their rules. However, it is not always computationally realistic to calculate quantities such as $\ln(23)$, or $\sin(10)$, or even the naive e^2. In fact, modern computers can only approximate these numbers up to a certain precision, and make error. Then, not to mention the integration of $\ln\left(x^2\right)$, $\sin\left(x^3\right)$ or e^{x^2}, all of which we have no answer to at the moment, with elementary means.

However, there is one particular group of functions we really enjoy differentiating or integrating – the polynomials. The rules are simple – the power rule! It cannot get better than this. Thus, it would be a crazy but brilliant idea to see if we can approximate functions using polynomials, that is,

$$f(x) \approx c_0 + c_1 (x-a) + c_2 (x-a)^2 + \cdots + c_n (x-a)^n$$

about certain point $x = a$. Or, maybe, with even more imagination, let's use an **infinite** polynomial, to represent f exactly.

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

But ugh, infinite series? Don't we worry about its convergence, since that's the topic we have been dealing with? What if the series diverges, so we get $f(x) =$ complete garbage. And now, this infinite series has x in it. What are the protocols to ensure that the equality above really is equality? Does the equality hold for every x in the domain of f or just a subset?

Power Series

A **power series about** $x = 0$ is a series of the form

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

A **power series about** $x = a$ is a series of the form

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

in which the **center** a and the **coefficients** $c_0, c_1, c_2, \ldots, c_n, \ldots$ are constants.

It is also customary to consider a power series as a function of x, namely,

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

from which we observe that f is infinitely differentiable (see it as an infinite-degree polynomial). Our principal question is:

For what x does the power series converge?

Of course, the answer should also depend on c_n. For x very large, we need c_n to control how fast x^n gets such that the series still converges.

Example 1. By the geometric series formula, we know exactly for what x does the following power series converges,

$$1 + x + x^2 + x^3 + \cdots + x^n + \cdots .$$

This is a geometric series with first term 1 and a ratio x. The convergence for geometric series posits that if $|x| < 1$, the series converges and converges to number

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + \cdots , \quad -1 < x < 1.$$

One may then consider the function

$$f(x) = \frac{1}{1-x}$$

and its exact power series representation

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

which is only valid for $|x| < 1$.

Here, we also say that the interval $(-1, 1)$ is the **interval of convergence** for the power series $\sum_{n=0}^{\infty} x^n$.

Example 2. Now, a slightly more complicated geometric series.

$$1 - \frac{1}{2}(x-2) + \frac{1}{4}(x-2)^2 + \cdots + \left(-\frac{1}{2}\right)^n (x-2)^n + \cdots = \sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n (x-2)^n .$$

This geometric series has first term 1 and a ratio of $-\frac{x-2}{2}$. Therefore, the convergence criteria of geometric series requires that

$$\left|-\frac{x-2}{2}\right| < 1 \iff 0 < x < 4.$$

On this interval, the series converges to the function

$$\frac{1}{1 - \left(-\frac{x-2}{2}\right)} = \frac{2}{x}.$$

In other words, the function $f(x) = \frac{2}{x}$ has a power series representation

$$\frac{2}{x} = \sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n (x-2)^n , \quad 0 < x < 4.$$

Since we are blessed with not only geometric series, but also its logical extension – the Ratio Test, we can find the **interval of convergence** for other series that are more involved.

Example 3. $\sum_{n=0}^{\infty} (-1)^{n-1} \frac{x^n}{n}$.

We compute the limiting ratio here while keeping x unknown. We know that the series converges absolutely if the limiting ratio is less than 1. So, let $a_n = (-1)^{n-1} \frac{x^n}{n}$, and thus

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\frac{x^{n+1}}{n+1}}{\frac{x^n}{n}}\right| = |x| \left|\frac{n}{n+1}\right| \to |x| .$$

This implies that if $|x| < 1$, the series converges by the Ratio Test.

Furthermore, if $|x| > 1$, the series diverges.

Note that at $x = 1$, we have the alternating harmonic series, which converges by the Alternating Series Test.

At $x = -1$, we have the negative harmonic series

$$-1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \cdots$$

which obviously diverges.

To conclude, the power series converges for $-1 < x \leq 1$ (note that we include $x = 1$).

Example 4. $\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$.

We perform the same set of analysis as above. Define $a_n = \frac{x^n}{n!}$. We compute the limiting ratio

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}}\right| = |x| \frac{n!}{(n+1)!} = \frac{|x|}{n+1} \to 0.$$

Note that this limit goes to 0 for **any** x. This implies then that the series converges absolutely for any x, namely, the interval of convergence is $(-\infty, \infty)$.

Example 5. $\sum_{n=0}^{\infty} n!x^n$.

Well, this series is already looking bad because the coefficients is blowing up very very quickly. To confirm, we perform the Ratio Test. Define $a_n = n!x^n$.

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

unless $x = 0$ a priori. This implies that the series diverges except $x = 0$.

Procedures in general:

(1) Use the Ratio Test (or the Root Test) to find the largest open interval where the series converges absolutely (just like the examples above),

$$|x - a| < R$$

or

$$a - R < x < a + R$$

Remark. $a = 2$ in Example 2 (look at the center of the interval $0 < x < 4$). In all other examples, $a = 0$. This a **is** the center of the series as introduced, but not all series are written in the convenient form as $\sum_{n=0}^{\infty} \left(-\frac{1}{2}\right)^n (x-2)^n$, so one must check carefully.

R is called the **radius of convergence** of the power series, which may be infinite (Example 5) or finite but nonzero (Example 1-4), or 0 (Example 6).

(2) If R is finite, test for convergence or divergence at each endpoint of the interval. Use a Comparison Test, the Integral Test, or the Alternating Series Test.
(3) If R is finite, the series diverges for $|x - a| > R$ because the n^{th} term does not approach zero for those values of x.

LECTURE XIII TAYLOR SERIES AND MACLAURIN SERIES

Motivation of Taylor Series

We now put our knowledge of series to use. The motivation of this section stems from linearization of a differentiable function $f(x)$ about a point $x = a$, namely,

$$f(x) \approx L(x) = f(a) + f'(a)(x-a).$$

We understand that the line is only exact at $x = a$, matching function value AND slope, and acts as an approximation of $f(x)$ only near $x = a$.

Can we improve this approximation? We obviously make some error with a linear approximation. So, how about more terms by increasing the power? Let's do a quadratization of f, that is, we say that f is locally approximated by a quadratic function,

$$Q(x) = c_0 + c_1(x-a) + c_2(x-a)^2.$$

So, can we find the coefficients c_0, c_1 and c_2 such that $Q(x)$ and $f(x)$ have matching value, first and second derivative, which is better than just a line? So, we force that

$$\begin{aligned} Q(a) &= f(a), \\ Q'(a) &= f'(a), \\ Q''(a) &= f''(a). \end{aligned}$$

Spelling out this system of equations, we have

$$\begin{aligned} Q(x) = c_0 + c_1(x-a) + c_2(x-a)^2 &\implies Q(a) = c_0 = f(a) \\ Q'(x) = c_1 + 2c_2(x-a) &\implies Q'(a) = c_1 = f'(a) \\ Q''(x) = 2c_2 &\implies Q''(a) = 2c_2 = f''(a) \implies c_2 = \frac{f''(a)}{2}. \end{aligned}$$

Thus,

$$Q(x) = f(a) + f'(a)(x-a) + f''(a)(x-a)^2.$$

Taylor Series

From the study of power series, we found that some functions can be represented by power series. But these functions all sort of have a geometric series flavor, e.g.

$$f(x) = \frac{1}{1-x}, \frac{1}{1+x}, etc.$$

Can a general function be represented by a power series?

We propose

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

for I some interval of convergence. These coefficients are to be determined based on what we know about f.

Via Term-by-Term differentiation, we have

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots + nc_n(x-a)^{n-1} + \cdots, \quad x \in I.$$
$$f''(x) = 2c_2 + 2\cdot 3c_3(x-a) + \cdots + n(n-1)c_n(x-a)^{n-2} + \cdots, \quad x \in I.$$
$$f'''(x) = 2\cdot 3c_3 + 2\cdot 3\cdot 4c_4(x-a) + \cdots + n(n-1)(n-2)c_n(x-a)^{n-3} + \cdots, \quad x \in I.$$

with the n^{th} derivative being

$$f^{(n)}(x) = n!a_n(x-a)^n + \text{higher order terms with } (x-a) \text{ as a factor.}$$

Now, evaluating at $x = a$, we find that

$$f(a) = c_0, \quad f'(a) = c_1, \quad f''(a) = 2c_2, \ldots, f^{(n)}(a) = n!a_n, \ldots.$$

Therefore, we find that the coefficients satisfy

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

and the power series representation of an infinitely differentiable function f is

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots, \quad x \in I$$

where $a \in I$. We say this power series is the **Taylor series generated by f at $x = a$**. If $a = 0$, then the series is also called the **Maclaurin series generated by f**, i.e.

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots, \quad x \in I.$$

The Most Important Taylor Series

Example.

$$f(x) = e^x.$$

This is easier than you think because the derivative of e^x at all orders is itself. Here, let's find the Maclaurin series of e^x.

$$f(0) = f'(0) = f''(0) = \ldots = f^{(n)}(0) = \ldots = 1.$$

Therefore,

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots$$
$$\iff \boxed{e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}}$$

Question: for what x does this power series representation hold for e^x?
Answer: do the Ratio Test. Let $a_n = \frac{x^n}{n!}$. Then

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}}\right| = \left|\frac{x}{n+1}\right| \to 0, \quad \text{for any } x$$

which then implies that the interval of convergence of this series is $(-\infty, \infty)$ (or an infinite radius of convergence). Note that this coincides with the study of Example 4 from Lecture 12 (MATH1023A_L12_PS.pdf).

Example.

$$f(x) = \sin x.$$

The job is to recognize a pattern in the derivative. Here, let's take $a = 0$, so we are looking for the Maclaurin series of $\sin x$.

$$\begin{aligned}
n &= 0 \quad f(0) = \sin 0 = 0\\
n &= 1 \quad f'(0) = \cos 0 = 1\\
n &= 2 \quad f''(0) = -\sin 0 = 0\\
n &= 3 \quad f'''(0) = -\cos 0 = -1
\end{aligned}$$

Or in general, we find that the derivatives of even order are always 0, while those of order order are always alternating between 1 and -1. More precisely,

$$\begin{aligned}
\sin x &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots\\
&= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots
\end{aligned}$$

In sigma notation, we put

$$\boxed{\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}}.$$

From the Maclaurin series of $\sin(x)$, we get that of $\cos(x)$ via Term-by-Term differentiation, namely,

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

From the Maclaurin series of e^x, we get that of e^{-x} for free. Just replace x with $-x$,

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \cdots = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^n}{n!}.$$

However, we can do FAR more than this. In fact, we also get the Maclaurin series for $\sin\left(x^2\right)$, or e^{-x^2}, or in general, any function composition $f(g(x))$. For example,

$$\sin\left(x^2\right) = \left(x^2\right) - \frac{\left(x^2\right)^3}{3!} + \frac{\left(x^2\right)^5}{5!} - \frac{\left(x^2\right)^7}{7!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2(2n+1)}}{(2n+1)!}$$

or

$$\begin{aligned}
e^{-x^2} &= 1 + \left(-x^2\right) + \frac{\left(-x^2\right)^2}{2!} + \frac{\left(-x^2\right)^3}{3!} + \frac{\left(-x^2\right)^4}{4!} + \cdots\\
&= 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \frac{x^8}{4!} + \cdots\\
&= \sum_{n=0}^{\infty} (-1)^{n+1} \frac{x^{2n}}{n!}.
\end{aligned}$$

Applications of Power and Taylor Series

Example. (Sum of the alternating harmonic series)
It suffices to identify the function f which has series representation

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

because the alternating harmonic series equals $f(1)$.
Taking a derivative kills n in the bottom. Thus, via Term-by-Term differentiation, we have

$$f'(x) = 1 - x + x^2 - x^3 + \cdots = \frac{1}{1+x}, \quad |x| < 1.$$

By direct integration, we find that

$$f(x) = \ln(1+x)$$

which implies that

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = f(1) = \ln(1+1) = \ln 2,$$

surprising, huh?

Example. Express $\int \sin\left(x^2\right) dx$ as a power series.
So, we found this last class!

$$\begin{aligned} \sin\left(x^2\right) &= \left(x^2\right) - \frac{\left(x^2\right)^3}{3!} + \frac{\left(x^2\right)^5}{5!} - \frac{\left(x^2\right)^7}{7!} + \cdots \\ &= x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+2}}{(2n+1)!} \end{aligned}$$

Therefore, we can integrate the power series term-by-term,

$$\begin{aligned} \int \sin\left(x^2\right) dx &= \int \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+2}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \int x^{4n+2} dx \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \frac{x^{4n+3}}{4n+3}. \end{aligned}$$

Example. (The Most Incredible Identity: Euler's Identity).
First, a few recall of important Maclaurin series

$$\begin{aligned} \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \\ \cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \end{aligned}$$

and

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

(you may observe already the similarity between them).

Now, magic happens. We make use a complex number $a + bi$ where a and b are real numbers and $i = \sqrt{-1}$. Note that

$$i^2 = -1, \quad i^3 = i^2 i = -i, \quad i^4 = i^2 i^2 = 1, \quad i^5 = i^4 i = i, \ldots.$$

Now, suppose we compute the Maclaurin series of the quantity

$$\begin{aligned}
e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} \cdots \\
&= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^6}{6!} + \cdots \\
&= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right) \\
&= \cos\theta + i\sin\theta.
\end{aligned}$$

The identity

$$e^{i\theta} = \cos\theta + i\sin\theta$$

is called the **Euler's identity**. For complex numbers $a + bi$, we find

$$e^{a+bi} = e^a e^{bi} = e^a \left(\cos(b) + i\sin(b)\right).$$

In particular, if $\theta = \pi$, we find the most intriguing identity

$$e^{i\pi} = -1 \iff e^{i\pi} + 1 = 0,$$

that is, an equation that combines five of the most important constants in mathematics.

LECTURE XVI PARAMETRIZATIONS OF PLANE CURVES

So far, we have learned about curves only via the graph of a 1D function, namely, the ordered pair $(x, y) = (x, f(x))$ where

$$y = f(x)$$

via some function f.

However, not all curves can be represented by a function, since a function must pass the vertical line test. Simple shapes such as a full circle easily fails the test. So, what can we do about these curves? Function representations are nice to have since we know so much about them. We need a small adjustment.

Definition. If x and y are given as functions,

$$x = f(t), \quad y = g(t)$$

over an interval I of t-values, then the set of points $(x, y) = (f(t), g(t))$ defined by these equations is a parametric curve. The equations are **parametric equations** for the curve.

This definition implies that we can describe a curve now using **two** single-variable functions, rather than one.

Example 1. Sketch the curve defined by the parametric equations

$$x = t^2, \quad y = t + 1, \quad -\infty < t < \infty.$$

Going back to basics, we set up a table for the function values and inputs.

TABLE 11.2 Values of $x = t^2$ and $y = t + 1$ for selected values of t.

t	x	y
-3	9	-2
-2	4	-1
-1	1	0
0	0	1
1	1	2
2	4	3
3	9	4

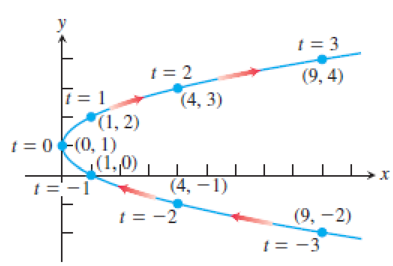


FIGURE 11.3 The curve given by the parametric equations $x = t^2$ and $y = t + 1$ (Example 2).

Now, you may note that we can still find a relationship between x and y. First, we find that $t = \pm\sqrt{x}$, which means

$$y = \pm\sqrt{x} + 1.$$

Indeed, this is describing **two** pieces of curves, rather than 1, though the upper one does look like the function $y = \sqrt{x} + 1$.

Parametrization of Circles

Example 2. The equation associated with a circle is

$$x^2 + y^2 = r^2$$

where r is the radius. So, how can we parametrize $x = f(t)$ and $y = g(t)$ such that the expression above still holds.

Consider

$$x = r\cos t, \quad y = r\sin t, \quad 0 \le t \le 2\pi.$$

Then, trig identity carries through, and we see the formula stands. This parametrization describes a motion that begins at the point $(r, 0)$ (evaluating both quantities at $t = 0$), traverses the circle $x^2 + y^2 = r^2$ **once counterclockwise**, and returns to $(r, 0)$ (evaluating at $t = 2\pi$).

Remark. What motion would it describe if $0 \le t \le \pi$? It's still on the same circle, but only counterclockwise 180 degrees.

Remark. What motion would it describe if $0 \le t \le 4\pi$? Still on the same circle, but counterclockwise **two times** around.

Remark. Is this the **only** parametrization of a circle? No, in fact, to describe the same circle, one may do

$$x = r\cos(2t), \quad y = r\sin(2t)$$

because $x^2 + y^2 = r^2$ is still the same circle. However, the range of t-values will describe the **motion**, in addition to the curve the particle travels on. If $0 \le t \le \pi$, then we have reproduced the same **motion** as the example because the argument of the trigonometric functions are now $2t$.

However, if $0 \le t \le 2\pi$, note that the argument goes from 0 to 4π, which indicates that you have gone around counterclockwise twice around the circle.

Parametrizaiton of a Line

Example. Find a parametrization for the line through the point (a, b) having slope m.

We know already that the equation of this line in Cartesian coordinates is

$$y - b = m(x - a).$$

So, define

$$t = x - a \implies x = a + t, \quad y = b + mt, \quad -\infty < t < \infty.$$

This is a valid parametrization, but not unique. There are various other parametrizations.

Example. Find a parametrization for the line that starts at $(0, 3)$ and ends at $(4, 0)$.

One way to perform this parametrization is via **a convex combination** of the two points.

$$x = 0 \cdot (1 - t) + 4 \cdot t = 4t$$
$$y = 3 \cdot (1 - t) + 0 \cdot t = 3(1 - t)$$

where $0 \le t \le 1$. Clearly, when $t = 0$, we are at the point $(x, y) = (0, 3)$ and when $t = 1$, we are at $(4, 0)$.

In general, to find the parametrization between two points (x_1, y_1) and (x_2, y_2), do

$$x = x_1 (1 - t) + x_2 t,$$
$$y = y_2 (1 - t) + y_2 t,$$

where $0 \leq t \leq 1$.

LECTURE XVII CALCULUS WITH PARAMETRIC CURVES

Now that we have seen a few examples of parametrized curves with equations

$$x = f(t), \quad y = g(t),$$

a few natural questions come about. Ultimately, we want to know more properties of the curves. Formerly, curves we know were described by functions $y = h(x)$. We are curious of the smoothness of the curves, by studying its continuity and then various orders of derivatives. So, instead of being given a function $y = h(x)$, we are given the parametric equations. Can we still find the slope of the tangent line at any given point (x, y)?

Tangents and Areas

A parametrized curve $x = f(t)$ and $y = g(t)$ is **differentiable** at t if f and g are differentiable at t. We make use of the chain rule

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

given that all three derivatives exist. If $dx/dt \neq 0$, we can divide it over to obtain the slope of the tangent

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

Furthermore, if the parametric equations define y as a twice-differentiable function of x, we can apply Chain rule again to find

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

Example. Find the tangent to the curve

$$x = \sec t, \quad y = \tan t, \quad -\pi/2 < t < \pi/2,$$

at the point $\left(\sqrt{2}, 1\right)$ that corresponds to $t = \pi/4$.

Solution. It suffices to find dy/dt and dx/dt.

$$\frac{dy}{dt} = \sec t \tan t, \quad \frac{dx}{dt} = \sec^2 t.$$

Now, don't evaluate at $t = \pi/4$ so quickly. We assemble the derivative first and note for simplifications,

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{\sec^2 t}{\sec t \tan t} = \frac{\sec t}{\tan t} = \frac{1/\cos t}{\sin t/\cos t} = \frac{1}{\sin t}.$$

Thus,

$$\frac{dy}{dx}\,\Big|_{(\sqrt{2},1)} = \frac{1}{\sin\left(\frac{\pi}{4}\right)} = \frac{1}{\sqrt{2}/2} = \sqrt{2}.$$

The equation of the tangent line is given by the point-slope form

$$y - 1 = \sqrt{2}\left(x - \sqrt{2}\right) \implies y = \sqrt{2}x - 1.$$

Example. (second derivative) Find d^2y/dx^2 as a function of t if $x = t - t^2$ and $y = t - t^3$.

Solution. So, the goal is to find $\frac{d^2y}{dx^2} = \frac{dy'/dt}{dx/dt}$. We need to express y' as a function of t. To find y', we have

$$y' = \frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{1-3t^2}{1-2t}.$$

Thus,

$$\frac{dy'}{dt} = \frac{d}{dt}\left(\frac{1-3t^2}{1-2t}\right) = \frac{-6t\,(1-2t) - \left(1-3t^2\right)(-2)}{(1-2t)^2} = \frac{-6t+12t^2+2-6t^2}{(1-2t)^2} = \frac{2-6t+6t^2}{(1-2t)^2}.$$

Altogether,

$$\frac{d^2y}{dx^2} = \frac{dy'/dt}{dx/dt} = \frac{\frac{2-6t+6t^2}{(1-2t)^2}}{1-2t} = \frac{2-6t+6t^2}{(1-2t)^3}.$$

Area Enclosed by a Parametric Curve

We can also find areas enclosed under a curve. In Calculus II, we were blessed with a function $y = h(x)$ so a definite integral does the job, i.e.,

$$A = \int_a^b y dx = \int_a^b h(x)\,dx.$$

Here, we are only given $x = f(t)$ and $y = g(t)$ (because they are more powerful in describing curves that don't pass the vertical line test). The general approach then is to re-express dx in terms of dt, namely,

$$dx = \left(\frac{dx}{dt}\right) dt.$$

It's basically a substitution.

$$A = \int_a^b y dx = \int_c^d y\frac{dx}{dt}dt$$

where the interval $[c, d]$ for t corresponds to the interval $[a, b]$ for x.

Example. Show that the area of a circle of radius 2 is 4π, using the parametric equations

$$x = 2\cos t, \quad y = 2\sin t, \quad 0 \le t \le 2\pi.$$

Solution. We just need to find the area of the northeast sector, and multiply by 4. This region corresponds to the curve $y(x)$ for $0 \le x \le 2$. Meanwhile, this also corresponds to the curve $y(t)$ (note the dependence on t) for $0 \le t \le \pi/2$.

Now, the change of bounds via

$$x = 2\cos t$$

is really subtle.

$$\begin{aligned} x = 0 &\iff t = \frac{\pi}{2} \\ x = 2 &\iff t = 0 \end{aligned}$$

Thus,

$$\begin{aligned}
A &= 4\int_0^2 y dx \\
&= 4\int_{\frac{\pi}{2}}^0 y\left(\frac{dx}{dt}\right)dt \\
&= 4\int_{\frac{\pi}{2}}^0 2\sin t\left(-2\sin t\right)dt \\
&\overset{(\dagger)}{=} 16\int_0^{\pi/2} \sin^2 t dt \\
&= 16\int_0^{\pi/2} \frac{1-\cos 2t}{2}dt \\
&= 16\left[\frac{t}{2}-\frac{\sin 2t}{4}\right]_0^{\pi/2} \\
&= 16\left[\frac{\pi}{4}\right] \\
&= 4\pi.
\end{aligned}$$

where at ($\dagger$), we flipped the integrating bound, thus absorbing the negative sign.

Arc Length of a Curve

Recall that from Calculus II, we learn how to compute the arc length where the arc is described by the graph of a function $y = f(x)$, namely,

$$L = \int_a^b \sqrt{1+\left[f'(x)\right]^2}dx.$$

The derivation of this quantity relies on the infinitesimal manifestation of Pythagorus. More precisely, the arc length differential is given by

$$ds = \sqrt{(dx)^2 + (dy)^2}$$

and

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

Now, since parametrization allows us to describe more curves than function alone does, we need a more general formula for arc length. Given parametric equations

$$x = f(t), \quad y = g(t), t \in [a,b],$$

the arc length over the interval $[a,b]$ is

$$\begin{aligned}
L &= \int_a^b \sqrt{(dx)^2+(dy)^2} \\
&= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2}dt \\
&= \int_a^b \sqrt{\left(f'(t)\right)^2+\left(g'(t)\right)^2}dt
\end{aligned}$$

if f and g are differentiable functions of t. For a more detailed derivation, please refer to your textbook.

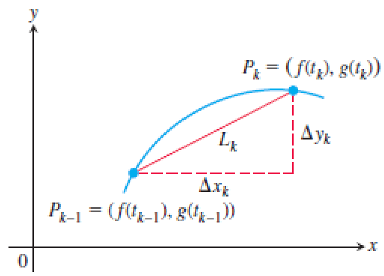


FIGURE 11.17 The arc $P_{k-1}P_k$ is approximated by the straight-line segment shown here, which has length $L_k = \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2}$.

Definition. If a curve C is defined parametrically by $x = f(t)$ and $y = g(t)$, $a \le t \le b$, where f' and g' are continous and not simultaneously zero on $[a, b]$, and C is traversed exactly once as t increases from $t = a$ to $t = b$, then **the length of** C is the definite integral

$$L = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2} dt.$$

Example. Show that the perimeter of a circle of radius r is $2\pi r$ using the parametric equations

$$x = r\cos t, \quad y = r\sin t, \quad 0 \le t \le 2\pi.$$

Solution.

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{[-r\sin t]^2 + [r\cos t]^2} dt \\ &= \int_0^{2\pi} \sqrt{r\left(\sin^2 t + \cos^2 t\right)} dt \\ &= \int_0^{2\pi} r dt \\ &= 2\pi r. \end{aligned}$$

Area of Surfaces of Revolution

Recall again from Calculus II that the area of a surface of revolution relies on the understanding of arc length. Given parametric equations

$$x = f(t), \quad y = g(t), t \in [a, b],$$

(1) **Revolution about the x-axis** ($y \ge 0$):

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

(2) **Revolution about the y-axis** ($x \geq 0$):

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

Example. Find the surface area of a sphere of radius r. (We know it's $4\pi r^2$, confirm it).

Solution. It suffices to consider the surface area of the northern hemisphere. Note that the northern hemisphere can be obtained by revolving the northeast sector of a circle of radius r about the y-axis. This sector has parametric equations

$$x = r\cos t, \quad y = r\sin t, \quad 0 \leq t \leq \frac{\pi}{2}.$$

Thus,

$$\begin{aligned}
S &= \int_0^{\pi/2} 2\pi x \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \\
&= \int_0^{\pi/2} 2\pi r \cos t \sqrt{(-r\sin t)^2 + (r\cos t)^2}\, dt \\
&= 2\pi r \int_0^{\pi/2} \cos t \sqrt{r^2\left(\sin^2 t + \cos^2 t\right)}\, dt \\
&= 2\pi r^2 \int_0^{\pi/2} \cos t\, dt \\
&= 2\pi r^2 \left[\sin t\right]_0^{\pi/2} \\
&= 2\pi r^2.
\end{aligned}$$

Remember that this is the surface area of the northern hemisphere. Multiplying by 2, we obtain the desired result.

LECTURE XVIII POLAR COORDINATES

A Short Note on Area of Surfaces of Revolution

Recall again from Calculus II that the area of a surface of revolution relies on the understanding of arc length. Given parametric equations

$$x = f(t), \quad y = g(t), t \in [a, b],$$

(1) **Revolution about the x-axis** $(y \geq 0)$:

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

(2) **Revolution about the y-axis** $(x \geq 0)$:

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

Example. Find the surface area of a sphere of radius r. (We know it's $4\pi r^2$, confirm it).

Solution. It suffices to consider the surface area of the northern hemisphere. Note that the northern hemisphere can be obtained by revolving the northeast sector of a circle of radius r about the y-axis. This sector has parametric equations

$$x = r\cos t, \quad y = r\sin t, \quad 0 \leq t \leq \frac{\pi}{2}.$$

Thus,

$$\begin{aligned}
S &= \int_0^{\pi/2} 2\pi x \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \\
&= \int_0^{\pi/2} 2\pi r \cos t \sqrt{(-r\sin t)^2 + (r\cos t)^2} dt \\
&= 2\pi r \int_0^{\pi/2} \cos t \sqrt{r^2\left(\sin^2 t + \cos^2 t\right)} dt \\
&= 2\pi r^2 \int_0^{\pi/2} \cos t dt \\
&= 2\pi r^2 \left[\sin t\right]_0^{\pi/2} \\
&= 2\pi r^2.
\end{aligned}$$

Remember that this is the surface area of the northern hemisphere. Multiplying by 2, we obtain the desired result.

Polar Coordinates

We have seen **one set** of parametric equations for a circle of radius r,

$$x = r\cos t, \quad y = r\sin t, \quad 0 \leq t \leq 2\pi.$$

Here, r is the **constant** radius, or, more precisely, the distance of any point on the circle to the **origin**.

Can we describe curves that have this "circular" feature? Maybe an ellipse, maybe a heart, or maybe something even more complicated such as the cardioid, or a three-leaved rose? Clearly, the distance from any point on the curve to the origin is no longer constant – in fact, this distance depends on where you are on the curve. Let's now devise a special parametrization to characterize this relationship.

Definition. Fix an **origin** O (called the **pole**) and an **initial ray** from O (usually chosen as the x-axis). Each point on the plane (with Cartesian coordinate (x, y)) can be located by assigning to it a **polar coordinate pair** (r, θ), in which r gives the **directed distance** from O to P and θ gives the **directed angle** from the initial ray to ray OP. We label this point

$$P(r, \theta).$$

Now, here comes the fun. A Cartesian coordinate (x, y) is unique to every point in space – namely, each point corresponds to exactly **one** Cartesian coordinate. However, the angle θ associated with a given point is not unique, and therefore, each point has **infinitely many** pairs of polar coordinates.

$$P\left(2, \frac{\pi}{6}\right)$$

represents the point where we rotate the initial ray **counterclockwise** by $\pi/6$ and then move **forward** 2 units. This exactly point can also be obtained by rotating **clockwise** $-11\pi/6$ units and move forward 2 units, thus

$$P\left(2, \frac{\pi}{6}\right) = P\left(2, -\frac{11\pi}{6}\right).$$

Another example would be $P\left(2, \frac{7\pi}{6}\right)$, which is turning the ray $7\pi/6$ units counterclockwise, and move forward 2 units. However, this very point can also be obtained by just turning $\pi/6$ radians counterclockwise from the initial ray and going **backward** 2 units. Thus,

$$P\left(2, \frac{7\pi}{6}\right) = P\left(-2, \frac{\pi}{6}\right).$$

More generally, given $P(r, \theta)$, the corresponding coordinate pairs of P are

$$(r, \theta + 2n\pi), \quad n = 0, \pm 1, \pm 2, \ldots$$
$$(-r, \pi - \theta + 2n\pi), \quad n = 0, \pm 1, \pm 2, \ldots$$

Example. Graph the following sets of points whose polar coordinates satisfy the following conditions

(1) $1 \leq r \leq 2$ and $0 \leq \theta \leq \pi/2$.
(2) $-3 \leq r \leq 2$ and $\theta = \pi/4$.
(3) $\frac{2\pi}{3} \leq \theta \leq \frac{5\pi}{6}$ (with no restriction on r)

The equations relating polar and Cartesian Coordinates is

$$x = r\cos\theta, \quad y = r\sin\theta, \quad r^2 = x^2 + y^2, \quad \tan\theta = \frac{y}{x}.$$

This can be viewed as a "double substitution". In the following examples, some curves are more simply expressed with polar coordinates, while others are not.

Example.

$$\begin{aligned} r\cos\theta = 2 &\iff x = 2 \\ r^2\cos\theta\sin\theta = 4 &\iff xy = 4 \\ r = 1 - \cos\theta &\iff x^4 + y^4 + 2x^2y^2 + 2x^3 + 2xy^2 - y^2 = 0 \end{aligned}$$

These are the contrived examples from your textbook. Here are a few more practical ones.

Example. Find a polar equation for the circle

$$x^2 + (y-3)^2 = 9.$$

This is a circle with radius 3, centered at $(0, 3)$.

A general strategy is to know that $x^2 + y^2$ is your dear friend because

$$x^2 + y^2 = r^2\cos^2\theta + r^2\sin^2\theta = r^2$$

which simplifies a lot of mess. Therefore, you should try your best to put yourself in position to see $x^2 + y^2$.

In this case, we expand the square $(y-3)^2$.

$$x^2 + y^2 - 6y + 9 = 0 \implies x^2 + y^2 - 6y = 0 \implies r^2 - 6r\sin\theta = 0.$$

Now, if $r = 0$, we are describing a curve that always has 0 distance to the origin, which is, just the origin. Thus, we exclude this case, and cancel r from both sides,

$$r - 6\sin\theta = 0$$

is the polar equation for the circle centered at $(0, 3)$ with radius 3.

LECTURE XIX AREA AND LENGTH OF CURVES IN POLAR COORDINATES

Area of Curves in Polar Coordinates

How do we calculate the area of a pizza slice? It is a sector of radius r and angle θ.

$$A = \frac{1}{2}r^2\theta.$$

In fact, you can obtain this formula by using parametrization of a circle, namely,

$$x = r\cos t, \quad y = r\sin t, \quad 0 \leq t \leq \theta.$$

Then,

$$\begin{aligned} A &= \int_\theta^0 r\sin t\left(-r\sin^2 t\right) dt \\ &= r^2\int_0^\theta \sin^2 t dt \\ &= r^2\int_0^\theta \frac{1+\cos 2t}{2} dt \\ &= r^2\left[\frac{t}{2}+\frac{\sin 2t}{4}\right]_0^\theta \\ &= \frac{1}{2}r^2\theta \end{aligned}$$

In practice, we approximate the area under a general curve using different pizza slices, where the directed distance r is a function of θ, i.e., $r = f(\theta)$. Therefore, we add these slices and formulate the following expression for the area of the fan-shaped region between the origin and the curve $r = f(\theta)$, when $\alpha \leq \theta \leq \beta$ and $\beta - \alpha \leq 2\pi$ (without repeating itself).

$$A = \int_\alpha^\beta \frac{1}{2}r^2(\theta)\, d\theta.$$

Example. Find the area of the region in the xy-plane enclosed by the cardioid $r = 2(1+\cos\theta)$.

Solution.

$$\begin{aligned}
\int_0^{2\pi} \frac{1}{2} r^2(\theta)\, d\theta &= \int_0^{2\pi} \frac{1}{2}\left[2(1+\cos\theta)\right]^2 d\theta \\
&= 2\int_0^{2\pi} \left(1 + 2\cos\theta + \cos^2\theta\right) d\theta \\
&= 2\int_0^{2\pi} \left(1 + 2\cos\theta + \frac{1+\cos 2\theta}{2}\right) d\theta \\
&= 2\left[\theta + 2\sin\theta + \left(\frac{\theta}{2} + \frac{\sin 2\theta}{4}\right)\right]_0^{2\pi} \\
&= 2\left[2\pi + 0 + \pi\right] \\
&= 6\pi.
\end{aligned}$$

Area between Two Polar Curves

Now, we have done problems such as: find the area between the curve $y = x^2$, $y = 4$ and $x = 0$. The idea is to find the point of intersection, and perform a careful integration of the difference between the expressions. Here, in polar coordinates, we can do the same, as long as we can identify the which curve sits on "top". The area that lies between two curves $r_1(\theta)$ and $r_2(\theta)$ from $\theta = \alpha$ to $\theta = \beta$ is given by

$$A = \int_\alpha^\beta \frac{1}{2} r_2^2(\theta)\, d\theta - \int_\alpha^\beta \frac{1}{2} r_1^2(\theta)\, d\theta = \int_\alpha^\beta \frac{1}{2}\left(r_2^2(\theta) - r_1^2(\theta)\right) d\theta.$$

Example. Find the area of the region that lies inside the circle $r = 1$ and outside the cardioid $r = 1 - \cos\theta$.

Solution. Always, sketch first. As in Calculus II area between curves, we need to find intersections by setting the equations equal. Here, we find that

$$1 = 1 - \cos\theta \implies \theta = -\pi/2, \pi/2.$$

With a further confirmation of the sketch, we find that the region we are interested in sits in the angle interval $-\pi/2 \le \theta \le \pi/2$.

How do we determine which curve is the outer one?

The outer curve is the one further away from the origin, that is, achieving larger directed distance. In this case, the outer curve is $r_2 = 1$ and the inner curve is

$r_1 = 1 - \cos\theta$. Thus,

$$\begin{aligned}
A &= \int_{-\pi/2}^{\pi/2} \frac{1}{2}\left(r_2^2 - r_1^2\right) d\theta \\
&\overset{\text{symmetry}}{=} \int_0^{\pi/2} \left(1^2 - (1-\cos\theta)^2\right) d\theta \\
&= \int_0^{\pi/2} \left(2\cos\theta - \cos^2\theta\right) d\theta \\
&= \int_0^{\pi/2} \left(2\cos\theta - \frac{1+\cos 2\theta}{2}\right) d\theta \\
&= 2\sin\theta - \frac{\theta}{2} - \frac{\sin 2\theta}{4} \Big|_0^{\pi/2} \\
&= \left(2 - \frac{\pi}{4} - 0\right) - (0 - 0 - 0) \\
&= 2 - \frac{\pi}{4}
\end{aligned}$$

Arc Length of Curves in Polar Coordinates

Note that all we did with polar coordinates is to provide a special parametrization, i.e.,

$$\begin{aligned}
x &= r\cos\theta = f(\theta)\cos\theta, \\
y &= r\sin\theta = f(\theta)\sin\theta.
\end{aligned}$$

Thus, the arc length formula is no other than

$$\begin{aligned}
L &= \int_\alpha^\beta \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\left(f'(\theta)\cos\theta - f(\theta)\sin\theta\right)^2 + \left(f'(\theta)\sin\theta + f(\theta)\cos\theta\right)^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\begin{array}{l} \left[f'(\theta)\right]^2\cos^2\theta + [f(\theta)]^2\sin^2\theta \cancel{- 2f'(\theta)\cos\theta f(\theta)\sin\theta} \\ + \left[f'(\theta)\right]^2\sin^2\theta + [f(\theta)]^2\cos^2\theta \cancel{+ 2f'(\theta)\sin\theta f(\theta)\cos\theta} \end{array}} d\theta \\
&= \int_\alpha^\beta \sqrt{[f'(\theta)]^2 + [f(\theta)]^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\left[\frac{dr}{d\theta}\right]^2 + [r]^2} d\theta
\end{aligned}$$

Example. Find the length of the cardioid $r = 1 - \cos\theta$.

Solution. The ingredients for the arc length formula is

$$r(\theta) = 1 - \cos\theta, \quad \frac{dr}{d\theta} = \sin\theta.$$

Thus,

$$\begin{aligned}
L &= \int_0^{2\pi} \sqrt{\left[\frac{dr}{d\theta}\right]^2 + [r]^2} d\theta \\
&= \int_0^{2\pi} \sqrt{\sin^2\theta + [1-\cos\theta]^2} d\theta \\
&= \int_0^{2\pi} \sqrt{\sin^2\theta + 1 - 2\cos\theta + \cos^2\theta} d\theta \\
&= \sqrt{2}\int_0^{2\pi} \sqrt{1-\cos\theta} d\theta \\
&= \sqrt{2}\int_0^{2\pi} \sqrt{2\sin^2\left(\frac{\theta}{2}\right)} d\theta \\
&= 2\int_0^{2\pi} \sin\left(\frac{\theta}{2}\right) d\theta \\
&= 2\left[-2\cos\left(\frac{\theta}{2}\right)\right]_0^{2\pi} \\
&= 2\left[2-(-2)\right] \\
&= 8.
\end{aligned}$$

We use the identity

$$\cos(2t) = 1 - 2\sin^2(t) \implies 2\sin^2(t) = 1 - \cos(2t)$$

and now set $t = \theta/2$.

LECTURE XX WORKED EXAMPLES IN POLAR COORDINATES

Since graphing curves in polar coordinates is essential to area and arc length computation, we should learn it in a more general setting. First, it would be useful to observe symmetry in the curve and thus significantly reduce pain. Think of this as an advil for polar graphs.

Symmetry Tests for Polar Graphs in the Cartesian xy-Plane.

(1) **Symmetry about the x-axis**: If the point (r, θ) lies on the graph, then the point $(r, -\theta)$ or $(-r, \pi - \theta)$ lies on the graph.
(2) **Symmetry about the y-axis:** If the point (r, θ) lies on the graph, then the point $(r, \pi - \theta)$ or $(-r, -\theta)$ lies on the graph.
(3) **Symmetry about the origin:** If the point (r, θ) lies on the graph, then the point $(-r, \theta)$ or $(r, \pi + \theta)$ lies on the graph.

Example. Note the symmetry in the following curves in polar coordinates without plotting.

(1) The cardioid $r = 1 - \cos\theta$.

Solution. Check the symmetry one-by-one. Let (r, θ) be a point on the curve.

(a) (x-**axis**) Plug in now $(r, -\theta)$ and see if the equation stands?

$$r(-\theta) = 1 - \cos(-\theta) = 1 - \cos(\theta)$$

YES!

(b) (y-**axis**) Now, check $(r, \pi - \theta)$.

$$r(\pi - \theta) = 1 - \cos(\pi - \theta).$$

How to find $\cos(\pi - \theta)$? We use HW0, the angle formula,

$$\begin{aligned}\cos(\pi - \theta) &= \cos(\pi)\cos(-\theta) - \sin(\pi)\sin(-\theta) \\ &= -\cos\theta\end{aligned}$$

and thus

$$r(\pi - \theta) = 1 - \cos(\pi - \theta) = 1 + \cos\theta,$$

oops, not the original equation. The shape is NOT symmetric about the y-**axis**.

(c) (origin) Plug in $(-r, \theta)$. But we find that

$$-r = -(1 - \cos\theta) \neq 1 - \cos\theta$$

which is NOT the original equation. Thus, this curve is NOT symmetric about the origin.

(2) The cardioid $r = 1 + \sin\theta$.

Solution. Check the symmetry one-by-one. Let (r, θ) be a point on the curve.

(a) (x-**axis**) Plug in now $(r, -\theta)$ and see if the equation stands?

$$r(-\theta) = 1 + \sin(-\theta) = 1 - \sin(\theta)$$

YES!

(b) (y-**axis**) Now, check $(r, \pi - \theta)$.

$$r(\pi - \theta) = 1 + \sin(\pi - \theta).$$

How to find $\sin(\pi - \theta)$? We use HW0, the angle formula,

$$\begin{aligned} \sin(\pi - \theta) &= \sin(\pi)\cos(-\theta) + \cos(\pi)\sin(-\theta) \\ &= (0)\cos\theta + (-1)(-\sin\theta) \\ &= \sin\theta. \end{aligned}$$

and thus

$$r(\pi - \theta) = 1 + \sin(\pi - \theta) = 1 + \sin\theta,$$

oops, not the original equation. The shape is NOT symmetric about the y-**axis**.

(c) (origin) Plug in $(-r, \theta)$. But we find that

$$-r = -(1 + \sin\theta) \neq 1 + \sin\theta$$

which is NOT the original equation. Thus, this curve is NOT symmetric about the origin.

(3) The four-leaved rose $r = 1 + \cos(4\theta)$.

Solution. Check the symmetry one-by-one. Let (r, θ) be a point on the curve.

(a) (x-**axis**) Plug in now $(r, -\theta)$ and see if the equation stands?

$$r(-\theta) = 1 + \cos(-4\theta) = 1 + \cos(4\theta)$$

YES!

(b) (y-**axis**) Now, check $(r, \pi - \theta)$.

$$r(\pi - \theta) = 1 + \cos(4(\pi - \theta)).$$

How to find $\cos(2(\pi - \theta))$? We use HW0, the angle formula,

$$\begin{aligned} \cos(4(\pi - \theta)) = \cos(4\pi - 4\theta) &= \cos(4\pi)\cos(-4\theta) - \sin(4\pi)\sin(-4\theta) \\ &= \cos 4\theta \end{aligned}$$

and thus

$$r(\pi - \theta) = \cos(4(\pi - \theta)) = \cos 4\theta,$$

YES! The shape is symmetric about the y-**axis**.

(c) (**origin**) Plug in $(r, \pi + \theta)$. But we find that

$$\begin{aligned} r(\pi + \theta) &= 1 + \cos(4(\pi + \theta)) \\ &= 1 + \cos(4\pi)\cos(4\theta) - \sin(4\pi)\sin(4\theta) \\ &= 1 + \cos(4\theta) \end{aligned}$$

which IS the original equation. Thus, this curve IS symmetric about the origin.

Slope.

To describe a polar curve, we must consider the radial distance r as a function of θ, namely, $r = f(\theta)$. Given the polar coordinate transformation,

$$\begin{aligned} x &= r\cos\theta = f(\theta)\cos\theta \\ y &= r\sin\theta = f(\theta)\sin\theta \end{aligned}$$

a set of parametric equations. If f is differentiable in θ, and if $dx/d\theta \neq 0$, then we can compute the derivative dy/dx (or slope),

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy/d\theta}{dx/d\theta} \\ &= \frac{\frac{d}{d\theta}(f(\theta)\sin\theta)}{\frac{d}{d\theta}(f(\theta)\cos\theta)} \\ &= \frac{f'(\theta)\sin\theta + f(\theta)\cos\theta}{f'(\theta)\cos\theta - f(\theta)\sin\theta} \end{aligned}$$

Note that from this formula alone, we can calculate the slope of the curve if it passes through the origin at some angle $\theta = \theta_0$. The origin corresponds to a radial distance of 0, that is, $f(\theta_0) = 0$. Thus,

$$\begin{aligned} \frac{dy}{dx}\Big|_{r=0,\theta=\theta_0} &= \frac{f'(\theta_0)\sin\theta_0 + \cancelto{0}{f(\theta_0)}\cos\theta_0}{f'(\theta_0)\cos\theta_0 - \cancelto{0}{f(\theta_0)}\sin\theta_0} \\ &= \frac{f'(\theta_0)\sin\theta_0}{f'(\theta_0)\cos\theta_0} \\ &= \tan\theta_0 \end{aligned}$$

This is the **slope at** $(0, \theta_0)$, and not just "slope at the origin" because a polar curve may pass through the origin more than once, with possibly different slopes at different θ-values, i.e., three-leaved rose.

Graph of Polar Curves.

Example. Graph the **cardioid** $r = 1 - \cos\theta$ in the Cartesian xy-plane.

Solution. The curve is symmetric about the x-axis because given a point (r, θ) on the graph, we have

$$r = 1 - \cos\theta \implies r = 1 - \cos(-\theta)$$

which implies $(r, -\theta)$ is also on the graph. This suggests that we just need to graph the curve above the x-axis.

As θ increases from 0 to π, $\cos\theta$ goes from 1 to -1, which then implies that $r = 1 - \cos\theta$ increases from a minimum value of 0 to a maximum value of 2.

Note that since $\cos\theta$ is 2π-periodic, this curve also starts repeating itself once $\theta = 2\pi$.

Area between Curves

Consider the area inside the cardioid $r = 2(1 + \cos\theta)$ but outside the circle $r = 2$.

First, we need to find the intersection points. Setting both equations equal, we have

$$2 = 2(1+\cos\theta) \implies 1 = 1+\cos\theta \implies \cos\theta = 0 \implies \theta = -\pi/2, \pi/2.$$

The outer curve is $r_2(\theta) = 2(1+\cos\theta)$ while the inner curve is $r_1(\theta) = 2$. Thus, this region of interest has area (using symmetry about x-axis)

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \frac{1}{2}\left(r_2^2(\theta) - r_1^2(\theta)\right) d\theta &= 2\int_0^{\pi/2} \frac{1}{2}\left(r_2^2(\theta) - r_1^2(\theta)\right) d\theta \\
&= \int_0^{\pi/2} \left[4(1+\cos\theta)^2 - 2^2\right] d\theta \\
&= 4\int_0^{\pi/2} \left[1 + 2\cos\theta + \cos^2\theta - 1\right] d\theta \\
&= 4\int_0^{\pi/2} \left[2\cos\theta + \frac{1+\cos 2\theta}{2}\right] d\theta \\
&= 4\left[2\sin\theta + \frac{\theta}{2} + \frac{\sin 2\theta}{4}\right]_0^{\pi/2} \\
&= 4\left[2\sin\frac{\pi}{2} + \frac{\pi}{4} + \frac{\sin\pi}{4} - 0 - 0 - 0\right] \\
&= 4\left[2 + \frac{\pi}{4}\right] \\
&= 8 + \pi.
\end{aligned}$$

Arc Length of Curves in Polar Coordinates

Note that all we did with polar coordinates is to provide a special parametrization, i.e.,

$$\begin{aligned}
x &= r\cos\theta = f(\theta)\cos\theta, \\
y &= r\sin\theta = f(\theta)\sin\theta.
\end{aligned}$$

Thus, the arc length formula is no other than

$$\begin{aligned}
L &= \int_\alpha^\beta \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\left(f'(\theta)\cos\theta - f(\theta)\sin\theta\right)^2 + \left(f'(\theta)\sin\theta + f(\theta)\cos\theta\right)^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\begin{array}{l} \left[f'(\theta)\right]^2\cos^2\theta + [f(\theta)]^2\sin^2\theta \cancel{- 2f'(\theta)\cos\theta f(\theta)\sin\theta} \\ + \left[f'(\theta)\right]^2\sin^2\theta + [f(\theta)]^2\cos^2\theta \cancel{+ 2f'(\theta)\sin\theta f(\theta)\cos\theta} \end{array}} d\theta \\
&= \int_\alpha^\beta \sqrt{[f'(\theta)]^2 + [f(\theta)]^2} d\theta \\
&= \int_\alpha^\beta \sqrt{\left[\frac{dr}{d\theta}\right]^2 + [r]^2} d\theta
\end{aligned}$$

Example. Find the length of the cardioid $r = 1 - \cos\theta$.

Solution. The ingredients for the arc length formula is

$$r(\theta) = 1 - \cos\theta, \quad \frac{dr}{d\theta} = \sin\theta.$$

Thus,

$$\begin{aligned}
L &= \int_0^{2\pi} \sqrt{\left[\frac{dr}{d\theta}\right]^2 + [r]^2} d\theta \\
&= \int_0^{2\pi} \sqrt{\sin^2\theta + [1-\cos\theta]^2} d\theta \\
&= \int_0^{2\pi} \sqrt{\sin^2\theta + 1 - 2\cos\theta + \cos^2\theta} d\theta \\
&= \sqrt{2}\int_0^{2\pi} \sqrt{1-\cos\theta} d\theta \\
&= \sqrt{2}\int_0^{2\pi} \sqrt{2\sin^2\left(\frac{\theta}{2}\right)} d\theta \\
&= 2\int_0^{2\pi} \sin\left(\frac{\theta}{2}\right) d\theta \\
&= 2\left[-2\cos\left(\frac{\theta}{2}\right)\right]_0^{2\pi} \\
&= 2\left[2-(-2)\right] \\
&= 8.
\end{aligned}$$

We use the identity

$$\cos(2t) = 1 - 2\sin^2(t) \implies 2\sin^2(t) = 1 - \cos(2t) \implies$$

and now set $t = \theta/2$.

LECTURE XXI VECTORS AND THE GEOMETRY OF SPACE

Coordinates in 3D

The study of single variable calculus has been finished (at least temporarily). We now embark on one of the liberating journey of Multivariable Calculus.

We are used to seeing the Cartesian coordinates (x, y) that represent a point in the xy-plane. Very clearly, this is NOT enough to describe the visible physical world. If you need a reminder of the number of dimensions we live in, just look at the corner of a wall, where the floor represents the familiar xy-plane, and the vertical line represents a **new axis**, usually referred to as the z-axis.

With the added dimension, we describe a point in 3D space using the triple (x, y, z). If x, y, z are all real numbers, we often write $x \in \mathbb{R}$, or $(x, y, z) \in \mathbb{R}^3 = \mathbb{R}\times\mathbb{R}\times\mathbb{R}$, where $\mathbb{R}$ represents the set of all real numbers. $\mathbb{R}^3 = \mathbb{R}\times\mathbb{R}\times\mathbb{R}$ represents the **Cartesian product**, i.e., the set of all ordered pairs of real numbers (in this case, $\mathbb{R}^3$ is a set of triples who components are real).

Basic Objects in 3D

The coordinate axis has the following expression

(1) x-axis: $(x, 0, 0)$, or $y = z = 0$.
(2) y-axis: $(0, y, 0)$, or $x = z = 0$.
(3) z-axis: $(0, 0, z)$, or $x = y = 0$.

Note that these axes are **perpendicular** to each other. Furthermore, if you want a line parallel to one of the axis, just need to specify the two other variables. For example, the line $x = 2$, $z = 5$ is parallel to the y-axis.

Now, in addition to the principal axes, where at least two coordinates are zero, we also have the case where only one variable is zero. This situation gives rise to **coordinate planes**.

(1) xy-plane: $(x, y, 0)$, or $z = 0$.
(2) xz-plane: $(x, 0, z)$, or $y = 0$.
(3) yz-plane: $(0, y, z)$, or $x = 0$.

Note that these planes are **perpendicular** to each other.

Just like in 2D where parallel lines are possible, we may consider parallel planes in 3D. Consider the **coordinate planes** as bases of comparison. The points in a plane perpendicular to the x-axis all have the same x-coordinate, this being the number at which that plane cuts the x-axis; the y- and z-coordinates of points in this plane can be **any** numbers. For example, the plane $x = 2$ can also be written more precisely, using set notation,

$$\{x = 2\} = \{x, y, z \in \mathbb{R} \mid x = 2\}.$$

Example. Equations and inequalities in 3D.

(1) $z \geq 0$: the half-space consisting of the points on and above the xy-plane.
(2) $x = -3$: the plane perpendicular to the x-axis at $x = -3$. This plane lies parallel to the yz-plane and 3 units behind it.

(3) $z = 0$, $x \leq 0$, $y \geq 0$: the second quadrant of the xy-plane.
(4) $x \geq 0$, $y \geq 0$, $z \geq 0$: the first octant.
(5) $-1 \leq y \leq 1$: the vertical slab between the planes $y = -1$ and $y = 1$ (planes included).
(6) $y = -2$, $z = 2$: the line in which the planes $y = -2$ and $z = 2$ intersect. Alternatively, it is the line through the point $(0, -2, 2)$, parallel to the x-axis.

With these preparations, we can begin to describe a few more complex curves in space.

Example. (Circles at a height) The equations

$$x^2 + y^2 = 4, \quad z = 3$$

represents the circle of radius 2, centered at the z-axis, in the plane of $z = 3$ (or at a height of $z = 3$).

Distance in 3D

Suppose we are given two points in 3D, $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$, the distance between the two points has a similar Pythagorus feel,

$$|P_1P_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Example. The distance between $P_1 = (2, 1, 5)$ and $P_2 = (-2, 3, 0)$ is

$$\begin{aligned} |P_1P_2| &= \sqrt{(-2-2)^2 + (3-1)^2 + (0-5)^2} \\ &= \sqrt{16 + 4 + 25} \\ &= \sqrt{45}. \end{aligned}$$

Note that you can also do

$$\begin{aligned} |P_2P_1| &= \sqrt{(2-(-2))^2 + (3-1)^2 + (5-0)^2} \\ &= \sqrt{45} \end{aligned}$$

because we are squaring the coordinate differences anyways, so the order doesn't matter. But make sure that you stay consistent with which order.

There is a special surface whose points to a particular point in space is always a constant. It is a sphere. By the description alone, define the **center** of the **sphere** (x_0, y_0, z_0), then the points (x, y, z) living on the surface of the sphere satisfies

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = a^2$$

for some a, the radius of the sphere.

Example. Find the center and radius of the sphere

$$x^2 + y^2 + z^2 + 3x - 4z + 1 = 0.$$

Solution. This is an exercise of "Completing the square", a technique very useful in mathematical procedures found in engineering and sciences.

$$\begin{aligned}
&\left(x^2+3x\right)+y^2+\left(z^2-4z\right)+1=0\\
\Longrightarrow\ &\left(x^2+3x+\frac{9}{4}-\frac{9}{4}\right)+y^2+\left(z^2-4z+4-4\right)+1=0\\
\Longrightarrow\ &\left(x+\frac{3}{2}\right)^2+y^2+(z-2)^2-\frac{9}{4}-4+1=0\\
\Longrightarrow\ &\left(x+\frac{3}{2}\right)^2+y^2+(z-2)^2=\frac{21}{4}.
\end{aligned}$$

Thus, the center of the sphere is $\left(-\frac{3}{2},0,2\right)$, and the radius is $\frac{\sqrt{21}}{2}$.

LECTURE XXII VECTORS

Vectors

Earlier, we said a point in 3D space can be represented by the Cartesian coordinate (x, y, z). Suppose now we have the coordinates of another point (x_0, y_0, z_0). How do I know the position of this new point **relative** to the previous point? The easiest way is to connect the two points with a line segment, and give the line a direction. Thus, you can say, certain point is "southwest" of another, etc, etc, at a distance of certain units away. We call this **directed line segment** a **vector**. The starting point of this direction is called the **initial point**, while the destination point is called **terminal point**. The distance between the two points can be calculated using the distance formula.

Two vectors are same if they have the same direction and same magnitude (length). Note here that the two points provided above are completely arbitrary. Therefore, we may find examples of vectors with different initial points and terminal points that are in fact equal.

Definition. If $\boldsymbol{v}$ is a two-dimensional vectorin the plane equal to the vector with initial point at the origin and terminal point (v_1, v_2), then the **component form** of $\boldsymbol{v}$ is

$$\boldsymbol{v} = \langle v_1, v_2 \rangle .$$

Similarly, a 3D parallel is

$$\boldsymbol{v} = \langle v_1, v_2, v_3 \rangle .$$

The magnitude or length of the vector $\boldsymbol{v}$, connecting the two points (x_1, y_1, z_1) and (x_2, y_2, z_2) is

$$|\boldsymbol{v}| = \sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Example. Find the component form and length of the vector with initial point $P(-3, 4, 1)$ and terminal point $Q(-5, 2, 2)$.

$$\boldsymbol{v} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) = \langle -5 - (-3), 2 - 4, 2 - 1 \rangle = \langle -2, -2, 1 \rangle .$$

The length is

$$|\boldsymbol{v}| = \sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{(-2)^2 + (-2)^2 + (1)^2} = 3.$$

Vector Algebra Operations

Two principal operations involving vectors are **vector addition** and **scalar multiplication**.
Let $\boldsymbol{u} = \langle u_1, u_2, u_3 \rangle$ and $\boldsymbol{v} = \langle v_1, v_2, v_3 \rangle$ be vectors, and k a scalar. Then,

$$\begin{aligned} \boldsymbol{u} + \boldsymbol{v} &= \langle u_1 + v_1, u_2 + v_2, u_3 + v_3 \rangle . \\ k\boldsymbol{u} &= \langle ku_1, ku_2, ku_3 \rangle . \end{aligned}$$

The first law is also called **principal of superposition of vectors in** $\mathbb{R}^3$. The sum $\boldsymbol{u} + \boldsymbol{v}$ is also a vector, called the **resultant vector**.

Example. Let $\boldsymbol{u} = \langle -1, 3, 1 \rangle$ and $\boldsymbol{v} = \langle 4, 7, 0 \rangle$. Find the components of

(1) $2\boldsymbol{u} + 3\boldsymbol{v}$

Solution.

$$2\boldsymbol{u} + 3\boldsymbol{v} = 2\langle -1, 3, 1 \rangle + 3\langle 4, 7, 0 \rangle = \langle -2, 6, 2 \rangle + \langle 12, 21, 0 \rangle = \langle 10, 27, 2 \rangle .$$

(2) $\boldsymbol{u} - \boldsymbol{v}$

Solution. $\boldsymbol{u} - \boldsymbol{v} = \langle -1, 3, 1 \rangle - \langle 4, 7, 0 \rangle = \langle -1 - 4, 3 - 7, 1 - 0 \rangle = \langle -5, -4, 1 \rangle .$

(3) $\left|\frac{1}{2}\boldsymbol{u}\right|$ (the length).

Solution. $\left|\frac{1}{2}\boldsymbol{u}\right| = \left|\left\langle -\frac{1}{2}, \frac{3}{2}, \frac{1}{2} \right\rangle\right| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{2}\sqrt{11}.$

Properties of Vector Operations

Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ be vectors and a, b be scalars.

(1) (additive commutativity) $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
(2) (additive associativity) $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
(3) (additive identity) $\boldsymbol{u}+\boldsymbol{0}=\boldsymbol{u}$.
(4) (additive inverse) $\boldsymbol{u}+(-\boldsymbol{u})=0$.
(5) $0\boldsymbol{u}=0$.
(6) $1\boldsymbol{u}=\boldsymbol{u}$.
(7) $a\,(b\boldsymbol{u})=(ab)\,\boldsymbol{u}$.
(8) $a\,(\boldsymbol{u}+\boldsymbol{v})=a\boldsymbol{u}+a\boldsymbol{v}$.
(9) $(a+b)\,\boldsymbol{u}=a\boldsymbol{u}+b\boldsymbol{u}$.

Unit Vectors

A vector $\boldsymbol{v}$ of length 1 is called a **unit vector**. The **standard unit vectors are**

$$i=\langle 1,0,0\rangle\,,\quad j=\langle 0,1,0\rangle\,,\quad k=\langle 0,0,1\rangle\,.$$

Any vector $\boldsymbol{v}=\langle v_1,v_2,v_3\rangle$ can be written as a **linear combination** of the standrrd unit vectors

$$\begin{aligned}\boldsymbol{v}&=\langle v_1,v_2,v_3\rangle\\&=\langle v_1,0,0\rangle+\langle 0,v_2,0\rangle+\langle 0,0,v_3\rangle\\&=v_1\langle 1,0,0\rangle+v_2\langle 0,1,0\rangle+v_3\langle 0,0,1\rangle\\&=v_1\boldsymbol{i}+v_2\boldsymbol{j}+v_3\boldsymbol{k}.\end{aligned}$$

The component form for the vector from $P_1\,(x_1,y_1,z_1)$ to $P_2\,(x_2,y_2,z_2)$ is

$$\overrightarrow{P_1P_2}=(x_2-x_1)\,\boldsymbol{i}+(y_2-y_1)\,\boldsymbol{j}+(z_2-z_1)\,\boldsymbol{k}.$$

Any nonzero vector has a nonzero length, and

$$\left|\frac{1}{|\boldsymbol{v}|}\boldsymbol{v}\right|=\frac{1}{|\boldsymbol{v}|}\,|\boldsymbol{v}|=1$$

which suggests that $\boldsymbol{v}/\,|\boldsymbol{v}|$ is a unit vector in the direction of $\boldsymbol{v}$, called the **direction** of the nonzero vector $\boldsymbol{v}$.

Example. Find a unit vector $\boldsymbol{u}$ in the direction of the vector from $P_1\,(1,0,1)$ to $P_2\,(3,2,0)$.

(1) Find the vector in component form or ijk form.

$$\overrightarrow{P_1P_2}=(3-1)\,\boldsymbol{i}+(2-0)\,\boldsymbol{j}+(0-1)\,\boldsymbol{k}=2\boldsymbol{i}+2\boldsymbol{j}-\boldsymbol{k}.$$

(2) Find the length of $\overrightarrow{P_1P_2}$.

$$\left|\overrightarrow{P_1P_2}\right|=\sqrt{(2)^2+(2)^2+(-1)^2}=3.$$

(3) Find the unit vector

$$\boldsymbol{u}=\frac{\overrightarrow{P_1P_2}}{\left|\overrightarrow{P_1P_2}\right|}=\frac{2\boldsymbol{i}+2\boldsymbol{j}-\boldsymbol{k}}{3}=\frac{2}{3}\boldsymbol{i}+\frac{2}{3}\boldsymbol{j}-\frac{1}{3}\boldsymbol{k}.$$

In practice, it is often useful to express the length and the direction of a vector, its two important features, in a separate fashion as follows,

$$\boldsymbol{v}=|\boldsymbol{v}|\,\frac{\boldsymbol{v}}{|\boldsymbol{v}|}.$$

The example above can be written as,

$$\overrightarrow{P_1P_2}=3\left(\frac{2}{3}\boldsymbol{i}+\frac{2}{3}\boldsymbol{j}-\frac{1}{3}\boldsymbol{k}\right)=3\left\langle\frac{2}{3},\frac{2}{3},-\frac{1}{3}\right\rangle.$$

If a question asks for the **direction** of a vector, it suffices to give the unit vector in the of the vector.

Example. A 75-N weight is suspended by two wires, making an angle of 55° and 40° against the ceiling, respectively. Find the forces $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ acting in both wires.

Solution. Forces in each component must balance. The weight is directly downward, so it must balance the total force felt in the sum of the vertical component of the two wires. Now, the components of the forces can be found using basic trigonometry,

$$\begin{aligned}\boldsymbol{F}_1&=\langle -\,|\boldsymbol{F}_1|\cos 55^\circ,|\boldsymbol{F}_1|\sin 55^\circ\rangle\,,\\\boldsymbol{F}_2&=\langle |\boldsymbol{F}_2|\cos 40^\circ,|\boldsymbol{F}_2|\sin 40^\circ\rangle\,.\end{aligned}$$

Why does the first component of $\boldsymbol{F}_1$ have a negative sign? It's because we have designated up and right as positive directions. Note that the horizontal component of $\boldsymbol{F}_1$ must be opposite of that of $\boldsymbol{F}_2$ to balance, thus it must be of an opposite direction.

By force balance we have $\langle \boldsymbol{F}_1 + \boldsymbol{F}_2 \rangle = \langle 0, 75 \rangle$ (where 75 is the force going **up** to balance the weight, and up is positive). This forms a system of equations for $|\boldsymbol{F}_1|$ and $|\boldsymbol{F}_2|$, i.e.,

$$\begin{aligned} -|\boldsymbol{F}_1|\cos 55^\circ + |\boldsymbol{F}_2|\cos 40^\circ &= 0 \\ |\boldsymbol{F}_1|\sin 55^\circ + |\boldsymbol{F}_2|\sin 40^\circ &= 75 \end{aligned}$$

Via some calculations, we found that

$$|\boldsymbol{F}_1| = \frac{75}{\sin 55^\circ + \cos 55^\circ \tan 40^\circ} \approx 57.67\,N,$$

and

$$|\boldsymbol{F}_2| = \frac{75\cos 55^\circ}{\sin 55^\circ \cos 40^\circ + \cos 55^\circ \sin 40^\circ} \approx 43.18\,N.$$

The force vectors are then

$$\begin{aligned} \boldsymbol{F}_1 &= \langle -|\boldsymbol{F}_1|\cos 55^\circ, |\boldsymbol{F}_1|\sin 55^\circ \rangle \approx \langle -33.08, 47.24 \rangle, \\ \boldsymbol{F}_2 &= \langle |\boldsymbol{F}_2|\cos 40^\circ, |\boldsymbol{F}_2|\sin 40^\circ \rangle \approx \langle 33.08, 27.76 \rangle. \end{aligned}$$

LECTURE XXIII THE DOT PRODUCT AND PROJECTIONS

The Dot Product

Consider two vectors in 2D

$$\boldsymbol{u} = \langle u_1, u_2 \rangle\,,$$
$$\boldsymbol{v} = \langle v_1, v_2 \rangle\,.$$

Now, we know that the components of $\boldsymbol{u}$ (resp. $\boldsymbol{v}$) can be further represented by trigonometry, that is,

$$\boldsymbol{u} = \langle |\boldsymbol{u}| \cos\alpha, |\boldsymbol{u}| \sin\alpha \rangle$$
$$\boldsymbol{v} = \langle |\boldsymbol{v}| \cos\beta, |\boldsymbol{v}| \sin\beta \rangle$$

where α and β are the angle made by $\boldsymbol{u}$ and $\boldsymbol{v}$ against the positive x-axis, respectively. A prime example of this decomposition is the wires with a weight example from last class (and your HW).

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are not in the same direction, then $\alpha \neq \beta$, so a natural question would be to find the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$, written as $\theta = \alpha - \beta$. One may consider the size of $\theta = \alpha - \beta$ as a measure of how similar $\boldsymbol{u}$ and $\boldsymbol{v}$ are; certainly if $\alpha = \beta$, then $\boldsymbol{u}$ and $\boldsymbol{v}$ point in the same direction, and they are considered more similar than if the vectors point in different directions.

To find $\theta = \alpha - \beta$, we need it to show up explicitly in some kind of formula. Observing from the components of $\boldsymbol{u}$ and $\boldsymbol{v}$, we may incline to multiply the components together and add them so that α and β may combine. Indeed,

$$|\boldsymbol{u}| \cos\alpha\, |\boldsymbol{v}| \cos\beta + |\boldsymbol{u}| \sin\alpha\, |\boldsymbol{v}| \sin\beta = |\boldsymbol{u}|\, |\boldsymbol{v}| \left(\cos\alpha\cos\beta + \sin\alpha\sin\beta\right) = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\alpha - \beta\right) = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\theta\right)$$

which contains big information about θ. The LHS here is actually the sum of the component-wise product,

$$u_1 v_1 + u_2 v_2 = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\theta\right).$$

This "lucky" product seems to imply a great deal about the relationship/similarity between the two vectors, which deserves a definition on its own.

Definition. The dot product, or **the scalar product**, or **the inner product,** of vectors $\boldsymbol{u} = \langle u_1, u_2, u_3 \rangle$ and $\boldsymbol{v} = \langle v_1, v_2, v_3 \rangle$ is

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

The calculation above involving the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is characterized as a Theorem,

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\theta\right) \iff \theta = \cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}|\, |\boldsymbol{v}|}\right).$$

When $\theta = \pi/2$, $\cos\left(\pi/2\right) = 0$ which implies that the dot product is 0. This implies that $\boldsymbol{u}$ and $\boldsymbol{v}$ are perpendicular/orthogonal to each other.

Definition. Vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

This is obvious because $\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\theta\right) = |\boldsymbol{u}|\, |\boldsymbol{v}| \cos\left(\pi/2\right) = 0$.

Properties of the Dot Product

If $\boldsymbol{u}$, v and $\boldsymbol{w}$ are any vectors and c is a scalar, then

(1)
$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$$

(2)
$$\left(c\boldsymbol{u}\right) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \left(c\boldsymbol{v}\right) = c\left(\boldsymbol{u} \cdot \boldsymbol{v}\right).$$

(3)
$$\boldsymbol{u} \cdot \left(\boldsymbol{v} + \boldsymbol{w}\right) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}.$$

(4)
$$\boldsymbol{u} \cdot \boldsymbol{u} = |\boldsymbol{u}|^2.$$

(5)
$$\boldsymbol{u} \cdot 0 = 0 \cdot \boldsymbol{u} = 0.$$

Example. Find the angle between the two vectors.

(1) $\boldsymbol{v} = \boldsymbol{i} + \boldsymbol{j}$, $\boldsymbol{u} = -\boldsymbol{j}$.

Solution. We expect the answer to be $3\pi/4$. Just draw the picture.

We find $|\boldsymbol{v}| = \sqrt{2}, |\boldsymbol{u}| = 1$,and

$$\boldsymbol{u} \cdot \boldsymbol{v} = \langle 0, -1 \rangle \cdot \langle 1, 1 \rangle = -1.$$

Thus,

$$\cos\theta = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}|\,|\boldsymbol{v}|} = -\frac{1}{(\sqrt{2})\,(1)} = -\frac{\sqrt{2}}{2}.$$

Yes, $\theta = 3\pi/4$ is a solution.

(2) $\boldsymbol{v} = 5\boldsymbol{i} + \boldsymbol{j}$, $\boldsymbol{u} = 2\boldsymbol{i} + \sqrt{17}\boldsymbol{j}$.

Solution. We don't expect anything.

$$|\boldsymbol{v}| = \sqrt{5^2 + 1^2} = \sqrt{26}$$
$$|\boldsymbol{u}| = \sqrt{2^2 + 17} = \sqrt{21}$$

and

$$\boldsymbol{u} \cdot \boldsymbol{v} = \langle 5, 1 \rangle \cdot \left\langle 2, \sqrt{17} \right\rangle = 10 + \sqrt{17}.$$

Thus,

$$\cos\theta = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{u}|\,|\boldsymbol{v}|} = \frac{10 + \sqrt{17}}{\sqrt{26}\sqrt{21}}$$

and

$$\theta = \cos^{-1}\left(\frac{10 + \sqrt{17}}{\sqrt{26}\sqrt{21}}\right) \approx 52.81°.$$

To me, this is a useless problem.

Remark. Interpretations of the sign of a dot product between $\boldsymbol{u}$ and $\boldsymbol{v}$.

(1) If $\boldsymbol{u} \cdot \boldsymbol{v} > 0$, this implies that the angle θ is acute, or $0 < \theta < \frac{\pi}{2}$.
(2) If $\boldsymbol{u} \cdot \boldsymbol{v} < 0$, this implies that the angle θ is obtuse, or $\frac{\pi}{2} < \theta < \pi$.

Projections

The notion of projections first came with optics. When you are standing under the sun, you see your shadow casted on the ground. You projected yourself onto the ground using the sunlight. Mathematically, there is a very strong paralell as follows.

Given two vectors $\boldsymbol{u} = \overrightarrow{PQ}$ and $\boldsymbol{v} = \overrightarrow{PS}$, the **vector projection** of $\boldsymbol{v}$ onto a nonzero vector $\boldsymbol{u} = \overrightarrow{PS}$ is the vector $\boldsymbol{w} = \overrightarrow{PR}$ determined by dropping a perpendicular from Q to the line PS. The notation for this vector is

$$\operatorname{proj}_{\boldsymbol{v}}(\boldsymbol{u}) \qquad \text{the vector projection of } \boldsymbol{u} \text{ onto } \boldsymbol{v}.$$

Mathematically speaking, the projection takes magnitude $|\boldsymbol{u}|\cos\theta$ in the direction of $\boldsymbol{v}$, where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. So, the magnitude-direction presentation of a vector comes in handy here (recall $\boldsymbol{v} = |\boldsymbol{v}|\,\frac{\boldsymbol{v}}{|\boldsymbol{v}|}$). This implies that

$$\begin{aligned}
\operatorname{proj}_{\boldsymbol{v}}(\boldsymbol{u}) &= |\boldsymbol{u}|\cos\theta\frac{\boldsymbol{v}}{|\boldsymbol{v}|} \\
&\overset{\text{dot product formula}}{=} \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{v}|}\frac{\boldsymbol{v}}{|\boldsymbol{v}|} \\
&= \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{|\boldsymbol{v}|^2}\right)\boldsymbol{v}.
\end{aligned}$$

This derivation tells us that to compute the projection vector, we don't need to explicitly find θ.

We also say that **the scalar component of $\boldsymbol{u}$ in the direction of $\boldsymbol{v}$** is the scalar

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

Example. Find $\boldsymbol{v} \cdot \boldsymbol{u}$, $|\boldsymbol{v}|$, $|\boldsymbol{u}|$, $\cos\theta$ where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$, the scalar component of $\boldsymbol{u}$ in the direction of $\boldsymbol{v}$, and the vector $\operatorname{proj}_{\boldsymbol{v}}(\boldsymbol{u})$, where

$$\boldsymbol{v} = 5\boldsymbol{j} - 3\boldsymbol{k};$$
$$\boldsymbol{u} = \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}.$$

Solution.

$$\boldsymbol{v} \cdot \boldsymbol{u} = \langle 0, 5, -3 \rangle \cdot \langle 1, 1, 1 \rangle = 0 + 5 - 3 = 2.$$

$$|\boldsymbol{v}| = \sqrt{0^2 + 5^2 + (-3)^2} = \sqrt{34},$$
$$|\boldsymbol{u}| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}.$$

To find $\cos\theta$, we use the dot product result

$$\boldsymbol{v}\cdot\boldsymbol{u}=|\boldsymbol{v}|\,|\boldsymbol{u}|\cos\theta \implies \cos\theta=\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{|\boldsymbol{v}|\,|\boldsymbol{u}|}=\frac{2}{\sqrt{102}}.$$

The scalar component of $\boldsymbol{u}$ in the direction of $\boldsymbol{v}$ is

$$\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{|\boldsymbol{v}|}=\frac{2}{\sqrt{34}}.$$

The vector projection is

$$\begin{aligned}\operatorname{proj}_{\boldsymbol{v}}(\boldsymbol{u}) &= \left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{|\boldsymbol{v}|^2}\right)\boldsymbol{v}\\ &= \left(\frac{2}{34}\right)\langle 0,5,-3\rangle\\ &= 0\boldsymbol{i}+\frac{5}{17}\boldsymbol{j}-\frac{6}{17}\boldsymbol{k}.\end{aligned}$$

An application of the dot product is to find the total work done by a force.

Suppose you are to push a box from a point P to another point Q, with force $\boldsymbol{F}$. Let $\boldsymbol{D}=\overrightarrow{PQ}$ be the displacement vector. If θ is the angle between $\boldsymbol{F}$ and $\boldsymbol{D}$, the work done by this force is

$$W=|\boldsymbol{F}|\cos\theta\,|\boldsymbol{D}|=\boldsymbol{F}\cdot\boldsymbol{D}.$$

Example. If $|\boldsymbol{F}|=40\,\mathrm{N}$ (newtons), $|\boldsymbol{D}|=3\,\mathrm{m}$, and $\theta=60°$, the work done by $\boldsymbol{F}$ in acting from P to Q is

$$W=\boldsymbol{F}\cdot\boldsymbol{D}=|\boldsymbol{F}|\,|\boldsymbol{D}|\cos\theta=(40)(3)\cos 60°=(120)(1/2)=60\,\mathrm{J}\quad(\text{joules}).$$

A Cute Problem (Optional)

Problem. Show that the indicated diagonal of the parallelogram determined by vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ bisects the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ if $|\boldsymbol{u}|=|\boldsymbol{v}|$.

Proof. The diagonal is the vector $\boldsymbol{u}+\boldsymbol{v}$ by superposition. Let α be the angle between $\boldsymbol{u}$ and $\boldsymbol{u}+\boldsymbol{v}$, and β be that between $\boldsymbol{v}$ and $\boldsymbol{u}+\boldsymbol{v}$. They satisfy

$$\begin{aligned}\cos\alpha &= \frac{\boldsymbol{u}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{u}|\,|\boldsymbol{u}+\boldsymbol{v}|},\\ \cos\beta &= \frac{\boldsymbol{v}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{v}|\,|\boldsymbol{u}+\boldsymbol{v}|}.\end{aligned}$$

Now, if $\alpha=\beta$, we must have

$$\frac{\boldsymbol{u}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{u}|\,|\boldsymbol{u}+\boldsymbol{v}|}=\frac{\boldsymbol{v}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{v}|\,|\boldsymbol{u}+\boldsymbol{v}|}.$$

We can divide $|\boldsymbol{u}+\boldsymbol{v}|$ from both sides, or otherwise if it is zero, it means that $\boldsymbol{u}$ and $\boldsymbol{v}$ are opposite from each other, and thus do not form a parallelogram.

$$\frac{\boldsymbol{u}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{u}|}=\frac{\boldsymbol{v}\cdot(\boldsymbol{u}+\boldsymbol{v})}{|\boldsymbol{v}|}.$$

Now, computing the dot product, we have

$$\frac{|\boldsymbol{u}|^2+\boldsymbol{u}\cdot\boldsymbol{v}}{|\boldsymbol{u}|}=\frac{|\boldsymbol{v}|^2+\boldsymbol{u}\cdot\boldsymbol{v}}{|\boldsymbol{v}|}.$$

Cross multiply, we have

$$\begin{aligned}|\boldsymbol{u}|\left(|\boldsymbol{v}|^2+\boldsymbol{u}\cdot\boldsymbol{v}\right)-|\boldsymbol{v}|\left(|\boldsymbol{u}|^2+\boldsymbol{u}\cdot\boldsymbol{v}\right)&=0\\ \implies |\boldsymbol{u}|\,|\boldsymbol{v}|^2+|\boldsymbol{u}|\,\boldsymbol{u}\cdot\boldsymbol{v}-|\boldsymbol{v}|\,|\boldsymbol{u}|^2-|\boldsymbol{v}|\,\boldsymbol{u}\cdot\boldsymbol{v}&=0\\ \implies |\boldsymbol{u}|\,|\boldsymbol{v}|\,(|\boldsymbol{v}|-|\boldsymbol{u}|)+\boldsymbol{u}\cdot\boldsymbol{v}\,(|\boldsymbol{u}|-|\boldsymbol{v}|)&=0\\ \implies (|\boldsymbol{u}|\,|\boldsymbol{v}|-\boldsymbol{u}\cdot\boldsymbol{v})\,(|\boldsymbol{v}|-|\boldsymbol{u}|)&=0.\end{aligned}$$

So, we have two options,

$$|\boldsymbol{u}|\,|\boldsymbol{v}|-\boldsymbol{u}\cdot\boldsymbol{v}=0 \implies \theta=0$$

which means that the $\boldsymbol{u}$ and $\boldsymbol{v}$ are colinear, and thus there is no angle to bisect (or say you bisect the 0 degree). The other option is $|\boldsymbol{v}|=|\boldsymbol{u}|$ as desired. □

LECTURE XXIV THE CROSS PRODUCT AND THE EQUATION OF LINES AND PLANES

Now, with the help of the dot product, we are able to find the relationship between two vectors via

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}|\,|\boldsymbol{v}| \cos\theta.$$

When studying lines in the plane, this θ may help us identify how tilted one line is to another (which may be the positive x-axis). The geometric parallel of a line in 3D is a plane. We now want to study how tilted/inclined a plane is.

In 2D, two points determine a line. Thus, the parallel statement for 3D is that two lines determine a plane. Given two lines, you can **always** find a plane that contain them both.

There is a single vector $\boldsymbol{n}$ that is perpendicular to every point on any given plane. This vector is called **the normal vector** of the plane. To express this orthogonality and therefore the equation of a plane, consider an arbitrary point (x, y, z) and a given point (x_0, y_0, z_0) on the plane. Then, through these two points forms the vector

$$\boldsymbol{r} = \langle x - x_0, y - y_0, z - z_0 \rangle .$$

Therefore, this orthogonality condition,

$$\boldsymbol{n} \cdot \boldsymbol{r} = \boldsymbol{n} \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0,$$

provides an equation for the plane, that is, a plane is a set of points (x, y, z) that satisfy this equation.

So, how do we find $\boldsymbol{n}$, the normal vector? We need two lines/vectors in the plane because $\boldsymbol{n}$ is perpendicular to both lines.

Definition. Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ that are not colinear (one is not a scalar multiple of another), then,

$$\boldsymbol{u} \times \boldsymbol{v} = (|\boldsymbol{u}|\,|\boldsymbol{v}| \sin\theta)\, \boldsymbol{n}$$

where $\boldsymbol{n}$ is the unit vector perpendicular to $\boldsymbol{u}$ and $\boldsymbol{v}$ and points in the direction of the right thumb, as your fingers curl through the angle θ from $\boldsymbol{u}$ to $\boldsymbol{v}$.

Remark. Nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel if and only if

$$\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{0}.$$

Of course, when $\theta = 0$ (parallel) or $\theta = \pi$ (antiparallel), $\sin\theta = 0$.

This formula alone still does not help us fully determine the vector $\boldsymbol{u} \times \boldsymbol{v}$. However, we may derive the following properties, which eventually help us compute $\boldsymbol{u} \times \boldsymbol{v}$.

Properties of the Cross Product. Let $\boldsymbol{u}, \boldsymbol{v}$ and $\boldsymbol{w}$ be any vectors in $\mathbb{R}^3$, and r, s are scalars, then

(1) $(r\boldsymbol{u}) \times (s\boldsymbol{v}) = (rs)\,(\boldsymbol{u} \times \boldsymbol{v})$.
(2) (distributivity over addition) $\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w}$.
(3) (anticommutativity) $\boldsymbol{v} \times \boldsymbol{u} = -\,(\boldsymbol{u} \times \boldsymbol{v})$.

(4) $(\boldsymbol{v}+\boldsymbol{w})\times\boldsymbol{u}=\boldsymbol{v}\times\boldsymbol{u}+\boldsymbol{w}\times\boldsymbol{u}$.
(5) $\boldsymbol{0}\times\boldsymbol{u}=\boldsymbol{0}$.
(6) $\boldsymbol{u}\times(\boldsymbol{v}\times\boldsymbol{w})=(\boldsymbol{u}\cdot\boldsymbol{w})\,\boldsymbol{v}-(\boldsymbol{u}\cdot\boldsymbol{v})\,\boldsymbol{w}$.

The properties we use to calculate $\boldsymbol{u}\times\boldsymbol{v}$ is the scalar distributive law, item 1, the additive distributive law, item 2 and the anticommutativity, item 3. But before we do, we may note that the cross product of the principal unit vectors is in fact doable. Consider

$$\boldsymbol{i}=\langle 1,0,0\rangle\,,\quad \boldsymbol{j}=\langle 0,1,0\rangle\,,\quad \boldsymbol{k}=\langle 0,0,1\rangle\,.$$

Think of them as the x-axis, y-axis and z-axis.

So, what is $\boldsymbol{i}\times\boldsymbol{j}$? We look in the xy-plane. Curl our fingers from the x-axis to the y-axis, we find our thumb pointing up. Thus,

$$\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k}.$$

Doing this for other axes, we find, collectively,

(1) $\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k}=-(\boldsymbol{j}\times\boldsymbol{i})$.
(2) $\boldsymbol{j}\times\boldsymbol{k}=\boldsymbol{i}=-(\boldsymbol{k}\times\boldsymbol{i})$.
(3) $\boldsymbol{k}\times\boldsymbol{i}=\boldsymbol{j}=-(\boldsymbol{i}\times\boldsymbol{k})$.

Furthermore, since $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are colinear to themselves respectively, we must have

(1) $\boldsymbol{i}\times\boldsymbol{i}=0$.
(2) $\boldsymbol{j}\times\boldsymbol{j}=0$.
(3) $\boldsymbol{k}\times\boldsymbol{k}=0$.

Explicit Computation of the Cross Product.

So, we are now in full position to compute explicitly the cross product $\boldsymbol{u}\times\boldsymbol{v}$. Recall the component representation of vectors,

$$\begin{aligned}\boldsymbol{u}&=u_1\boldsymbol{i}+u_2\boldsymbol{j}+u_3\boldsymbol{k}\\ \boldsymbol{v}&=v_1\boldsymbol{i}+v_2\boldsymbol{j}+v_3\boldsymbol{k}\end{aligned}$$

We perform the cross product using the distributive laws and anticommutativity. Note that the

$$\begin{aligned}\boldsymbol{u}\times\boldsymbol{v}&=(u_1\boldsymbol{i}+u_2\boldsymbol{j}+u_3\boldsymbol{k})\times(v_1\boldsymbol{i}+v_2\boldsymbol{j}+v_3\boldsymbol{k})\\
&=\cancelto{0}{u_1\boldsymbol{i}\times v_1\boldsymbol{i}}+\boxed{u_2\boldsymbol{j}\times v_1\boldsymbol{i}}+\underline{u_3\boldsymbol{k}\times v_1\boldsymbol{i}}\\
&\quad+\boxed{u_1\boldsymbol{i}\times v_2\boldsymbol{j}}+\cancelto{0}{u_2\boldsymbol{j}\times v_2\boldsymbol{j}}+\underbrace{u_3\boldsymbol{k}\times v_2\boldsymbol{j}}\\
&\quad+\underline{u_1\boldsymbol{i}\times v_3\boldsymbol{k}}+\underbrace{u_2\boldsymbol{j}\times v_3\boldsymbol{k}}+\cancelto{0}{u_3\boldsymbol{k}\times v_3\boldsymbol{k}}\\
&=\boxed{u_1\boldsymbol{i}\times v_2\boldsymbol{j}+u_2\boldsymbol{j}\times v_1\boldsymbol{i}}\\
&\quad+\underline{u_3\boldsymbol{k}\times v_1\boldsymbol{i}+u_1\boldsymbol{i}\times v_3\boldsymbol{k}}\\
&\quad+\underbrace{u_3\boldsymbol{k}\times v_2\boldsymbol{j}+u_2\boldsymbol{j}\times v_3\boldsymbol{k}}\\
&=(u_1v_2-u_2v_1)\,\boldsymbol{k}+(u_3v_1-u_1v_3)\,\boldsymbol{j}+(u_2v_3-u_3v_2)\,\boldsymbol{i}\\
&=(u_2v_3-u_3v_2)\,\boldsymbol{i}-(u_1v_3-u_3v_1)\,\boldsymbol{j}+(u_1v_2-u_2v_1)\,\boldsymbol{k}\end{aligned}$$

The way to remember this is, whichever component you are trying to compute, ignore that component in $\boldsymbol{u}$ and $\boldsymbol{v}$ and simply take a cross difference of the remaining entries. For example, when looking at the first component $\boldsymbol{i}$, we ignore the first component of $\boldsymbol{u}$ and $\boldsymbol{v}$, and just consider

$$(u_2, u_3) \quad (v_2, v_3).$$

Take $u_2v_3 - u_3v_2$, some kind of a cross difference. So on and so forth.

Another way to look at is via the determinant of a matrix,

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2v_3 - u_3v_2)\,\boldsymbol{i} - (u_1v_3 - u_3v_1)\,\boldsymbol{j} + (u_1v_2 - u_2v_1)\,\boldsymbol{k}.$$

Equation of the Plane Strategy.

Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, and a point on the plane (x_0, y_0, z_0),

(1) Find $\boldsymbol{n} = \langle A, B, C\rangle$, the normal vector perpendicular to both $\boldsymbol{u}$ and $\boldsymbol{v}$ by computing $\boldsymbol{n} = \boldsymbol{u} \times \boldsymbol{v}$.
(2) Form a vector that lives in the plane

$$\boldsymbol{r} = \langle x - x_0, y - y_0, z - z_0\rangle.$$

(3) Exercise the **orthogonality condition**

$$\boldsymbol{n} \cdot \boldsymbol{r} = \langle A, B, C\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = A\,(x - x_0) + B\,(y - y_0) + C\,(z - z_0) = 0.$$

Example. Find the equation of the plane that passes through $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$.

Solution. We need two vectors in this plane. We are blessed with three points.

$$\boldsymbol{u} = (0, 1, 0) - (1, 0, 0) = \langle -1, 1, 0\rangle$$
$$\boldsymbol{v} = (0, 0, 1) - (1, 0, 0) = \langle -1, 0, 1\rangle$$

Now, the cross product

$$\begin{aligned} \boldsymbol{n} = \boldsymbol{u} \times \boldsymbol{v} &= (u_2v_3 - u_3v_2)\,\boldsymbol{i} - (u_1v_3 - u_3v_1)\,\boldsymbol{j} + (u_1v_2 - u_2v_1)\,\boldsymbol{k} \\ &= ((1)\,(1) - (0)\,(0))\,\boldsymbol{i} - ((-1)\,(1) - (0)\,(-1))\,\boldsymbol{j} + ((-1)\,(0) - (1)\,(-1))\,\boldsymbol{k} \\ &= \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k} \\ &= \langle 1, 1, 1\rangle. \end{aligned}$$

Thus, given an arbitrary point (x, y, z) on the plane,

$$\boldsymbol{r} = \langle x - 1, y - 0, z - 0\rangle$$

is a vector that lives on the plane. Thus, the orthogonality condition is

$$\boldsymbol{n} \cdot \boldsymbol{r} = \langle 1, 1, 1\rangle \cdot \langle x - 1, y - 0, z - 0\rangle = 0$$

which simplifies to

$$x - 1 + y + z = 0 \implies x + y + z = 1.$$

Remark. The region bounded by this plane and the coordinate planes is called a simplex, an item that finds mysterious but important applications in network/graph theory, topological data science, etc, all of which intersect with the whitehot “Machine Learning”.