# Mathematics II

*Class Notes, Example Problems (Answers ver.)*

*Written by* :

*Lee Sihoo* / *Lee Hojun*

*KOREA SCIENCE ACADEMY OF KAIST*

2025-12-04

# Preface

Welcome to **Mathematics II**, a comprehensive exploration of advanced mathematical concepts that will expand your understanding of how mathematics describes and analyzes the world around us.

## About This Textbook

This textbook has been thoughtfully designed to guide you through four fundamental areas of advanced mathematics:

**Polar Coordinates, Parametric Equations, and Vectors** — Moving beyond the rectangular coordinate system to explore dynamic representations of position, motion, and force in multi-dimensional space.

**Matrices** — Discovering powerful tools for organizing data, solving linear systems, and transforming geometric space through elegant algebraic operations.

**Conic Sections** — Investigating the beautiful curves that emerge from intersecting planes with cones, revealing the geometric patterns that govern planetary orbits and architectural designs.

**Probability and Statistics** — Transitioning from deterministic mathematics to the realm of uncertainty, developing tools to quantify chance and extract meaning from complex datasets.

## How to Use This Book

This textbook follows a structured learning approach designed to build deep understanding. You will encounter various block types throughout:

**Definition blocks** introduce key concepts and terminology with precise mathematical language.

**Example blocks** present problems that illustrate the application of concepts, followed by detailed step-by-step solutions.

**Solution blocks** provide complete, worked-out answers with clear reasoning at each step.

**Theorem blocks** state important mathematical results that you can rely upon in your work.

**Proof blocks** demonstrate the logical reasoning behind theorems, helping you understand why results are true.

**Note blocks** offer helpful insights, common pitfalls to avoid, and connections between ideas.

## A Note on Learning

Mathematics is not a spectator sport. While reading through definitions and examples is important, true understanding comes from *active engagement*. We encourage you to:

- Work through each example *before* looking at the solution
- Try to understand *why* each step follows from the previous one
- Connect new concepts to what you already know
- Practice additional problems to reinforce your understanding
- Ask questions to your instructor when something is unclear

The solutions provided are detailed and show the complete thought process, but they should serve as a guide and verification tool, not a substitute for your own problem-solving efforts.

## Looking Ahead

The mathematics you encounter in this course is not merely *abstract theory*—it forms the foundation for advanced work in *physics*, *engineering*, *computer science*, *economics*, and countless other fields. From the elliptical orbits of satellites to the encryption protecting digital communications, from statistical analysis of scientific data to the rendering of 3D computer graphics, these mathematical tools are actively shaping our modern world.

We hope this textbook serves as both a rigorous introduction to these powerful ideas and an invitation to see mathematics as a lens through which to understand the elegant patterns underlying our universe.

---

*Lee Sihoo* & *Lee Hojun*

*Korea Science Academy of KAIST*

# Table of Contents

## Chapter 14 *Probabilities & Statistics*

Chapter 08

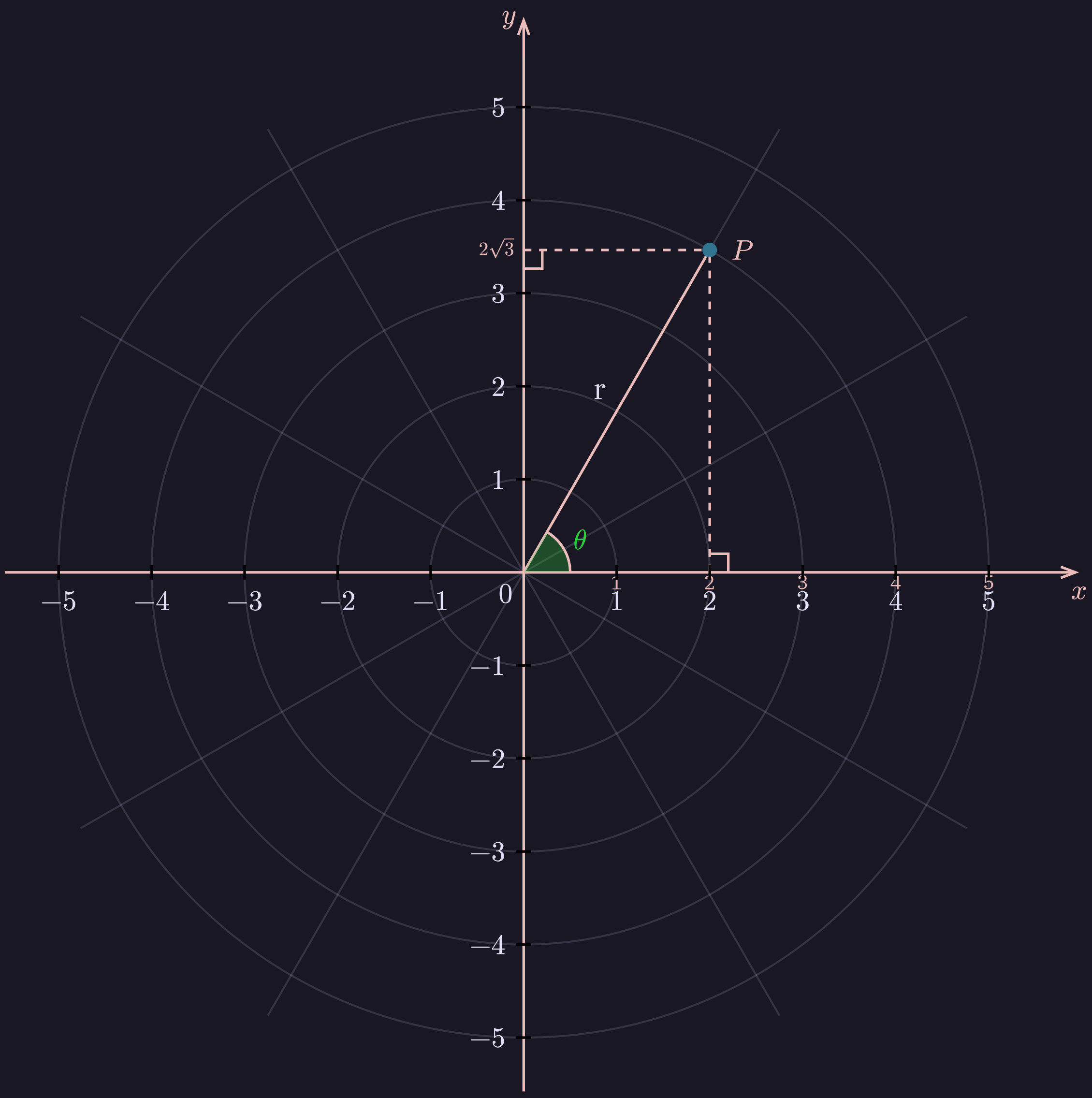
# Polar Coordinates, Parametric Equations & Vectors

*Until now, our study of functions has been confined to the static grid of the Rectangular coordinate system. However, the dynamic physical universe—filled with spiraling galaxies and directional forces—rarely conforms to such rigid constraints. In this chapter, we expand our mathematical lexicon to include Polar Coordinates, Parametric Equations, and Vectors, freeing us from the limitations of Cartesian graphs. By redefining position through distance and angle and introducing time as a driving parameter, we gain the sophisticated framework necessary to analyze motion, force, and form in a complex, multi-dimensional world.*

Article 08.01

# *Polar Coordinates*

## Polar Coordinates




The cartesian coordinate of dot $P$ is $\left(2, 2\sqrt{3}\right)$.
In polar coordinate system, the coordinate of dot $P$ is $\left(4, \frac{\pi}{3}\right)$.

**Definition | Polar Coordinates**

In this section, we will study the polar coordinate system (**극좌표계**). The polar coordinate system uses distances and directions to specify the location of a point in the plane.

To set up this system, we choose a fixed point $O$ in the plane called the **pole** (or origin, **극점**, **원점**) and draw from $O$ a half-line called the **polar axis** (**극축**). Then each point $P$ can be assigned polar coordinates $P(r, \theta)$, where

- $r$ is the distance from $O$ to $P$,
- $\theta$ is the angle between the polar axis and the segment $OP$.

We use the convention that $\theta$ is positive if measured in a counterclockwise direction from the polar axis or negative if measured in a clockwise direction. If $r$ is negative, then $P(r, \theta)$ is defined to be the point that lies $|r|$ units from the pole in the direction opposite to that given by $\theta$.

### Example | Plotting Points

Plot the points whose polar coordinates are given.
(a) $(1, \frac{3}{4}\pi)$
(b) $(3, -\frac{\pi}{6})$
(c) $(3, 3\pi)$
(d) $(-4, \frac{\pi}{4})$

#### Solution 1 |

(a) Point at distance 1, angle $135^\circ$.
(b) Point at distance 3, angle $-30^\circ$.
(c) Point at distance 3, angle $540^\circ \equiv 180^\circ$.
(d) Point at distance 4, angle $45^\circ + 180^\circ = 225^\circ$ (opposite direction).

### Example | Exercise

Explain why any polar coordinate $P(r, \theta)$ represents the same point as

$$P(r, \theta + 2n\pi) \quad \text{and} \quad P(-r, \theta + (2n+1)\pi), \quad n \in \mathbb{Z}.$$

#### Solution 1 |

**First representation:** $P(r, \theta + 2n\pi)$

Since angles are measured from the polar axis, adding $2n\pi$ (where $n \in \mathbb{Z}$) represents a complete rotation(s) around the origin. After completing full rotation(s), we return to the same direction. The distance $r$ from the pole remains unchanged, so the point is identical to $P(r, \theta)$.

**Second representation:** $P(-r, \theta + (2n+1)\pi)$

The term $(2n+1)\pi$ represents rotation by an odd multiple of $\pi$, which means rotating by $\pi, 3\pi, 5\pi, \ldots$ radians. Each of these rotations places us in the **opposite direction** from $\theta$.

When $r$ is negative, we move $|r|$ units in the direction **opposite** to the angle. So $P(-r, \theta + (2n+1)\pi)$ means:
- Start at angle $\theta + (2n+1)\pi$ (which is opposite to $\theta$)
- Move $|r|$ units in the opposite direction of that angle
- This brings us to the direction of $\theta$ at distance $|r| = r$

Therefore, both representations describe the same point as $P(r, \theta)$.

### Example | Finding other representations

Find two other polar coordinate representations of $P\left(2, \frac{\pi}{3}\right)$ with $r > 0$ and two with $r < 0$.

#### Solution 1 |

$r > 0$:

$$\ldots, \left(2, -11\frac{\pi}{3}\right), \left(2, -5\frac{\pi}{3}\right), \left(2, 7\frac{\pi}{3}\right), \left(2, 13\frac{\pi}{3}\right), \ldots$$

$r < 0$:

$$\ldots, \left(-2, -8\frac{\pi}{3}\right), \left(-2, -2\frac{\pi}{3}\right), \left(-2, 4\frac{\pi}{3}\right), \left(-2, 10\frac{\pi}{3}\right), \ldots$$

## Relationship Between Polar and Rectangular Coordinates

1. To change from polar to rectangular coordinates, use the formulas
$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$
2. To change from rectangular to polar coordinates, use the formulas
$$r^2 = x^2 + y^2, \quad \tan\theta = \frac{y}{x}(x \neq 0)$$
or
$$r^2 = x^2 + y^2, \quad \cos\theta = \frac{x}{r} \text{ and } \sin\theta = \frac{y}{r}.$$

## Example | Coordinate Conversion

(a) Find rectangular coordinates for the point that has polar coordinates $(4, 2\frac{\pi}{3})$.
(b) Find polar coordinates for the point that has rectangular coordinates $(2, -2)$.

### Solution 1 |

(a)

$$x = 4\cos\left(2\frac{\pi}{3}\right) = 4 \cdot \left(-\frac{1}{2}\right) = -2$$

$$y = 4\sin\left(2\frac{\pi}{3}\right) = 4 \cdot \left(\frac{\sqrt{3}}{2}\right) = 2\sqrt{3}$$

$$\therefore \left(-2, 2\sqrt{3}\right)$$

(b)

$$r = \pm\sqrt{2^2 + (-2)^2} = \pm 2\sqrt{2}$$

$$\tan\theta = -\frac{2}{2} = -1$$

$$\theta = 3\frac{\pi}{4} + n\pi, \left(-\frac{\pi}{4}\right) + n\pi, \quad n \in \mathbb{Z}$$

$$\therefore \left(2\sqrt{2}, -\frac{\pi}{4} + 2n\pi\right), \left(-2\sqrt{2}, 3\frac{\pi}{4} + 2n\pi\right), \quad n \in \mathbb{Z}$$

## Note | Remarks

1. In most cases, we assume $r > 0$ and $0 \leq \theta < 2\pi$ in polar coordinates. Under these conditions, each point corresponds uniquely to a single polar coordinate, establishing a one-to-one correspondence with Cartesian coordinates. This convention is often preferred.
2. The origin $(0, 0)$ corresponds to all polar coordinates of the form $(0, \theta)$ for any angle $\theta$. However, it is generally preferred not to consider the origin explicitly in polar coordinates.
3. Allowing $r \leq 0$ is mainly useful when sketching graphs in polar coordinates, as it simplifies the representation of certain curves.

## Polar Equations

**Definition | Polar Equation**

A **polar equation** (tr. **극방정식**) is an equation in the polar coordinates $r$ and $\theta$; similarly, a rectangular equation is an equation in the rectangular coordinates $x$ and $y$.

**Example | Converting to Polar Equation**

Express the equation $x^2 = 4y$ in polar coordinates.

**Solution 1 |**

(Case 1) $r \neq 0$ Since $r^2 = x^2 + y^2 = 4y + y^2$ and $y \geq 0$, we have $y \neq 0$ and therefore $x \neq 0$, which implies $\cos\theta \neq 0$. Then we get

$$x^2 = 4y$$

$$(r\cos\theta)^2 = 4r\sin\theta$$

$$r\cos^2\theta = 4\sin\theta$$

$$r = \frac{4\sin\theta}{\cos^2\theta} = 4\sec\theta\tan\theta$$

(Case 2) $r = 0$ This corresponds to the origin $(0, 0)$, which also satisfies the equation $x^2 = 4y$. However, substituting $\theta = 0$ into the expression derived in Case 1 yields $r = 0$, which satisfies the condition in Case 2 as well. Therefore, we conclude that

$$r = 4\sec\theta\tan\theta.$$

**Example | Converting to Rectangular Equation**

Express each polar equation in rectangular coordinates. (a) $r = 5\sec\theta$ (b) $r = 2\sin\theta$ (c) $r = 2 + 2\cos\theta$

**Solution 1 |**

(a) A vertical line.

$$r = 5 \sec\theta \Rightarrow r\cos\theta = 5 \Rightarrow x = 5$$

(b) A circle.

$$r = 2\sin\theta \Rightarrow r^2 = 2r\sin\theta \Rightarrow x^2 + y^2 = 2y \Rightarrow x^2 + (y-1)^2 = 1$$

(c) A “Cardioid”.

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

$$\left(x^2 + y^2 - 2x\right)^2 = 4r^2 = 4(x^2 + y^2)$$

### Example | Distance Formula

Prove that the distance between the polar points $(r_1, \theta_1)$ and $(r_2, \theta_2)$ is

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_2 - \theta_1)}.$$

**Solution 1 |**

The corresponding points in rectangular coordinates for $(r_1, \theta_1)$ and $(r_2, \theta_2)$ are $(r_1\cos\theta_1, r_1\sin\theta_1)$ and $(r_2\cos\theta_2, r_2\sin\theta_2)$, respectively. Then the distance between them is

$$\begin{aligned} d &= \sqrt{\left(r_2\cos\theta_2 - r_1\cos\theta_1\right)^2 + \left(r_2\sin\theta_2 - r_1\sin\theta_1\right)^2} \\ &= \sqrt{r_1^2 + r_2^2 - 2r_1 r_2(\cos\theta_2\cos\theta_1 + \sin\theta_2\sin\theta_1)} \\ &= \sqrt{r_1^2 + r_2^2 - 2r_1 r_2\cos(\theta_2 - \theta_1)}. \end{aligned}$$

### Example | Exercise

(a) Use the Law of Cosines to prove the formula above.
(b) Find the distance between the points whose polar coordinates are $(3, 3\frac{\pi}{4})$ and $(-1, 7\frac{\pi}{6})$.

**Solution 1 |**

(a) Consider the triangle formed by the origin $O$ and the two points $P_1(r_1, \theta_1)$ and $P_2(r_2, \theta_2)$.

The three sides of this triangle have lengths:
- From $O$ to $P_1$: length $r_1$
- From $O$ to $P_2$: length $r_2$
- From $P_1$ to $P_2$: length $d$ (what we want to find)

The angle at the origin between the two radii is $\theta_2 - \theta_1$.

By the Law of Cosines:

$$d^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_2 - \theta_1)$$

Taking the square root:

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_2 - \theta_1)}$$

(b) Given: $P_1(3, 3\frac{\pi}{4})$ and $P_2(-1, 7\frac{\pi}{6})$

First, note that $P_2(-1, 7\frac{\pi}{6})$ with negative $r$ is equivalent to $P_2(1, 7\frac{\pi}{6} + \pi) = P_2(1, 13\frac{\pi}{6})$.

For simplicity, we can also write this as $P_2(1, \frac{\pi}{6})$ (since $13\frac{\pi}{6} - 2\pi = \frac{\pi}{6}$).

Using the distance formula:

$$d = \sqrt{3^2 + 1^2 - 2(3)(1)\cos\left(\frac{\pi}{6} - 3\frac{\pi}{4}\right)} \backslash = \sqrt{9 + 1 - 6\cos\left(-7\frac{\pi}{12}\right)} \backslash = \sqrt{10 - 6\cos\left(7\frac{\pi}{12}\right)}$$

Since $\cos(7\frac{\pi}{12}) = \cos(105^\circ) = -\sin(15^\circ) = -\frac{\sqrt{6}-\sqrt{2}}{4}$:

$$d = \sqrt{10 - 6 \cdot \left(-\frac{\sqrt{6} - \sqrt{2}}{4}\right)} \backslash = \sqrt{10 + \frac{3\left(\sqrt{6} - \sqrt{2}\right)}{2}} \backslash = \sqrt{10 + \frac{3\sqrt{6} - 3\sqrt{2}}{2}}$$

Numerically: $d \approx \sqrt{10 + 1.421} \approx \sqrt{11.421} \approx 3.38$

Article 08.02

# *Graphs of Polar Equations*

## Graphs of Polar Equations

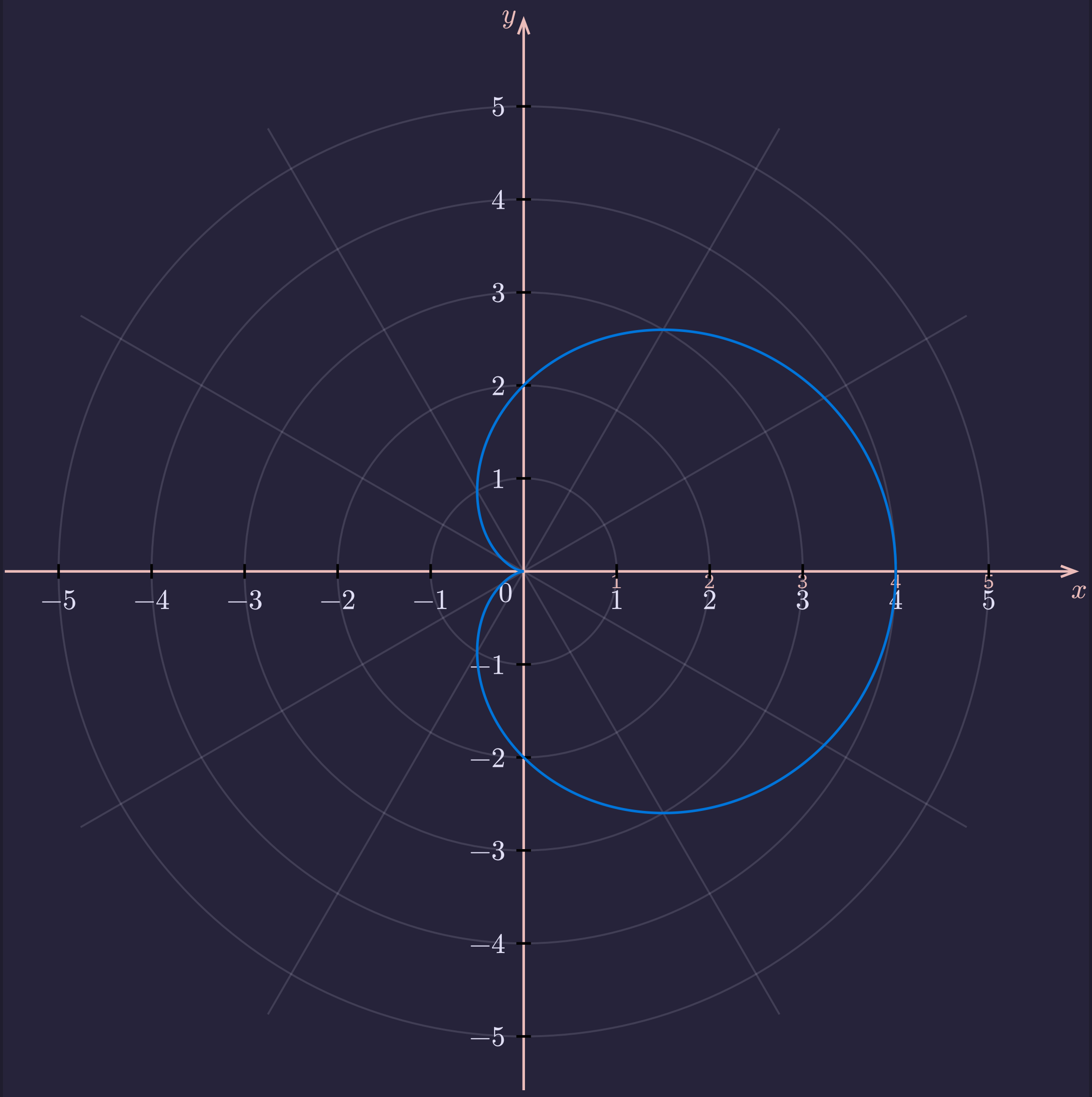
**DEFINITION | Graphing Polar Equations**

In this section, we learn how to sketch the graph of a polar equation $r = f(\theta)$. To plot points in polar coordinates, it is convenient to use a grid consisting of circles centered at the pole and rays emanating from the pole.

**EXAMPLE | A Cardioid**

Sketch a graph of $r = 2 + 2\cos\theta$.

**Solution 1 |**




In general, the graph of any equation of the form

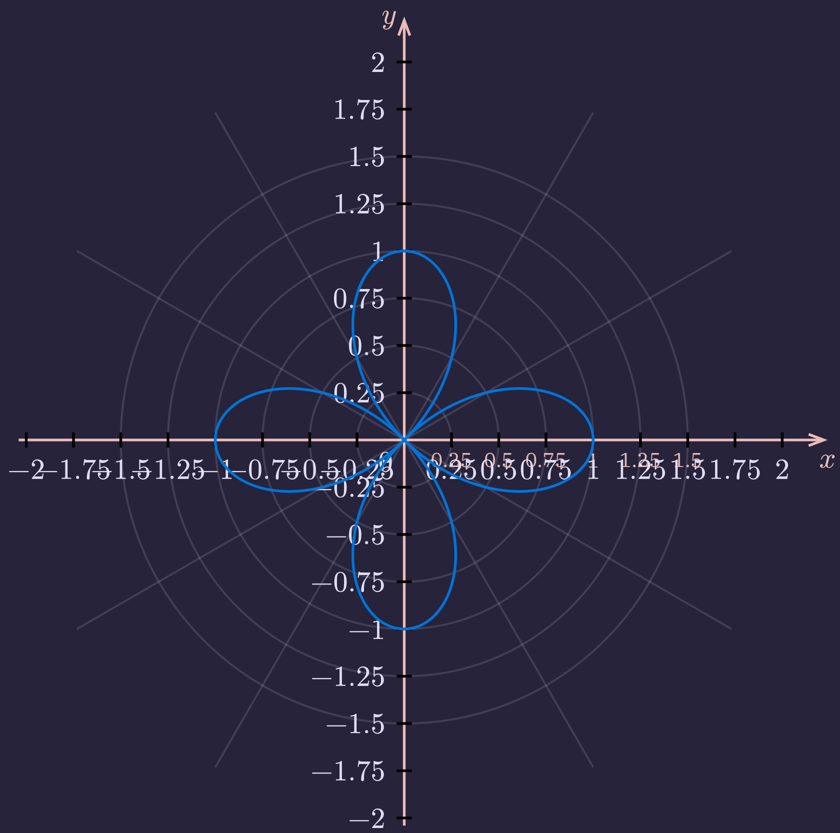
$$r = a(1 \pm \cos\theta) \quad \text{or} \quad r = a(1 \pm \sin\theta)$$

is a cardioid.

**Example | A Four-Leaved Rose**

Sketch the curve $r = \cos 2\theta$.

**Solution 1 |**




In general, the rose curves of the form

$$r = a\sin n\theta \quad \text{or} \quad r = a\cos n\theta$$

have $n$ petals if $n$ is odd, and $2n$ petals if $n$ is even.

### Example | Simple cases

Sketch the graphs of the following equations. (a) $r = 3$ (b) $\theta = \frac{\pi}{3}$ (c) $r = 2\sin\theta$

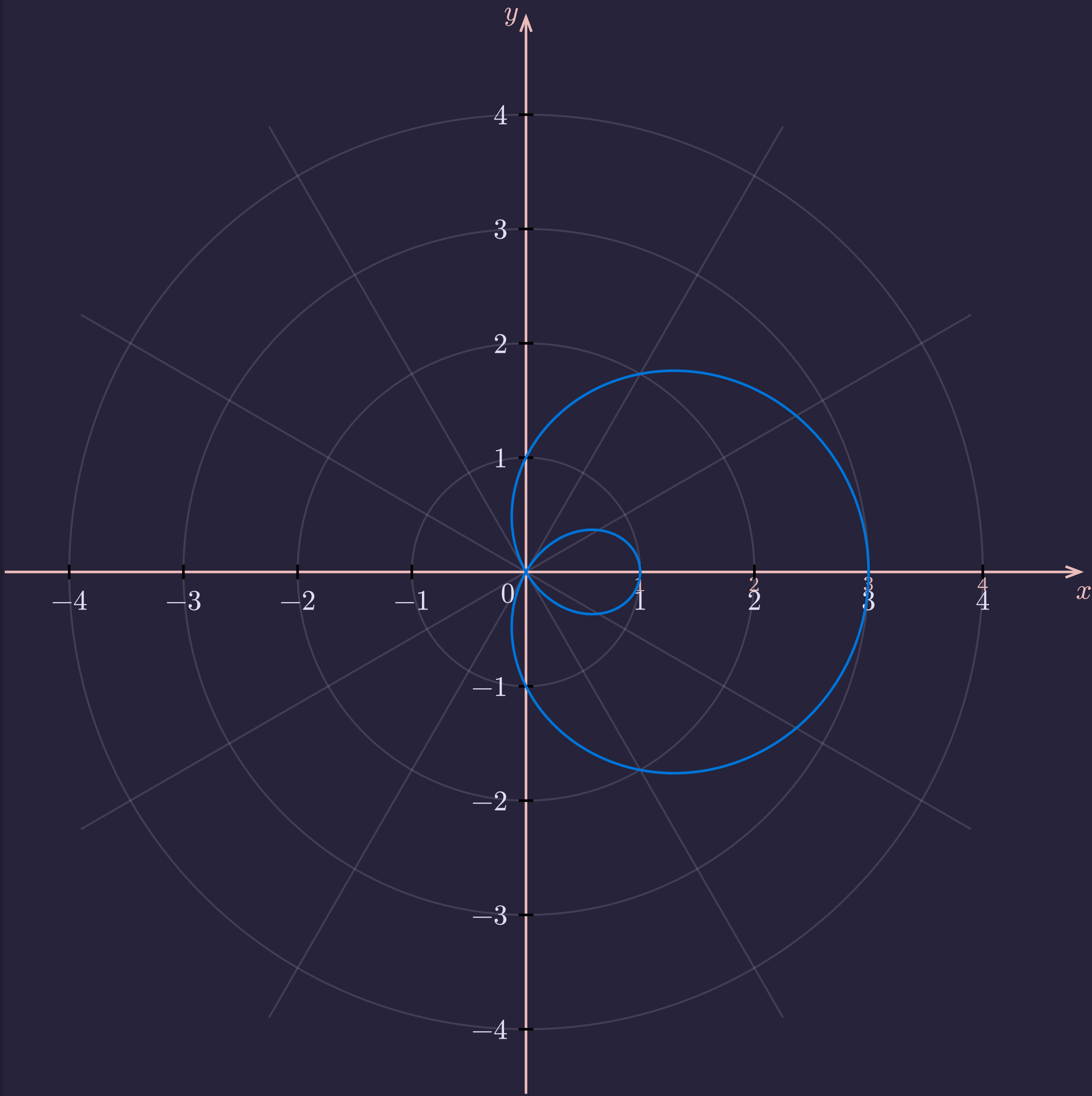
**Solution 1 |**

(a) $x^2 + y^2 = r^2 = 9$ (Circle radius 3) (b) $\frac{y}{x} = \tan\theta = \tan\left(\frac{\pi}{3}\right) = \sqrt{3} \Leftrightarrow y = \sqrt{3}x$ (Line) (c) $x^2 + (y-1)^2 = 1$ (Circle shifted up)

### Example | A Limacon

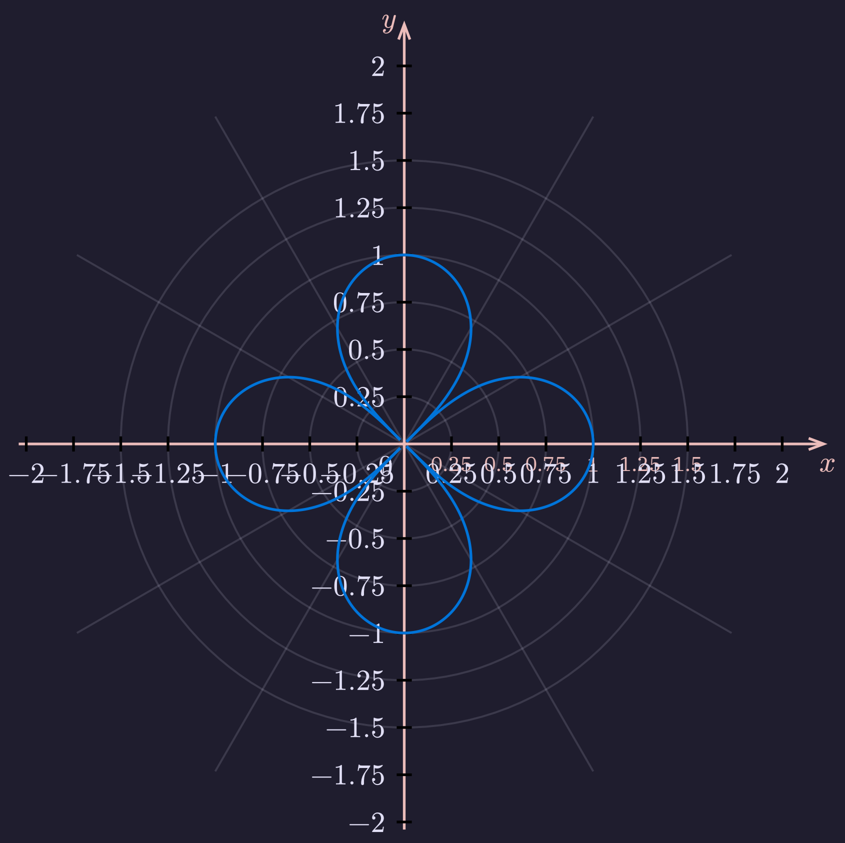
Sketch a graph of the equation $r = 1 + 2\cos\theta$.

**Solution 1 |**




**Example | Exercise (Lemniscate)**

Sketch a graph of the equation $r^2 = \cos 2\theta$.



## Symmetry

In graphing a polar equation $r = f(\theta)$, it's often helpful to take advantage of symmetry. We list three tests for symmetry.

| Symmetry | Test |
|---|---|
| With respect to the polar axis | The polar equation is unchanged if we replace $\theta$ by $-\theta$. |
| With respect to the pole | The polar equation is unchanged if we replace $r$ by $-r$ or $\theta$ by $\theta + \pi$. |
| With respect to the line $\theta = \frac{\pi}{2}$ | The polar equation is unchanged if we replace $\theta$ by $\pi - \theta$. |

Suppose $f(\theta) = f(-\theta)$ for all $\theta$. If a point $(r, \theta)$ lies on the graph, then the reflected point $(r, -\theta)$ also lie on the graph, since $r = f(\theta) = f(-\theta)$. Therefore, the graph of $r = f(\theta)$ is symmetric with respect to the polar axis.

Let's return to the previous example and consider $f(\theta) = 1 + 2\cos\theta$. Since $f(\theta) = f(-\theta)$ for all $\theta$, the graph of $r = 1 + 2\cos\theta$ is symmetric with respect to the polar axis. Therefore, it suffices to examine the graph for $\theta \in [0, \pi]$ to obtain the entire graph.

**Example | Exercise**

(a) Explain why the other two symmetry tests imply symmetry with respect to the pole and the line $\theta = \frac{\pi}{2}$, respectively. (b) Sketch the graphs of $r = \cos 2\theta$ and $r = 1 - \sin\theta$ using symmetry tests.

### Solution 1 |

(a) **Symmetry with respect to the pole:**

If the equation is unchanged when we replace $r$ by $-r$, then whenever $(r, \theta)$ is on the graph, so is $(-r, \theta)$. But $(-r, \theta)$ represents the same point as $(r, \theta + \pi)$, which is the reflection of $(r, \theta)$ through the origin (pole). Therefore, the graph is symmetric with respect to the pole.

Alternatively, if the equation is unchanged when we replace $\theta$ by $\theta + \pi$, then whenever $(r, \theta)$ is on the graph, so is $(r, \theta + \pi)$, which is the point rotated $180^{\circ}$ around the pole, giving symmetry with respect to the pole.

**Symmetry with respect to the line $\theta = \frac{\pi}{2}$:**

If the equation is unchanged when we replace $\theta$ by $\pi - \theta$, then whenever $(r, \theta)$ is on the graph, so is $(r, \pi - \theta)$. The angle $\pi - \theta$ is the reflection of $\theta$ across the line $\theta = \frac{\pi}{2}$ (the vertical axis). Therefore, the graph is symmetric with respect to the line $\theta = \frac{\pi}{2}$.

(b) **Graph of $r = \cos 2\theta$:**

Testing for symmetry:
- Polar axis: $f(-\theta) = \cos(-2\theta) = \cos(2\theta) = f(\theta)$ ✓ (symmetric)
- Pole: $f(\theta + \pi) = \cos(2\theta + 2\pi) = \cos(2\theta) = f(\theta)$ ✓ (symmetric)
- Line $\theta = \frac{\pi}{2}$: $f(\pi - \theta) = \cos(2\pi - 2\theta) = \cos(2\theta) = f(\theta)$ ✓ (symmetric)

Since it's symmetric about the polar axis, we only need to graph for $\theta \in [0, \pi]$. This is a rose curve with 4 petals (since $n = 2$ is even, giving $2n = 4$ petals).

**Graph of $r = 1 - \sin\theta$:**

Testing for symmetry:
- Polar axis: $f(-\theta) = 1 - \sin(-\theta) = 1 + \sin(\theta) \neq f(\theta)$ ×
- Pole: $f(\theta + \pi) = 1 - \sin(\theta + \pi) = 1 + \sin(\theta) \neq f(\theta)$ ×
- Line $\theta = \frac{\pi}{2}$: $f(\pi - \theta) = 1 - \sin(\pi - \theta) = 1 - \sin(\theta) = f(\theta)$ ✓ (symmetric)

Since it's symmetric about $\theta = \frac{\pi}{2}$, we can graph for $\theta \in [0, \pi]$ and reflect. This is a cardioid of the form $r = a(1 - \sin\theta)$ with the dimple pointing downward.

### Example | Intersection Points

Find all the intersection points of the curves given by the polar equations $r = \frac{1}{2}$ and $r = \cos 2\theta$.

**Solution 1 |**

(Case 1) $\cos 2\theta = \frac{1}{2}$

$$2\theta = (2n+1)\pi \pm 2\frac{\pi}{3}$$

$$n = 0, \cos\theta = \mp\frac{\sqrt{3}}{2}, \sin\theta = \frac{1}{2}$$

$$n = 1, \cos\theta = \pm\frac{\sqrt{3}}{2}, \sin\theta = -\frac{1}{2}$$

(Case 2) $\cos 2\theta = -\frac{1}{2}$

$$2\theta = (2n+1)\pi \pm \frac{\pi}{3}$$

$$n = 0, \cos\theta = \mp\frac{1}{2}, \sin\theta = \frac{\sqrt{3}}{2}$$

$$n = 1, \cos\theta = \pm\frac{1}{2}, \sin\theta = -\frac{\sqrt{3}}{2}$$

$$\therefore \frac{1}{2}\left(\pm\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \frac{1}{2}\left(\pm\frac{\sqrt{3}}{2}, -\frac{1}{2}\right), \frac{1}{2}\left(\pm\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \frac{1}{2}\left(\pm\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$$

Article 08.03

# *Polar Form of Complex Numbers; De Moivre's Theorem*

## Polar Form of Complex Numbers; De Moivre's Theorem

### Graphing Complex Numbers

We need two axes to graph complex numbers: one for the real part and one for the imaginary part. We call these the **real axis** (tr. **실수축**) and the **imaginary axis** (tr. **허수축**), respectively. The plane determined by these two axes is called the **complex plane** (tr. **복소평면**). To graph the complex number $a + bi$, we plot the ordered pair of numbers $(a, b)$ in this plane.

Recall that the absolute value of a real number can be thought of as its distance from the origin on the real number line. We define absolute value for complex numbers in a similar fashion. The **modulus** (or absolute value, **절댓값**) of the complex number $z = a + bi$ is

$$|z| = \sqrt{a^2 + b^2}.$$

**Example | Moduli**

Find the moduli of the complex numbers $3 + 4i$ and $8 - 5i$.

**Solution 1 |**

The modulus of a complex number $z = a + bi$ is defined as $|z| = \sqrt{a^2 + b^2}$.

For $3 + 4i$:

$$|3 + 4i| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5.$$

For $8 - 5i$:

$$|8 - 5i| = \sqrt{8^2 + (-5)^2} = \sqrt{64 + 25} = \sqrt{89}.$$

The set $C = \{z \in \mathbb{C} : |z| = 1\}$ is called the **unit circle**, and $D = \{z \in \mathbb{C} : |z| \leq 1\}$ is called the **(closed) unit disk**.

## Euler’s Formula

**Theorem | Euler's Formula**

$$e^{i\theta} = \cos\theta + i\sin\theta, \quad \theta \in \mathbb{R}$$

In particular, we have

$$e^{i\pi} + 1 = 0.$$

Euler’s formula is often called the most beautiful formula in the world. It brings together five of the most fundamental constants in mathematics—$e, i, \pi, 1$, and $0$—in a single equation.

**Note | Motivation**

It is well-known that

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

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

holds true for all $x \in \mathbb{R}$. If $e^x$ is well-defined for all $x \in \mathbb{C}$, then

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

## Polar Form of Complex Numbers

A complex number $z = a + bi$ has the **polar form** (or trigonometric form, tr. **극형식**)

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta},$$

where $r = \sqrt{a^2 + b^2}$ and $\tan\theta = \frac{b}{a}$ $(a \neq 0)$. The number $r$ is the modulus of $z$, and $\theta$ is called an **argument** (tr. **편각**) of $z$.

**Note | Remark**

The argument of $z$ is not unique, but any two arguments of $z$ differ by a multiple of $2\pi$.

**Example | Polar Form**

Write each complex number in polar form. (a) $1+i$ (b) $-4\sqrt{3}-4i$

**Solution 1 |**

(a) $1+i=\sqrt{2}(\cos(\frac{\pi}{4})+i\sin(\frac{\pi}{4}))$ (b) $-4\sqrt{3}-4i=8\left(\cos(7\frac{\pi}{6})+i\sin(7\frac{\pi}{6})\right)$

**Example | Properties of the Modulus**

Verify the property for the complex numbers $\omega$ and $z$. (a) $|\overline{z}|=|z|$ (b) $z\overline{z}=|z|^2$ (c) $|\omega z|=|\omega||z|$ (d) $|\frac{1}{z}|=\frac{1}{|z|}$ $(z\neq 0)$ (e) $|\frac{\omega}{z}|=\frac{|\omega|}{|z|}$ $(z\neq 0)$

**Solution 1 |**

(c) Let $\omega=a+bi$ and $z=c+di$. Then

$$\begin{aligned}|\omega z| &= |(ac-bd)+(ad+bc)i| \\ &= \sqrt{(ac-bd)^2+(ad+bc)^2} \\ &= \sqrt{a^2c^2+b^2d^2+a^2d^2+b^2c^2} \\ &= \sqrt{(a^2+b^2)(c^2+d^2)}.\end{aligned}$$

## Triangle Inequality

**Theorem | Triangle Inequality**

Let $z_1, z_2\in\mathbb{C}$. Then

$$|z_1+z_2|\leq|z_1|+|z_2|.$$

**Example | Exercise**

Let $z = \Re(z) + i\Im(z)$. Prove the following. (a) $\Re(z), \Im(z) \leq |z|$ (b) $z + \overline{z} = 2\Re(z)$ (c) $z - \overline{z} = 2i\Im(z)$

**Solution 1 |**

(a) Let $z = a + bi$ where $a = \Re(z), b = \Im(z)$. Then

$$|z| = \sqrt{a^2 + b^2} \geq \sqrt{a^2} = |a| \geq a = \Re(z)$$

Similarly, $|z| \geq |b| \geq b = \Im(z)$.

(b) Let $z = a + bi$. Then $\overline{z} = a - bi$, so

$$z + \overline{z} = (a + bi) + (a - bi) = 2a = 2\Re(z).$$

(c) $z - \overline{z} = (a + bi) - (a - bi) = 2bi = 2i\Im(z)$.

**Proof |**

It is enough to show that $|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$.

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} \\ &= |z_1|^2 + z_1\overline{z_2} + z_2\overline{z_1} + |z_2|^2 \\ &= |z_1|^2 + 2\Re(z_1\overline{z_2}) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1\overline{z_2}| + |z_2|^2 \\ &= (|z_1| + |z_2|)^2. \end{aligned}$$

**Example | Exercise**

Show that $|\sum_{i=1}^{n} z_i| \leq \sum_{i=1}^{n} |z_i|$. (Hint: Use induction.)

**Solution 1 |**

**Base case:** For $n = 1$, $|z_1| \leq |z_1|$ is trivially true.

**Inductive step:** Assume the statement holds for $n = k$:

$$|\sum_{i=1}^{k} z_i| \le \sum_{i=1}^{k} |z_i|.$$

For $n = k + 1$:

$$\begin{aligned}
|\sum_{i=1}^{k+1} z_i| &= |\sum_{i=1}^{k} z_i + z_{k+1}| \\
&\le |\sum_{i=1}^{k} z_i| + |z_{k+1}| \quad \text{(by triangle inequality)} \\
&\le \sum_{i=1}^{k} |z_i| + |z_{k+1}| \quad \text{(by inductive hypothesis)} \\
&= \sum_{i=1}^{k+1} |z_i|.
\end{aligned}$$

By mathematical induction, the statement holds for all $n \ge 1$.

**Note | Reverse Triangle Inequality**

$$||z_1| - |z_2|| \le |z_1 + z_2| \le |z_1| + |z_2|.$$

**Example | Zero Location**

Show that every zero $z_0$ of the polynomial

$$P(z) = z^4 - 2iz^3 + (1 + 3i)z^2 + 3z - 9$$

satisfies $|z_0| > 1$.

**Solution 1 |**

Suppose $|z_0| \le 1$. Then

$$\begin{aligned}
|P(z_0)| &\ge |-9| - |z_0^4 - 2iz_0^3 + (1 + 3i)z_0^2 + 3z_0| \\
&\ge 9 - (|z_0|^4 + 2|z_0|^3 + 2|z_0|^2 + 3|z_0|) \\
&\ge 9 - (1 + 2 + 2 + 3) = 1.
\end{aligned}$$

This contradicts $P(z_0) = 0$.

## Definition of the Exponential Function

We now define the exponential function for complex numbers. For $z = x + yi$, where $x, y \in \mathbb{R}$, we set

$$e^z := e^{x+yi} = e^x e^{yi} = e^x(\cos y + i \sin y).$$

## Properties of the Exponential Function

**Theorem | Properties**

For $z, z_1, z_2 \in \mathbb{C}$ and $n \in \mathbb{Z}$, the following properties hold:

1. $e^z \neq 0$
2. $e^{z_1+z_2} = e^{z_1} e^{z_2}$
3. $e^0 = 1, e^{-z} = \frac{1}{e^z}$
4. $(e^z)^n = e^{nz}$
5. (De Moivre's theorem) $\left(e^{i\theta}\right)^n = e^{in\theta}, \theta \in \mathbb{R}$
6. $\overline{e^z} = e^{\overline{z}}$
7. $e^z = 1 \Leftrightarrow z = 2n\pi i$

**Example | Exercise**

Prove that for $z_1, z_2 \in \mathbb{C}$ and $z_2 \neq 0$,

$$\frac{e^{z_1}}{e^{z_2}} = e^{z_1 - z_2}.$$

**Solution 1 |**

Using property 2 from the theorem (the exponential addition property):

$$e^{z_1 - z_2} \cdot e^{z_2} = e^{(z_1 - z_2) + z_2} = e^{z_1}.$$

Dividing both sides by $e^{z_2}$ (which is nonzero by property 1):

$$e^{z_1 - z_2} = \frac{e^{z_1}}{e^{z_2}}.$$

**Example | Complex Multiplication**

Let $z_1 = 2(\cos(\frac{\pi}{4}) + i \sin(\frac{\pi}{4}))$ and $z_2 = 5\left(\cos\left(\frac{\pi}{3}\right) + i \sin\left(\frac{\pi}{3}\right)\right)$. Find (a) $z_1 z_2$ and (b) $\frac{z_1}{z_2}$.

**Solution 1 |**

$z_1 = 2e^{\pi \frac{i}{4}}$, $z_2 = 5e^{\pi \frac{i}{3}}$. (a) $z_1 z_2 = 10e^{7\pi \frac{i}{12}} = 10(\cos(7\frac{\pi}{12}) + i \sin(7\frac{\pi}{12}))$ (b) $\frac{z_1}{z_2} = \frac{2}{5}e^{-\pi \frac{i}{12}} = \frac{2}{5}(\cos(\frac{\pi}{12}) - i \sin(\frac{\pi}{12}))$

## De Moivre's Theorem

**Theorem | De Moivre's Theorem**

For $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$,

$$z^n = r^n(\cos n\theta + i \sin n\theta) = r^n e^{in\theta}, \quad n \in \mathbb{Z}.$$

**Example | Calculation**

Find $\left(\frac{1}{2} + \frac{1}{2}i\right)^{10}$.

**Solution 1 |**

First, convert to polar form:

$$r = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}}$$

$$\theta = \arctan(1) = \frac{\pi}{4}$$

So $\frac{1}{2} + \frac{1}{2}i = \frac{1}{\sqrt{2}}e^{\pi \frac{i}{4}}$.

By De Moivre's theorem:

$$\begin{aligned}\left(\frac{1}{2}+\frac{1}{2}i\right)^{10} &= \left(\frac{1}{\sqrt{2}}\right)^{10} e^{10\pi\frac{i}{4}} \\ &= \frac{1}{32}e^{5\pi\frac{i}{2}} \\ &= \frac{1}{32}e^{\pi\frac{i}{2}} \quad (\text{since } e^{2\pi i}=1) \\ &= \frac{1}{32}\left(\cos\left(\frac{\pi}{2}\right)+i\sin\left(\frac{\pi}{2}\right)\right) \\ &= \frac{1}{32}i.\end{aligned}$$

## $n$th Root of Complex Numbers

Now we consider solving an equation of the form $z^n = c$. Let $z = Re^{i\omega}$ and $c = re^{i\theta}$. Then $R^n e^{in\omega} = re^{i\theta}$.

$$R = r^{\frac{1}{n}} \quad \text{and} \quad \omega = \frac{\theta + 2k\pi}{n}, \quad k \in \mathbb{Z}.$$

Thus, the solutions are given by

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

**Example | Roots**

Find the six sixth roots of $\omega = -64$, and graph these roots in the complex plane.

**Solution 1 |**

We want to solve $z^6 = -64$. First, write $-64$ in polar form:

$$-64 = 64(\cos\pi + i\sin\pi) = 64e^{\pi i}.$$

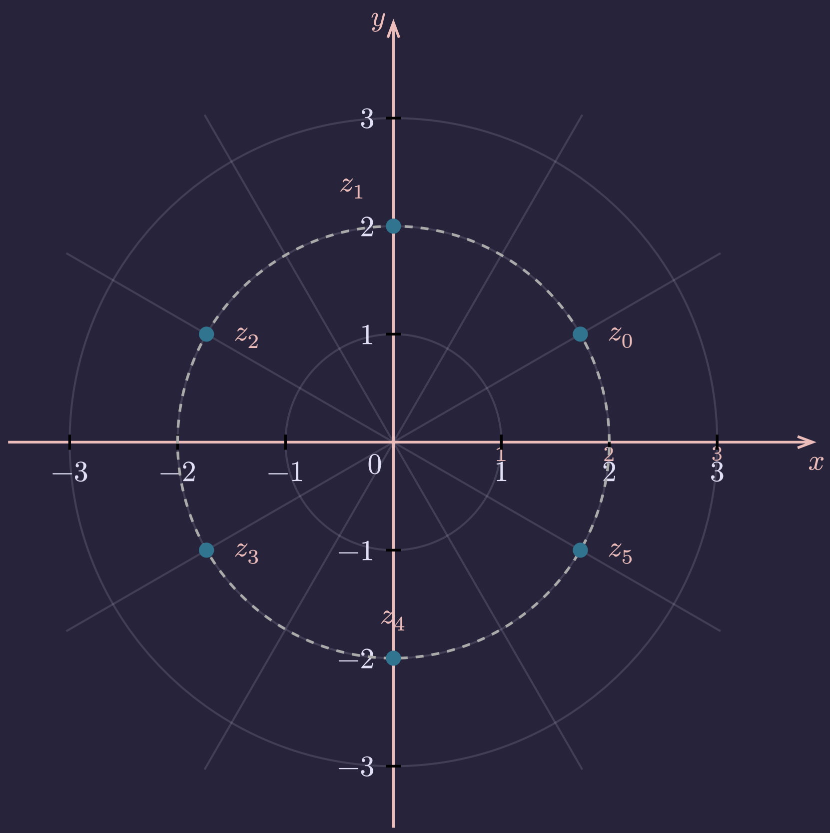
Let $z = re^{i\theta}$. Then $z^6 = r^6 e^{i6\theta}$. Equating the moduli and arguments:

$$r^6 = 64 \Rightarrow r = 2,$$

$$6\theta = \pi + 2k\pi \Rightarrow \theta_k = \frac{\pi + 2k\pi}{6}, \quad k = 0, 1, 2, 3, 4, 5.$$

The six roots are:

$$
\begin{aligned}
z_0 &= 2e^{i\frac{\pi}{6}} = 2\left(\cos\left(\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{6}\right)\right) = \sqrt{3} + i \\
z_1 &= 2e^{i3\frac{\pi}{6}} = 2e^{i\frac{\pi}{2}} = 2i \\
z_2 &= 2e^{i5\frac{\pi}{6}} = 2\left(\cos\left(5\frac{\pi}{6}\right) + i\sin\left(5\frac{\pi}{6}\right)\right) = -\sqrt{3} + i \\
z_3 &= 2e^{i7\frac{\pi}{6}} = 2\left(\cos\left(7\frac{\pi}{6}\right) + i\sin\left(7\frac{\pi}{6}\right)\right) = -\sqrt{3} - i \\
z_4 &= 2e^{i9\frac{\pi}{6}} = 2e^{i3\frac{\pi}{2}} = -2i \\
z_5 &= 2e^{i11\frac{\pi}{6}} = 2\left(\cos\left(11\frac{\pi}{6}\right) + i\sin\left(11\frac{\pi}{6}\right)\right) = \sqrt{3} - i
\end{aligned}
$$




## Complex Trigonometric Functions

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

**Example | Exercise**

How are $\sinh z$ and $\cosh z$ defined? Justify your answer.

**Solution 1 |**

By analogy with the definitions of $\cos z$ and $\sin z$:

$$\cosh z = \frac{e^z + e^{-z}}{2}, \quad \sinh z = \frac{e^z - e^{-z}}{2}.$$

These definitions are justified because they reduce to the standard hyperbolic functions when $z \in \mathbb{R}$, and they satisfy the same differential equations and identities as the real hyperbolic functions. For example:

- $\cosh^2 z - \sinh^2 z = 1$
- $\frac{d}{dz} \cosh z = \sinh z$ and $\frac{d}{dz} \sinh z = \cosh z$

**Example | Zeros**

Show that $\cos z = 0 \Leftrightarrow z = (2k+1)\frac{\pi}{2}$ for some $k \in \mathbb{Z}$.

**Solution 1 |**

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} = 0$$
$$\Leftrightarrow e^{iz} + e^{-iz} = 0$$
$$\Leftrightarrow e^{iz} = -e^{-iz}$$
$$\Leftrightarrow e^{2iz} = -1 = e^{\pi i}$$
$$\Leftrightarrow 2iz = (2k+1)\pi i \quad \text{for some} \quad k \in \mathbb{Z}$$
$$\Leftrightarrow z = (2k+1)\frac{\pi}{2}.$$

## Argument and the Principal Value

$\arg z$ is any argument. $\operatorname{Arg} z$ is the unique value $\theta$ such that $-\pi < \theta \leq \pi$.

**Example | Principal Value**

Find the principal argument $\operatorname{Arg}(z)$ for each complex number. (a) $1+i$ (b) $-5i$ (c) $-1+\sqrt{3}i$ (d) $-1$

**Solution 1 |**

Recall that $\operatorname{Arg}(z)$ is the unique angle $\theta$ such that $z = |z|e^{i\theta}$ and $-\pi < \theta \le \pi$.

(a) $z = 1 + i$. This point is in Quadrant I.

$$\tan\theta = \frac{1}{1} = 1 \Rightarrow \theta = \frac{\pi}{4}.$$

Since $\frac{\pi}{4} \in (-\pi, \pi]$, $\operatorname{Arg}(1+i) = \frac{\pi}{4}$.

(b) $z = -5i$. This point lies on the negative imaginary axis. The angle is $-\frac{\pi}{2}$.

$$\operatorname{Arg}(-5i) = -\frac{\pi}{2}.$$

(c) $z = -1 + \sqrt{3}i$. This point is in Quadrant II (since $x < 0, y > 0$).

$$\tan\alpha = |\frac{\sqrt{3}}{-1}| = \sqrt{3} \Rightarrow \alpha = \frac{\pi}{3}$$

(reference angle).

$$\theta = \pi - \frac{\pi}{3} = 2\frac{\pi}{3}.$$

$$\operatorname{Arg}\left(-1 + \sqrt{3}i\right) = 2\frac{\pi}{3}.$$

(d) $z = -1$. This point lies on the negative real axis. By definition, the argument is $\pi$.

$$\operatorname{Arg}(-1) = \pi.$$

## Complex Logarithm

$$\log z = \ln|z| + i \arg z, \quad z \neq 0.$$

$$\operatorname{Log} z = \ln|z| + i \operatorname{Arg} z.$$

**Example | Logarithm**

Find all values of (a) $\log(1 - i)$ (b) $\log(3 + 2i)$ (c) $\operatorname{Log}\left(-1 + \sqrt{3}i\right)$.

**Solution 1 |**

(a) $|1 - i| = \sqrt{2}$, $\arg(1 - i) = -\frac{\pi}{4} + 2k\pi$

$$\log(1-i) = \ln\left(\sqrt{2}\right) + i\left(-\frac{\pi}{4} + 2k\pi\right) = \frac{1}{2}\ln 2 + i\left(-\frac{\pi}{4} + 2k\pi\right), \quad k \in \mathbb{Z}$$

(b) $|3+2i| = \sqrt{13}$, $\arg(3+2i) = \arctan\left(\frac{2}{3}\right) + 2k\pi$

$$\log(3+2i) = \ln\left(\sqrt{13}\right) + i\left(\arctan\left(\frac{2}{3}\right) + 2k\pi\right), \quad k \in \mathbb{Z}$$

(c) $|-1+\sqrt{3}i| = 2$, $\operatorname{Arg}\left(-1+\sqrt{3}i\right) = 2\frac{\pi}{3}$

$$\operatorname{Log}\left(-1+\sqrt{3}i\right) = \ln 2 + i\left(2\frac{\pi}{3}\right)$$

## Complex Exponents

$$z^\alpha = e^{\alpha \log z}.$$

**Example | Exponents**

Find all values for (a) $i^{\frac{1}{2}}$ (b) $i^i$ (c) $5^{\frac{1}{2}}$.

**Solution 1 |**

(a) $i^{\frac{1}{2}} = e^{\frac{1}{2}\log i} = \pm e^{\pi\frac{i}{4}}$ (b) $i^i = e^{i \log i} = e^{-\left(\frac{\pi}{2}+2k\pi\right)}$ (c) $5^{\frac{1}{2}} = \pm\sqrt{5}$

Article 08.04

# *Plane Curves and Parametric Equations*

## Plane Curves and Parametric Equations

We can think of a curve as the path of a point moving in the plane; the $x$- and $y$-coordinates of the point are then functions of time $t$. Motivated by this idea, we define a **plane curve (평면곡선)** as the set of points $(f(t), g(t))$ where $f$ and $g$ are functions defined on an interval $I$. The equations

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

where $t \in I$, are **parametric equations (매개(변수)방정식)** for the curve, with **parameter (매개변수)** $t$.

**Example | Sketching**

Sketch the curve defined by the parametric equations

$$x = t^2 - 3t, \quad y = t - 1.$$

**Solution 1 |**

We can eliminate the parameter $t = y + 1$.

$$x = (y+1)^2 - 3(y+1) = y^2 - y - 2.$$

This is a parabola.

**Example | Moving Object**

The following parametric equations model the position of a moving object at time $t$:

$$x = \cos t, \quad y = \sin t, \quad t \geq 0.$$

Describe and graph the path.

**Solution 1 |**

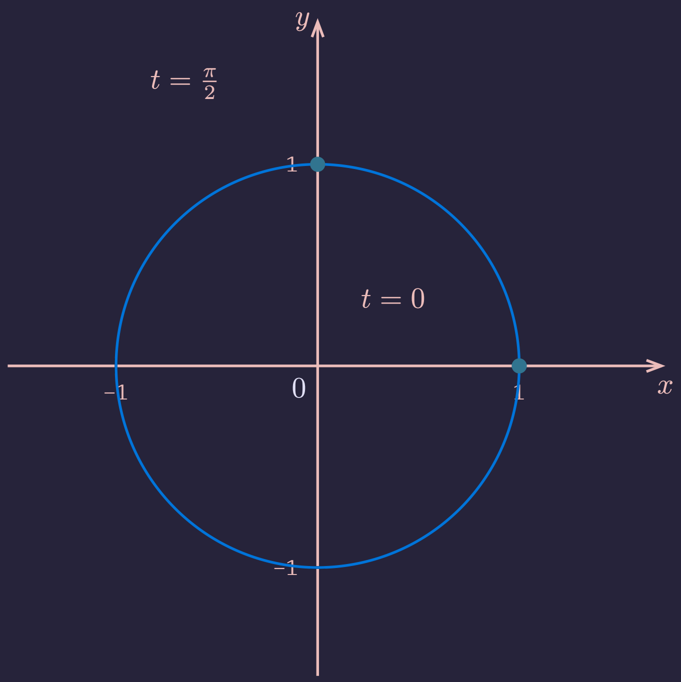
First, eliminate the parameter $t$ to find the Cartesian equation. Recall the identity $\cos^2 t + \sin^2 t = 1$. Substituting $x$ and $y$:

$$x^2 + y^2 = \cos^2 t + \sin^2 t = 1.$$

This is a circle of radius 1 centered at the origin $(0, 0)$.

To determine the direction of motion:
- At $t = 0$: $(x, y) = (\cos 0, \sin 0) = (1, 0)$.
- At $t = \frac{\pi}{2}$: $(x, y) = (\cos(\frac{\pi}{2}), \sin(\frac{\pi}{2})) = (0, 1)$.

Since the object moves from $(1, 0)$ to $(0, 1)$, it moves in a **counterclockwise** direction.




## Example | Eliminating Parameter

Eliminate the parameter, and sketch the graph of

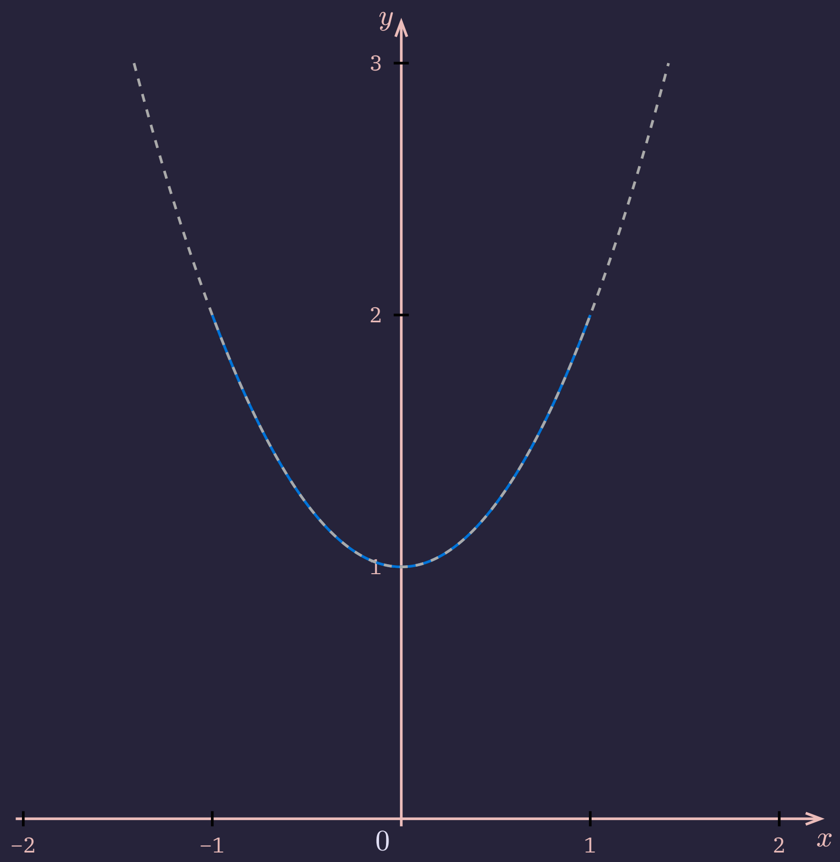
$$x = \sin t, \quad y = 2 - \cos^2 t.$$

**Solution 1 |**

We use the identity $\cos^2 t = 1 - \sin^2 t$. Substitute $\sin t = x$ into the equation for $y$:

$$y = 2 - (1 - \sin^2 t) = 2 - (1 - x^2) = 1 + x^2.$$

Now consider the domain. Since $x = \sin t$ and $-1 \leq \sin t \leq 1$, we must have $-1 \leq x \leq 1$. Thus, the graph is the segment of the parabola $y = x^2 + 1$ defined on the interval $[-1, 1]$.




## Finding Parametric Equations for a Curve

**Example | Line**

Find parametric equations for the line $y = 3x + 2$.

**Solution 1 |**

We can choose the parameter $t$ arbitrarily. The simplest choice is to let $x = t$. Then substituting into the equation gives $y = 3t + 2$. Thus, a set of parametric equations is:

$$x = t, \quad y = 3t + 2, \quad t \in \mathbb{R}.$$

(Note: Other choices are possible, e.g., $x = t - 1$, which would give $y = 3(t - 1) + 2 = 3t - 1$.)

**Example | Asteroid**

Find parametric equations for $x^{\frac{2}{3}} + y^{\frac{2}{3}} = 1$.

**Solution 1 |**

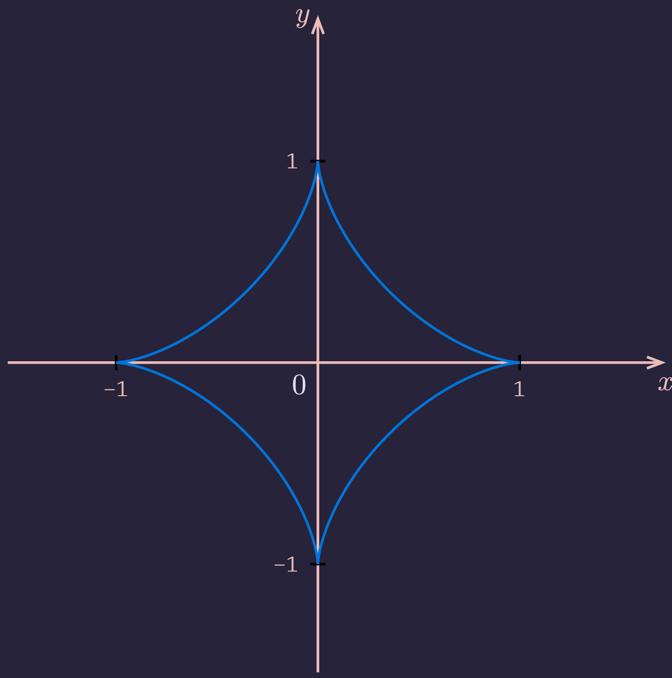
We want to use the Pythagorean identity $\cos^2\theta + \sin^2\theta = 1$. Let $x^{\frac{2}{3}} = \cos^2\theta$ and $y^{\frac{2}{3}} = \sin^2\theta$. Solving for $x$ and $y$:

$$\left(x^{\frac{2}{3}}\right)^{\frac{3}{2}} = \left(\cos^2\theta\right)^{\frac{3}{2}} \Rightarrow x = \cos^3\theta$$

$$\left(y^{\frac{2}{3}}\right)^{\frac{3}{2}} = \left(\sin^2\theta\right)^{\frac{3}{2}} \Rightarrow y = \sin^3\theta$$

So the parametric equations are:

$$x = \cos^3\theta, \quad y = \sin^3\theta, \quad 0 \leq \theta \leq 2\pi.$$




## Cycloid

As a circle rolls along a straight line, the curve traced out by a fixed point $P$ on the circumference of the circle is called a **cycloid**.

### Example | Cycloid Equation

If the circle has radius $r$ and rolls along the x-axis, with one position of the point $P$ being at the origin, find parametric equations for the cycloid.

#### Solution 1 |

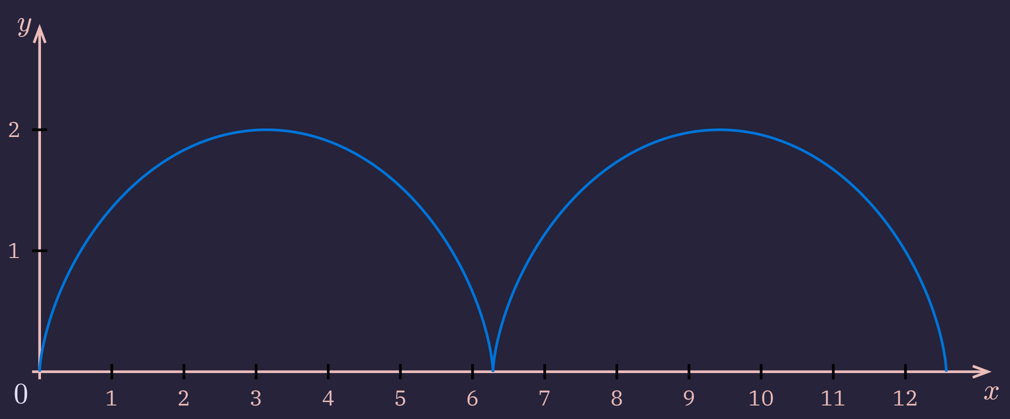
Let $\theta$ be the angle of rotation of the circle (in radians). Suppose the circle starts with $P$ at the origin $(0,0)$. When the circle has rotated by angle $\theta$, its center $C$ has moved horizontally by the arc length $r\theta$. So the coordinates of the center are $C(r\theta, r)$.

The point $P$ is at a distance $r$ from $C$. Relative to the center $C$, the position of $P$ is given by the angle $3\frac{\pi}{2} - \theta$ (since it starts at the bottom $3\frac{\pi}{2}$ and rotates clockwise by $\theta$). Vector $(\textit{CP}) = (r\cos(3\frac{\pi}{2} - \theta), r\sin(3\frac{\pi}{2} - \theta)) = (r(-\sin\theta), r(-\cos\theta))$.

The position of $P$ is $(\textit{OP}) = (\textit{OC}) + (\textit{CP})$:

$$x = r\theta - r\sin\theta = r(\theta - \sin\theta)$$

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




**Example | Track Problem**

A track consists of two semicircles of radius 1 and a rectangle of side lengths $\pi$ and 2. A circle of radius 1/2 rolls around the track. Find a parametrization of the trajectory of $P$.

**Solution 1 |**

The track can be divided into four segments:
1. **Bottom straight** ($0 \le t \le \pi$): The point $P$ moves along a line at height $\frac{1}{2}$.
2. **Right semicircle** ($\pi \le t \le 2\pi$): The circle rolls around the outer semicircle of radius 1.
3. **Top straight** ($2\pi \le t \le 3\pi$): The point moves along a line at height $\frac{3}{2}$.
4. **Left semicircle** ($3\pi \le t \le 4\pi$): The circle rolls around the outer semicircle.

The parametrization depends on which segment $t$ corresponds to. For the bottom straight section:

$$x(t) = t - \frac{1}{2}\sin(2t), \quad y(t) = \frac{1}{2}(1 - \cos(2t)), \quad 0 \le t \le \pi.$$

Similar cycloid-based expressions apply to other segments, with appropriate translations and rotations.

**Example | Polar to Parametric**

Consider the polar equation $r = \theta, 1 \leq \theta \leq 10\pi$. (a) Express the equation in parametric form. (b) Draw a graph.

**Solution 1 |**

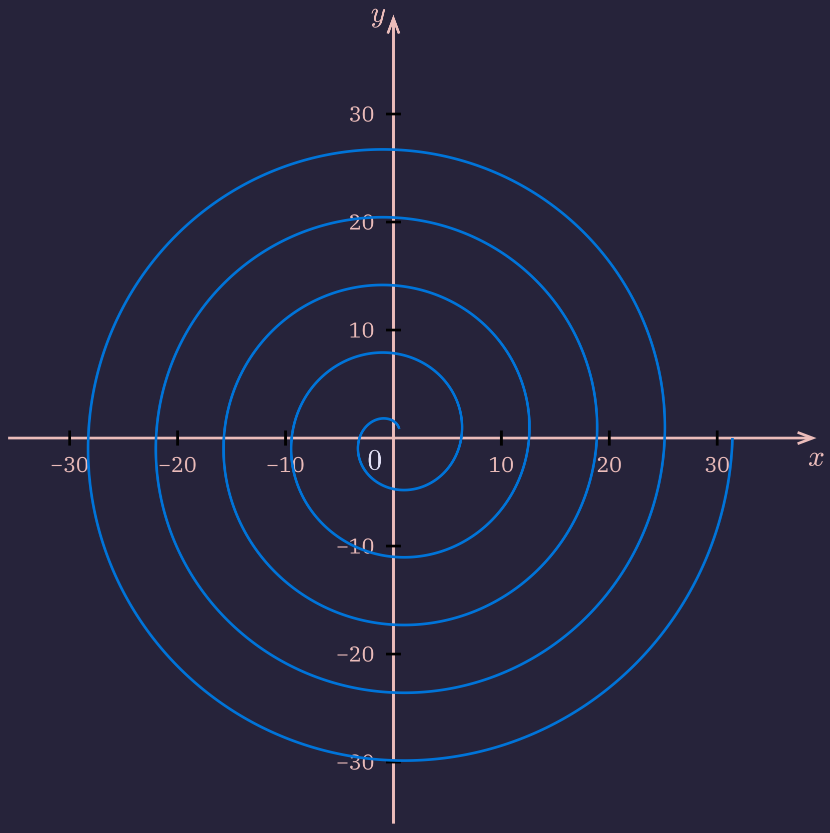
(a) The relationship between polar and Cartesian coordinates is $x = r\cos\theta$ and $y = r\sin\theta$. Substituting $r = \theta$:

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

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

where $1 \leq \theta \leq 10\pi$.

(b) This is a spiral.



Article 08.05

# Vectors

## Vectors

### Geometric Description of Vectors

A **vector** (**벡터**) in the plane is a line segment with an assigned direction. We sketch a vector with an arrow to specify the direction. We denote this vector by $(AB)$. Point $A$ is the **initial point** (**시점**), and $B$ is the **terminal point** (**종점**). The length of the line segment $AB$ is called the **magnitude** or **length** of the vector and is denoted by $|(AB)|$. We use boldface letters to denote vectors, e.g., $u = (AB)$. The zero vector, denoted by $0$, represents no displacement.

Two vectors are considered equal if they have equal magnitude and the same direction.

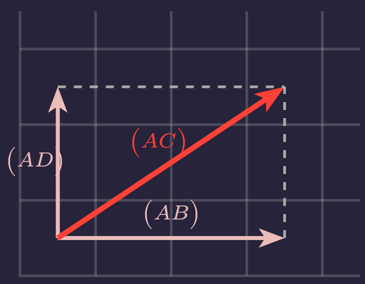
### Vector Arithmetic

- **Addition**: $(AC) = (AB) + (BC)$.
- **Scalar Multiplication**: $cv$ has magnitude $|c\|v|$ and direction same as $v$ if $c > 0$, opposite if $c < 0$.
- **Difference**: $u - v = u + (-v)$.

**Example | Parallelogram**

Let $ABCD$ be a parallelogram. Describe the following vectors using $A, B, C, D$. (a) $(AB) + (AD)$ (b) $(AB) - (BC)$ (c) $(AC) - (DC)$

**Solution 1 |**




(a) $(AC)$ (b) $(AB) - (AD) = (DB)$ (c) $(AC) + (CD) = (AD)$

### Properties of Vectors

| Vector Addition | Scalar Multiplication |
| --- | --- |

| | |
|---|---|
| $(u+v)+w=u+(v+w)$ | $c(u+v)=cu+cv$ |
| $u+v=v+u$ | $(c+d)u=cu+du$ |
| $u+0=u$ | $(cd)u=c(du)$ |
| $u+(-u)=0$ | $1u=u$ |

### Example | Exercise

Prove that $0u=0$ and $c0=0$.

#### Solution 1 |

**Part 1:** Prove $0u=0$:

$$0u=(0+0)u=0u+0u$$

Adding $-(0u)$ to both sides:

$$0=0u$$

**Part 2:** Prove $c0=0$:

$$c0=c(0+0)=c0+c0$$

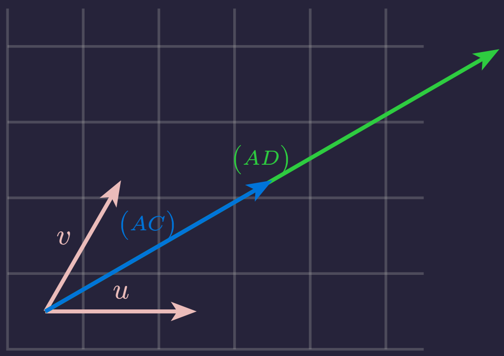
Adding $-(c0)$ to both sides:

$$0=c0$$

### Example | Hexagon

Let $ABCDEF$ be a regular hexagon, and let $u=(\overrightarrow{AB})$ and $v=(\overrightarrow{AF})$. Describe $(\overrightarrow{AC}),(\overrightarrow{AD}),(\overrightarrow{FC})$ in terms of $u,v$.

#### Solution 1 |




$(FC) = 2u$ $(AC) = 2u + v$ $(AD) = 2u + 2v$

## Vectors in the Coordinate Plane

We can represent $v$ as an ordered pair of real numbers, $v = \langle a_1, a_2 \rangle$. $v = \langle x_2 - x_1, y_2 - y_1 \rangle$ if initial point is $(x_1, y_1)$ and terminal is $(x_2, y_2)$.

### Example | Component Form

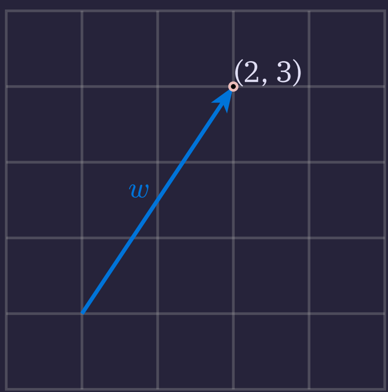
(a) Find component form of $u$ with initial $(-2, 5)$ and terminal $(3, 7)$. (b) Sketch $w = \langle 2, 3 \rangle$.

#### Solution 1 |

(a) The component form is found by subtracting the initial point coordinates from the terminal point coordinates:

$$u = \langle 3 - (-2), 7 - 5 \rangle = \langle 5, 2 \rangle$$

(b) To sketch $w = \langle 2, 3 \rangle$, we draw a vector from the origin $(0, 0)$ to the point $(2, 3)$.



### Example | Calculation

If $u = \langle 2, -3 \rangle$ and $v = \langle -1, 2 \rangle$, find $2u - 3v$.

**Solution 1 |**

$2\langle 2, -3 \rangle - 3\langle -1, 2 \rangle = \langle 4 + 3, -6 - 6 \rangle = \langle 7, -12 \rangle$

### Example | Linear Combination

In the given figure, find $a, b$ where $(OC) = a(OA) + b(OB)$. $(OA) = \langle -3, 1 \rangle$, $(OB) = \langle 1, 2 \rangle$, $(OC) = \langle -1, 5 \rangle$.

**Solution 1 |**

$\langle -1, 5 \rangle = a\langle -3, 1 \rangle + b\langle 1, 2 \rangle = \langle -3a + b, a + 2b \rangle$ $a = 1, b = 2$.

Article 08.06

# *The Dot & Cross Product*

## The Dot & Cross Product

### The Dot Product of Vectors

**Definition | Dot Product**

If $u = \langle a_1, a_2 \rangle$ and $v = \langle b_1, b_2 \rangle$ are vectors, then their dot product (**내적**), denoted by $u \cdot v$, is defined by

$$u \cdot v = a_1 b_1 + a_2 b_2.$$

**Example | Calculation**

Let $u = \langle 3, -2 \rangle$ and $v = \langle 4, 5 \rangle$. Find $u \cdot v$.

**Solution 1 |**

Using the definition of the dot product:

$$u \cdot v = (3)(4) + (-2)(5) = 12 - 10 = 2.$$

**Theorem | Properties of the Dot Product**

Let $u, v, w$ be vectors and $c$ be a scalar.
1. $u \cdot v = v \cdot u$
2. $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$
3. $(u + v) \cdot w = u \cdot w + v \cdot w$
4. $|u|^2 = u \cdot u$

**Example | Exercise**

Prove parts 1, 2 and 3 of the properties of the dot product.

**Solution 1 |**

Let $u = \langle u_1, u_2 \rangle, v = \langle v_1, v_2 \rangle, w = \langle w_1, w_2 \rangle$ and $c$ be a scalar.

**Property 1:** $u \cdot v = v \cdot u$

$$u \cdot v = u_1 v_1 + u_2 v_2 = v_1 u_1 + v_2 u_2 = v \cdot u.$$

**Property 2:** $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$

$$(cu) \cdot v = \langle cu_1, cu_2 \rangle \cdot \langle v_1, v_2 \rangle = cu_1 v_1 + cu_2 v_2 = c(u_1 v_1 + u_2 v_2) = c(u \cdot v).$$

Similarly, $u \cdot (cv) = u_1(cv_1) + u_2(cv_2) = c(u_1 v_1 + u_2 v_2) = c(u \cdot v)$.

**Property 3:** $(u + v) \cdot w = u \cdot w + v \cdot w$

$$\begin{aligned} (u + v) \cdot w &= \langle u_1 + v_1, u_2 + v_2 \rangle \cdot \langle w_1, w_2 \rangle \\ &= (u_1 + v_1) w_1 + (u_2 + v_2) w_2 \\ &= u_1 w_1 + v_1 w_1 + u_2 w_2 + v_2 w_2 \\ &= (u_1 w_1 + u_2 w_2) + (v_1 w_1 + v_2 w_2) \\ &= u \cdot w + v \cdot w. \end{aligned}$$

## Example | Identity

Show that $\|a \pm b\|^2 = \|a\|^2 + \|b\|^2 \pm 2a \cdot b$.

**Solution 1 |**

Using the property $|u|^2 = u \cdot u$ and the distributive laws:

$$\begin{aligned} \|a \pm b\|^2 &= (a \pm b) \cdot (a \pm b) \\ &= a \cdot a \pm a \cdot b \pm b \cdot a + b \cdot b \\ &= \|a\|^2 \pm 2(a \cdot b) + \|b\|^2 \\ &= \|a\|^2 + \|b\|^2 \pm 2a \cdot b. \end{aligned}$$

## Theorem | The Dot Product Theorem

If $\theta$ is the angle between two nonzero vectors $u$ and $v$, then

$$u \cdot v = |u\|v| \cos \theta.$$

### Example | Calculation

Let $\|b\| = 1$ and $\|a - 3b\| = \sqrt{13}$. If the angle between $a$ and $b$ is $60^{\circ}$, find $\|a\|$.

#### Solution 1 |

$\|a - 3b\|^2 = 13 \Rightarrow \|a\|^2 - 6a \cdot b + 9\|b\|^2 = 13$ $\|a\|^2 - 6\|a\| \cos 60^{\circ} + 9 = 13$ $\|a\|^2 - 3\|a\| - 4 = 0 \Rightarrow \|a\| = 4$.

### Definition | Orthogonal

Two nonzero vectors $u$ and $v$ are called **perpendicular**, or **orthogonal** (**수직, 직교**), if the angle between them is $\frac{\pi}{2}$, which implies $u \cdot v = 0$.

### Example | Angle Calculation

Find $u \cdot v$ and $\cos\theta$. (a) $u = \langle 1, 1\rangle, v = \langle 1, -2\rangle$ (b) $u = \langle 3, 1\rangle, v = \langle 1, -1\rangle$

#### Solution 1 |

The cosine of the angle is given by $\cos\theta = \frac{u \cdot v}{|u||v|}$.

(a)

$$u \cdot v = 1(1) + 1(-2) = -1$$

$$|u| = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad |v| = \sqrt{1^2 + (-2)^2} = \sqrt{5}$$

$$\cos\theta = \frac{-1}{\sqrt{2}\sqrt{5}} = -\frac{1}{\sqrt{10}}.$$

(b)

$$u \cdot v = 3(1) + 1(-1) = 2$$

$$|u| = \sqrt{3^2 + 1^2} = \sqrt{10}, \quad |v| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

$$\cos\theta = \frac{2}{\sqrt{10}\sqrt{2}} = \frac{2}{\sqrt{20}} = \frac{2}{2\sqrt{5}} = \frac{1}{\sqrt{5}}.$$

**Example | Recursive Sequence**

Let $a, b$ be two perpendicular unit vectors. Define $b_n$ recursively. Find $\cos\theta$ where $\theta = \sum \theta_n$.

**Solution 1 |**

$\cos\theta = 0$.

## The Component of u along v

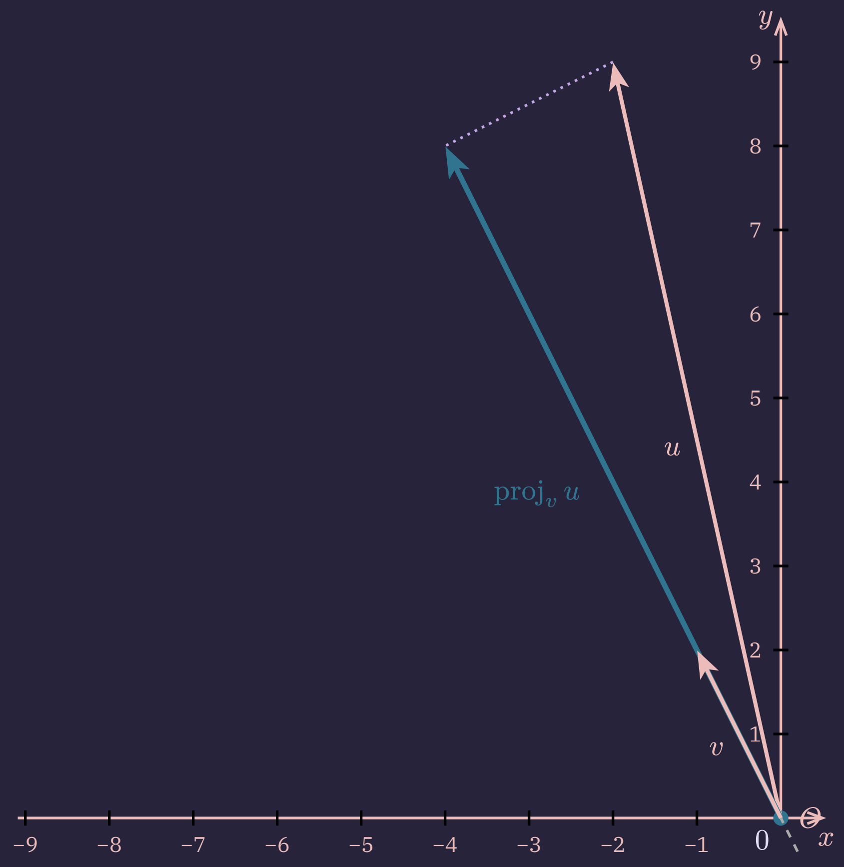
$$\operatorname{comp}_v u = |u| \cos\theta = \frac{u \cdot v}{|} v|$$

$$\operatorname{proj}_v u = \left(\frac{u \cdot v}{v \cdot v}\right) v$$

**Example | Projection**

Let $u = \langle -2, 9 \rangle$ and $v = \langle -1, 2 \rangle$. Find $\operatorname{proj}_v u$.

**Solution 1 |**




$\text{proj}_v\, u = \frac{20}{5}\langle -1, 2\rangle = \langle -4, 8\rangle.$

### Example | Work

A child pulls a wagon with force 100 N at 60° for 100 m. Find the work done.

#### Solution 1 |

The work done $W$ by a constant force $F$ moving an object through a displacement $D$ is given by $W = F \cdot D = |F\|D| \cos\theta$. Given: $|F| = 100$ N, $|D| = 100$ m, $\theta = 60^\circ$.

$$W = 100 \cdot 100 \cdot \cos(60^\circ) = 10000 \cdot \frac{1}{2} = 5000 \text{ J (Joules)}.$$

### Example | Normalize

Normalize (a) $\langle 1, -2, 2\rangle$ (b) $\langle 4, 1, -2\rangle$.

**Solution 1 |**

(a) $\frac{1}{3}\langle 1, -2, 2\rangle$ (b) $\frac{1}{\sqrt{21}}\langle 4, 1, -2\rangle$

**Example | Internal Division**

Let $A$ and $B$ be two points with position vectors $(OA) = a$ and $(OB) = b$. Find the position vector $p$ of the point $P$ that divides the segment $AB$ internally in the ratio $m : n$.

**Solution 1 |**

Since $P$ divides $AB$ in the ratio $m : n$, we have

$$(AP) : (PB) = m : n \Rightarrow n(AP) = m(PB).$$

Expressing these in terms of position vectors:

$$n(p - a) = m(b - p)$$

$$np - na = mb - mp$$

$$(m + n)p = mb + na$$

$$p = \frac{mb + na}{m + n}.$$

Article 08.07

# *Equations of Lines Planes*

## Equations of Lines Planes

### Equations of Lines in 3D

To specify a line, you need a point and a direction vector.

$$x = p + tv$$

**Definition | Parametric Equations of a Line**

The parametric equations of a line $l$ passing through $P(x_0, y_0, z_0)$ with direction $v = \langle v_1, v_2, v_3 \rangle$ are

$$\langle x, y, z \rangle = \langle x_0 + tv_1, y_0 + tv_2, z_0 + tv_3 \rangle, \quad t \in \mathbb{R}.$$

**Example | Line Equation**

Find parametric equations for line through $P(1, 2, 4)$ and $Q(3, 1, -1)$.

**Solution 1 |**

$(PQ) = \langle 2, -1, -5 \rangle$. $\langle x, y, z \rangle = \langle 1 + 2t, 2 - t, 4 - 5t \rangle$.

Cartesian Equation (Symmetric Equations):

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}$$

### Equations of Planes in 3D

A plane is determined by a point and a **normal vector** $n$.

$$n \cdot (x - p) = 0 \Rightarrow a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

**Example | Plane Equation**

Find Cartesian equation of plane through $P(1,0,1), Q(0,-1,4), R(2,0,0)$.

**Solution 1 |**

$(PQ) = \langle -1,-1,3\rangle, (PR) = \langle 1,0,-1\rangle$. $n = (PQ) \times (PR) = \langle 1,2,1\rangle$. $1(x-1) + 2(y-0) + 1(z-1) = 0 \Rightarrow x + 2y + z = 2$.

**Example | Parametric Plane**

Find parametric equations for plane containing $A(1,0,0), B(2,1,2), C(0,1,3)$.

**Solution 1 |**

$\langle x,y,z\rangle = \langle 1,0,0\rangle + s\langle 1,1,2\rangle + t\langle -1,1,3\rangle$.

## Relative Positions

- Two lines: Intersect, Parallel, or Skew.
- Two planes: Intersect (line) or Parallel.
- Line and Plane: Coincide, Intersect (point), or Parallel.

**Example | Intersection**

Determine if the line $l$ intersects the plane $\pi : x + 2y + 4z = 1$. The line $l$ is given by the symmetric equations: $\frac{x+1}{2} = \frac{y+1}{2} = \frac{2-z}{4}$.

**Solution 1 |**

First, convert the symmetric equations of the line to parametric equations. Let $\frac{x+1}{2} = \frac{y+1}{2} = \frac{2-z}{4} = t$. Then:

$$x = 2t - 1$$

$$y = 2t - 1$$

$$z = 2 - 4t$$

Substitute these into the equation of the plane:

$$(2t-1) + 2(2t-1) + 4(2-4t) = 1$$

$$2t - 1 + 4t - 2 + 8 - 16t = 1$$
$$-10t + 5 = 1$$
$$-10t = -4 \Rightarrow t = \frac{2}{5}$$

Since we found a unique value for $t$, the line intersects the plane. To find the point of intersection, substitute $t = \frac{2}{5}$ back into the parametric equations:

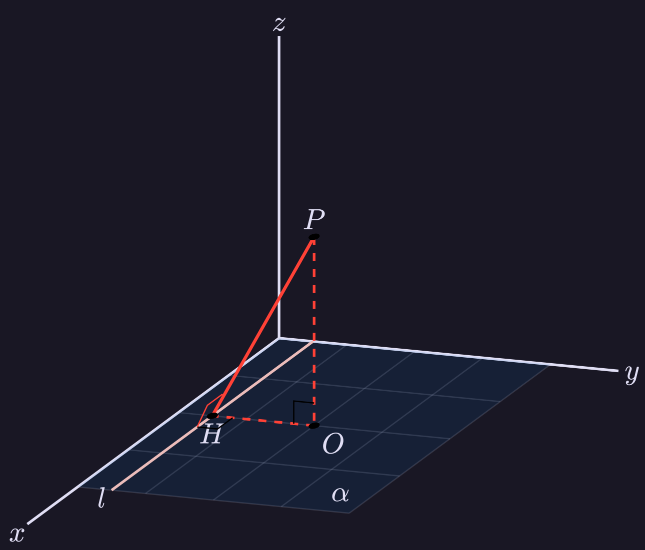
$$x = 2\left(\frac{2}{5}\right) - 1 = \frac{4}{5} - \frac{5}{5} = -\frac{1}{5}$$
$$y = 2\left(\frac{2}{5}\right) - 1 = \frac{4}{5} - \frac{5}{5} = -\frac{1}{5}$$
$$z = 2 - 4\left(\frac{2}{5}\right) = \frac{10}{5} - \frac{8}{5} = \frac{2}{5}$$

Thus, the intersection point is $(-\frac{1}{5}, -\frac{1}{5}, \frac{2}{5})$.

## Three Perpendiculars Theorem

**Theorem | Three Perpendiculars Theorem**

Let $\alpha$ be a plane, $l$ a line on $\alpha$. $H$ on $l$, $O$ on $\alpha$ not on $l$, $P$ not on $\alpha$.
1. If $PO \perp \alpha$ and $OH \perp l$, then $PH \perp l$.
2. If $PO \perp \alpha$ and $PH \perp l$, then $OH \perp l$.
3. If $PH \perp l, OH \perp l$, and $PO \perp OH$, then $PO \perp \alpha$.



**Example** | **Cuboid**

Consider a cuboid with vertices at coordinates $(x, y, z)$ where $x \in \{0, 2\}, y \in \{0, 3\}, z \in \{0, 4\}$. Let $A(0, 0, 0)$ and $G(2, 3, 4)$ be opposite vertices. Find the distance from vertex $A$ to the diagonal $BD$ on the face $z = 0$, where $B(2, 0, 0)$ and $D(0, 3, 0)$.

**Solution 1** |

The points are $A(0, 0, 0)$, $B(2, 0, 0)$, and $D(0, 3, 0)$. The line segment $BD$ lies in the $xy$-plane ($z = 0$). The vector $(\text{BD}) = D - B = \langle -2, 3, 0\rangle$. The line equation passing through $B$ and $D$ in the $xy$-plane is:

$$\frac{x-2}{-2} = \frac{y-0}{3} \Rightarrow 3x - 6 = -2y \Rightarrow 3x + 2y - 6 = 0.$$

The distance from $A(0, 0, 0)$ to the line $3x + 2y - 6 = 0$ (in 2D) is:

$$d = |3(0) + 2(0) - 6\frac{|}{\sqrt{3^2 + 2^2}} = | - 6\frac{|}{\sqrt{9+4}} = \frac{6}{\sqrt{13}}.$$

Alternatively, using the cross product area formula for triangle $ABD$: Area of $\triangle\, ABD = \frac{1}{2}\ |(\text{AB}) \times (\text{AD})|$. $(\text{AB}) = \langle 2, 0, 0\rangle$, $(\text{AD}) = \langle 0, 3, 0\rangle$. $(\text{AB}) \times (\text{AD}) = \langle 0, 0, 6\rangle$. Area $= \frac{1}{2}(6) = 3$. Base $BD = |(\text{BD})| = \sqrt{(-2)^2 + 3^2} = \sqrt{13}$. Height $h = \frac{2\times\ \text{Area}}{\text{Base}} = \frac{6}{\sqrt{13}}$.

## Projections in 3D

Projection of figure $F$ onto plane $\alpha$. Area of projected $S'$ is Area of $S$ times $\cos\theta$.

**Example** | **Tetrahedron**

Let $O(0, 0, 0)$, $A(3, 0, 0)$, $B(0, 3, 0)$, $C(0, 0, 3)$ be vertices of a tetrahedron. Find the area of the face $ABC$.

**Solution 1** |

The area of triangle $ABC$ is given by $\frac{1}{2}\ |(\text{AB}) \times (\text{AC})|$.

$$(AB) = B - A = \langle -3, 3, 0\rangle$$

$$(AC) = C - A = \langle -3, 0, 3\rangle$$

Calculate the cross product:

$$(AB) \times (AC) = \det\left(\begin{pmatrix} i & j & k \\ -3 & 3 & 0 \\ -3 & 0 & 3 \end{pmatrix}\right)$$
$$= i(9-0) - j(-9-0) + k(0-(-9))$$
$$= 9i + 9j + 9k = \langle 9, 9, 9 \rangle$$

The magnitude is:

$$|(AB) \times (AC)| = \sqrt{9^2 + 9^2 + 9^2} = \sqrt{3 \cdot 81} = 9\sqrt{3}.$$

Therefore, the area is:

$$\text{Area} = \frac{1}{2}\left(9\sqrt{3}\right) = \frac{9\sqrt{3}}{2}.$$

## Distance Formulas

- Point to Plane: $d = |ax_0 + by_0 + cz_0 - d\frac{|}{\sqrt{a^2+b^2+c^2}}$.
- Point to Line: $d = \|(AP) \times v\frac{\|}{\|}v\|$ (or using projection).
- Skew Lines: $d = |(p_1 - p_2) \cdot (m_1 \times m_2)\frac{|}{\|}m_1 \times m_2\|$.

### Example | Distance Point-Plane

Find the distance between the point $P(1, 2, 3)$ and the plane $x - 2y + 3z = 1$.

#### Solution 1 |

The distance $D$ from a point $P_0(x_0, y_0, z_0)$ to the plane $ax + by + cz + d = 0$ is given by:

$$D = |ax_0 + by_0 + cz_0 + d\frac{|}{\sqrt{a^2 + b^2 + c^2}}.$$

Rewrite the plane equation as $x - 2y + 3z - 1 = 0$. Here, $a = 1, b = -2, c = 3, d = -1$. $x_0 = 1, y_0 = 2, z_0 = 3$.

$$D = |1(1) + (-2)(2) + 3(3) - 1\frac{|}{\sqrt{1^2 + (-2)^2 + 3^2}}$$
$$= |1 - 4 + 9 - 1\frac{|}{\sqrt{1 + 4 + 9}}$$
$$= |5\frac{|}{\sqrt{14}} = \frac{5}{\sqrt{14}}.$$

## Example | Distance Skew Lines

Find the distance between the skew lines:

$$l_1 : x = 1 + t, y = t, z = 0$$

$$l_2 : x = 0, y = 1 + s, z = s$$

### Solution 1 |

Identify a point and direction vector for each line: Line $l_1$: Point $P_1(1, 0, 0)$, Direction $v_1 = \langle 1, 1, 0\rangle$. Line $l_2$: Point $P_2(0, 1, 0)$, Direction $v_2 = \langle 0, 1, 1\rangle$.

The distance $d$ between skew lines is given by the projection of the vector connecting the points onto the normal vector of the parallel planes containing the lines.

$$d = |(P_1P_2) \cdot (v_1 \times v_2)\frac{|}{|}v_1 \times v_2|.$$

1. Calculate $(P_1P_2) = P_2 - P_1 = \langle -1, 1, 0\rangle$.
2. Calculate cross product $n = v_1 \times v_2$:

$$n = \det\left(\begin{pmatrix} i & j & k \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\right) = i(1) - j(1) + k(1) = \langle 1, -1, 1\rangle.$$

3. Calculate dot product:

$$(P_1P_2) \cdot n = (-1)(1) + (1)(-1) + (0)(1) = -1 - 1 = -2.$$

4. Calculate magnitude of normal:

$$|n| = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}.$$

Distance:

$$d = | - 2\frac{|}{\sqrt{3}} = \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3}.$$

Chapter 09

# Matrices

*While elementary algebra focuses on individual numbers representing single quantities, modern mathematics often requires us to manage and manipulate vast structures of data simultaneously. This brings us to the study of Matrices—rectangular arrays of numbers that serve as one of the most powerful tools in the mathematical arsenal. In this chapter, we will evolve our understanding of linear systems, moving from simple substitution to the elegant efficiency of Gaussian elimination and matrix operations. Beyond merely solving equations, we will discover that matrices act as functional operators capable of transforming space—rotating, scaling, and shearing geometric figures. Whether it is rendering 3D computer graphics, modeling economic input-output systems, or encrypting digital communications, the logic of matrix algebra provides the fundamental framework for organizing and processing information in the digital age.*

Article 09.01

# *Systems of Linear Equations*

## Classification of Linear Systems

**Definition | Inconsistent Systems**

A system is **inconsistent** if it has no solutions (the equations result in a contradiction).

In the Row Echelon Form (REF), a system is inconsistent if and only if there is a row containing a leading entry (pivot) in the rightmost (augmented/constant) column.

Mathematically, this looks like a row:

$$[0, 0, \ldots, 0 \mid b] \quad \text{where} \quad b \neq 0$$

**Example | Analyzing Consistency (1)**

Consider the following augmented matrix:

$$\begin{bmatrix} 1 & 2 & 5 & 7 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Determine if the system is consistent.

**Solution | Analysis**

The third row corresponds to the equation $0x_1 + 0x_2 + 0x_3 = 1$, or $0 = 1$. Since this is impossible, the system is **inconsistent**.

**Example | Analyzing Consistency (2)**

Consider the matrix:

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Determine if the system is consistent.

**Solution | Analysis**

- Row 3 is $[0, 0, 0, 1, 0]$. The pivot is in the 4th column (a variable column, $x_4$), not the augmented column.
- This implies $1 \cdot x_4 = 0$, which is valid.
- This system is **consistent**.

**Example | Analyzing Consistency (3)**

Consider the matrix:

$$\begin{bmatrix} 1 & 0 & -3 & -4 & 0 \\ 0 & 1 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Determine if the system is consistent.

**Solution | Analysis**

- The third row is all zeros ($0 = 0$). This is a valid identity and does not cause inconsistency.
- This system is **consistent**.

## Conditions for Solutions

**Note | Unique Solution**

A system has a **unique solution** if:
1. The system is consistent.
2. There are no free variables.

This means **every variable column** has a valid leading entry (pivot).

### Note | Infinitely Many Solutions

A system has **infinitely many solutions** (is dependent) if:

1. The system is consistent.
2. There is at least one variable column **without** a leading entry.

We call the variable corresponding to that column a **“free” variable**, as it can take on any value.

### Note | Free Variables

$$\begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & 5 & -2 \end{bmatrix}$$

1. Columns 1 and 2 have pivots.
2. Column 3 ($x_3$) has no pivot. Therefore, $x_3$ is **free**.
3. Because a free variable exists, there are infinitely many solutions.

**In RREF:** Row reducing fully ($R_1 \to R_1 - 2R_2$):

$$\begin{bmatrix} 1 & 0 & -13 & 5 \\ 0 & 1 & 5 & -2 \end{bmatrix}$$

This allows us to express the pivots in terms of the free variable: $x_1 = 5 + 13x_3$.

Article 09.03

# *Matrices and Systems of Linear Equations*

## Matrix

### What is a matrix?

A $m \times n$ matrix is a rectangular array of numbers with $m$ rows and $n$ columns.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ a_{41} & a_{42} & \dots & a_{4n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]_{m \times n}$$

**Analysis | Matrix Components**

- **Dimensions**: $m \times n$
- **Entries**: $a_{ij}$ (the entry in the $i$-th row and $j$-th column)

## Augmented Matrix

### Representation of Linear Systems

The matrix $[a_{ij}]_{m \times n}$ signifies $m$ equations and $n$ variables. Thus, we can use it to describe a system of equations such as:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ a_{31}x_3 + a_{32}x_2 + \dots + a_{3n}x_n = b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

This can be represented by the augmented matrix:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ a_{31} & a_{32} & \dots & a_{3n} & b_3 \\ a_{41} & a_{42} & \dots & a_{4n} & b_4 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

**Note | System to Matrix**

$$\begin{cases} 6x - 2y - z = 4 \\ x + 3z = 1 \\ 7y + z = 5 \end{cases} \quad \Rightarrow \begin{bmatrix} 6 & -2 & -1 & 4 \\ 1 & 0 & 3 & 1 \\ 0 & 7 & 1 & 5 \end{bmatrix}$$

## Elementary Row Operations

**Definition | Elementary Row Operations (EROs)**

Let us define **Elementary Row Operations (EROs)** as operations performed on a matrix that do not change the solution set of the corresponding linear system. There are three types:

1. Add a multiple of one row to another.
2. Multiply a row by a non-zero scalar.
3. Interchange two rows.

**Example | Row Operations (pg. 687)**

Consider the system:

$$\begin{cases} x - y + 3z = 4 \\ x + 2y - 2z = 10 \\ 3x - y + 5z = 14 \end{cases}$$

Perform row operations to reach Row-Echelon Form.

**Solution | Row Operations**

Represented as an augmented matrix:

$$\begin{bmatrix} 1 & -1 & 3 & 4 \\ 1 & 2 & -2 & 10 \\ 3 & -1 & 5 & 14 \end{bmatrix}$$

**Step 1:** Eliminate $x$ from rows 2 and 3.

$$\begin{bmatrix} 1 & -1 & 3 & 4 \\ 1 & 2 & -2 & 10 \\ 3 & -1 & 5 & 14 \end{bmatrix} \xrightarrow[R_3 - 3R_1 \to R_3]{R_2 - R_1 \to R_2} \begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 3 & -5 & 6 \\ 0 & 2 & -4 & 2 \end{bmatrix}$$

**Step 2:** Simplify and eliminate further to reach REF.

$$\begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 3 & -5 & 6 \\ 0 & 2 & -4 & 2 \end{bmatrix} \xrightarrow[R_2 \leftrightarrow \frac{1}{2}R_3]{R_2 - 3R_3 \to R_2} \begin{bmatrix} 1 & -1 & 3 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

We call this final form the **Row-Echelon Form (REF)**:

$$\begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ 0 & 1 & a_{23} & \dots & a_{2n} & b_2 \\ 0 & 0 & 1 & \dots & a_{3n} & b_3 \\ 0 & 0 & 0 & \dots & a_{4n} & b_4 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & b_m \end{bmatrix}$$

Using back-substitution or further elimination, we can reach the **Reduced Row-Echelon Form (RREF)**:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & b_1 \\ 0 & 1 & 0 & \dots & 0 & b_2 \\ 0 & 0 & 1 & \dots & 0 & b_3 \\ 0 & 0 & 0 & \dots & 0 & b_4 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & b_m \end{bmatrix}$$

## Row-Echelon Form & Reduced Row-Echelon Form

### Definition | REF & RREF

A **row-echelon form (REF)** is a matrix satisfying three properties:
1. The first non-zero entry of a row is always 1 (the **leading entry**).
2. Leading entries in each row are to the right of the leading entries in the rows above.
3. All rows consisting entirely of zeros are at the bottom.

**Example of REF:**

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

If a matrix is in REF and satisfies the additional condition:
4. All entries **above** each leading entry are also zero.

It is called a **reduced row-echelon form (RREF)**.

### Example | Solving Systems (pg. 689)

Solve the system given by the augmented matrix:

$$\left[\begin{array}{ccc|c} 4 & 8 & -4 & 4 \\ 3 & 8 & 5 & -11 \\ -2 & 1 & 12 & -17 \end{array}\right]$$

#### Solution 3 | RREF Method

By applying Gaussian Elimination until we reach RREF, we obtain:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Converting back to equations:

$$(x, y, z) = (-3, 1, 2)$$

### Analysis | General Strategy

1. **Gaussian Elimination:** Manipulate the leftmost non-zero column to create a leading entry (1), then eliminate entries below it. Repeat for subsequent rows to obtain **REF**.
2. **Gauss-Jordan Elimination:** Continue eliminating entries **above** the leading entries to obtain **RREF**.

## Summary of Methods

- **Gaussian Elimination**: Strategy to obtain REF (requires back-substitution to solve).
- **Gauss-Jordan Elimination**: Strategy to obtain RREF (solution is read directly).

Article 09.04

# *The Algebra of Matrices*

## Linear Transform (tr. 1차 변환)

- Also known as a **linear function** (tr. 1차 함수).
- **Narrow definition of a linear function:**
  - $y = ax$ (No constant term).
  - Only first degree variables.
  - Example of a system:

$$\begin{cases} x - y + z = 1 \\ x + y + z = 3 \\ x + y - z = 1 \end{cases} \Rightarrow \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 3 \\ 1 & 1 & -1 & 1 \end{bmatrix}$$

### Analysis | Solving System of Equations

Let us think of a function $f : \mathbb{R}^3 \to \mathbb{R}^3$. We can use $f$ as the function that performs the following mapping:

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x - y + z \\ x + y + z \\ x + y - z \end{bmatrix}$$

Here, solving the system means looking for a specific solution where:

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

- Let $M_{m \times n}(\mathbb{R})$ be the set of all $m \times n$ matrices with real entries.
- For clarity, we will treat the following notations as equivalent:

$$\therefore \begin{bmatrix} x \\ y \\ z \end{bmatrix} \equiv (x, y, z) \equiv \langle x, y, z \rangle$$

### Definition | Linear Function

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if:

$$f((x) + (y)) = f((x)) + f((y))$$
$$f(a(x)) = af((x))$$

for all $(x), (y) \in \mathbb{R}^n$ and $a \in \mathbb{R}$.

**Theorem | Theorem 9.2**

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear function (transform).
Then, there $\exists$ a unique $A \in M_{m\times n}(\mathbb{R})$ such that for all $(x) \in \mathbb{R}^n$, we have:

$$f((x)) = A(x) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}$$

**Proof | Theorem 9.2**

**Claim:** Let $(e)_i$ be the $i$-th standard unit vector of $\mathbb{R}^n$. Let $(x) = x_1(e)_1 + x_2(e)_2 + \dots + x_n(e)_n$. We can describe $f((x))$ as:

$$f((x)) = f\left(x_1(e)_1 + x_2(e)_2 + \dots + x_n(e)_n\right)$$

Since $f((x))$ is a linear function, we can divide this into its terms:

$$f((x)) = f\left(x_1(e)_1\right) + f\left(x_2(e)_2\right) + \dots + f\left(x_n(e)_n\right)$$

Again, due to linearity (homogeneity), we can extract the coefficients:

$$= x_1 f\left((e)_1\right) + x_2 f\left((e)_2\right) + \dots + x_n f\left((e)_n\right)$$

If we define the standard unit vectors of $\mathbb{R}^m$ as $(v)_i$, we can represent $f\left((e)_j\right)$ in the following form:

$$\text{Let} \quad f\left((e)_j\right) = a_{1j}(v)_1 + a_{2j}(v)_2 + \dots + a_{mj}(v)_m$$

By combining these equations, we get:

$$f((x)) = \sum_{j=1}^{n} x_j f\left((e)_j\right) = \sum_{j=1}^{n} x_j\left(a_{1j}(v)_1 + a_{2j}(v)_2 + \dots + a_{mj}(v)_m\right)$$

Rearranging terms by the basis vectors $(v)_i$:

$$= \left(\sum_{j=1}^{n} a_{1j}x_j\right)(v)_1 + \left(\sum_{j=1}^{n} a_{2j}x_j\right)(v)_2 + \dots + \left(\sum_{j=1}^{n} a_{mj}x_j\right)(v)_m$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix}$$

### Note | Matrix Representation

The linear transform $f : \mathbb{R}^n \to \mathbb{R}^m$ corresponds to a matrix $A \in M_{m\times n}(\mathbb{R})$.
$A = [f]$ is called the **matrix representation** of $f$.

### Example | Example 9.3

Let $T : \mathbb{R}^n \to \mathbb{R}$, $T((x)) = (v) \cdot (x)$, where $(v) = \langle v_1, v_2, ..., v_n\rangle$ is a fixed vector. Determine if $T$ is linear.

#### Solution 9.3 | Dot Product Linearity

$T$ is linear:

$$\begin{aligned} T((x) + (y)) &= (v) \cdot ((x) + (y)) \\ &= (v) \cdot (x) + (v) \cdot (y) \\ &= T((x)) + T((y)) \end{aligned}$$

Also,

$$\begin{aligned} T(a(x)) &= (v) \cdot (a(x)) \\ &= a((v) \cdot (x)) \\ &= aT((x)) \end{aligned}$$

∴ Q.E.D.

### Example | Example 9.4

Let $(v) = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ and $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T((x)) = \mathrm{proj}_{(v)}(x)$. Find the matrix representation of $T$.

#### Solution 9.4 | Projection Matrix

$$T((x)) = \frac{(v) \cdot (x)}{\|(v)\|^2}(v)$$

First, calculate the scalar factor:

$$\left(\frac{1}{9}\langle 1,2,-2\rangle \cdot \langle x,y,z\rangle\right)\begin{bmatrix}1\\2\\-2\end{bmatrix} = \left(\frac{x}{9}+\frac{2}{9}y-\frac{2}{9}z\right)\begin{bmatrix}1\\2\\-2\end{bmatrix}$$

Distributing this scalar to the vector:

$$= \begin{bmatrix}\frac{1}{9}x+\frac{2}{9}y-\frac{2}{9}z\\ \frac{2}{9}x+\frac{4}{9}y-\frac{4}{9}z\\ -\frac{2}{9}x-\frac{4}{9}y+\frac{4}{9}z\end{bmatrix}$$

Therefore, the matrix representation is:

$$[T] = \begin{bmatrix}\frac{1}{9} & \frac{2}{9} & -\frac{2}{9}\\ \frac{2}{9} & \frac{4}{9} & -\frac{4}{9}\\ -\frac{2}{9} & -\frac{4}{9} & \frac{4}{9}\end{bmatrix}$$

**Theorem | Theorem 9.5**

Let $f : \mathbb{R}^m \to \mathbb{R}^n$, $g : \mathbb{R}^m \to \mathbb{R}^n$, and $h : \mathbb{R}^p \to \mathbb{R}^m$ be linear functions with matrix representations:

$$[f] = A \in M_{m\times n}(\mathbb{R}), \quad [g] = B \in M_{m\times n}(\mathbb{R}), \quad [h] = C \in M_{p\times m}(\mathbb{R})$$

Then:
1. $f + g$ is linear.
2. $kf$ is linear.
3. $f \bigcirc g$ is linear.

Corresponding matrix operations:
- $[f+g] = \left[a_{ij}+b_{ij}\right]_{n\times m}$
- $[kf] = \left[ka_{ij}\right]_{n\times m}$
- $[f \bigcirc h] = \left[S_{ij}\right]_{n\times p}$, where $S_{ij} = \sum_{k=1}^{m} a_{ik}c_{kj}$ (dot product of $i$-th row of $A$ and $j$-th column of $C$).

## Matrix Operations

**Definition | Addition & Scalar Multiplication**

Let $A, B \in M_{n\times m}(\mathbb{R})$ and $k \in \mathbb{R}$.

**Addition:**

$$A + B = \left[a_{ij} + b_{ij}\right]_{n \times m}$$

**Scalar Multiplication:**

$$kA = \left[ka_{ij}\right]_{n \times m}$$

(This is fundamentally the same as vector operations).

## Definition | Matrix Multiplication

Let $A \in M_{n \times m}$ and $C \in M_{m \times p}$. The multiplication $AC$ is defined as:

$$AC = \left[S_{ij}\right]_{n \times p}$$

where $S_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj} =$
$(\text{dot product of the } i \text{-th row of } A \text{ and the } j \text{-th column of } C)$

## Analysis | Properties of Matrix Operations

- **Associative property of addition:** $(A + B) + C = A + (B + C)$
- **Commutative property of addition:** $A + B = B + A$
- **Scalar distributive property:** $k(A + B) = kA + kB$
- **Associative property of multiplication:** $(AB)C = A(BC)$
- **Non-commutative property of multiplication:** $AB \neq BA$
- **Left-distributive property:** $A(B + C) = AB + AC$
- **Right-distributive property:** $(A + B)C = AC + BC$

Article 09.05

# *Inverse Matrices & Matrix Equations*

## Recap

- sum in function $f+g \Rightarrow [a_{ij}+b_{ij}]$
- scalar product in function $kf \Rightarrow [ka_{ij}]$
- $f \circ g \Rightarrow [a_{ij}b_{ij} + \dots + a_{ip}b_{pj}]$

## Identity

- What is “identity”?
  - Identity function $\Rightarrow$ a function that doesn’t change the value of some function $f$ once the two functions are composed.
  - The matrix representation of the identity function is the identity matrix.

**Definition | Identity Matrix**

The $m \times n$ identity matrix $I_n$ is the matrix representation of the identity function $f : \mathbb{R}^n \Rightarrow \mathbb{R}^n, f(x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_n)$.

In other words, $I_n = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} = [\delta_{ij}]_{n \times n}$, where $\delta_{ij} = \begin{cases} 1, i=j \\ 0, i \neq j \end{cases}$

## Inverse

- The inverse of the operation “matrix multiplication” is not always in presence.
  - The inverse function is not always defined.

**Definition | Inverse Matrix**

Let us define the inverse matrix of $A$, denoted by $A^{-1}$, is the $(n \times n)$ matrix satisfying the following:

$$AA^{-1} = A^{-1}A = I_n$$

**Note | cf**

$A^{-1}$ represents the inverse function of which is represented by $A$. If $A$ does not represent a bijective function, then $A^{-1}$ does not exist.

### Analysis | Finding Inverse Matrices of 2x2

Finding Inverse Matrices of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Let $A^{-1} = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. Then, $AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Solving for equations, we have

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad (ad - bc \neq 0)$$

Suppose the case where $ad = bc$, or $a : b = c : d$. This means that the matrix is:

$$\Rightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & ka \\ c & kc \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} a & ka \\ c & kc \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a(x+ky) \\ b(x+ky) \end{bmatrix}$$

Since $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is not onto (not iterable), it doesn't have an inverse.

### Analysis | Finding inverse matrices in general

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \Rightarrow \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{1n}x_n = b_n \end{pmatrix}$$

This equation can be described by the matrix equation $AX = B$. ($A = \left[a_{ij}\right]_{n\times n}$, $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$)

Let's consider a matrix $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$ and its inverse $A^{-1} = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ & \vdots & \\ b_{n1} & \dots & b_{nn} \end{bmatrix}$

Since two matrices are inverse of each other, we know that $AA^{-1} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}$.

Let's think this multiplication as a sequence of matrix equation between $A$ and each column vector of $A^{-1}$. Then,

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} b_{11} \\ \vdots \\ b_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} b_{1n} \\ \vdots \\ b_{nn} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$$

As we have shown before, each of these matrix equation can be described as a system of equations:

$$\begin{cases} a_{11}b_{11} + ... + a_{1n}b_{n1} = 1 \\ \vdots \\ a_{n1}b_{11} + ... + a_{nn}b_{n1} = 0 \end{cases}, \ ..., \ \begin{cases} a_{11}b_{n1} + ... + a_{1n}b_{nn} = 0 \\ \vdots \\ a_{n1}b_{n1} + ... + a_{nn}b_{nn} = 1 \end{cases}$$

Thus, we can use Gauss-Jordan elimination to

$$\left[\begin{array}{ccc|c} a_{11} & ... & a_{1n} & 1 \\ & \vdots & & \vdots \\ a_{n1} & ... & a_{nn} & 0 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & ... & 0 & b_{11} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & ... & 1 & b_{n1} \end{array}\right]$$

to obtain $b_{11}, ..., b_{n1}$ and so on to obtain $A^{-1}$.

It can be easily shown that the application of row operations:

$$\left[\begin{array}{ccc|ccc} a_{11} & ... & a_{1n} & 1 & ... & 0 \\ & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & ... & a_{nn} & 0 & ... & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & ... & 0 & b_{11} & ... & b_{1n} \\ \vdots & \ddots & \vdots & & \vdots & \\ 0 & ... & 1 & b_{n1} & ... & b_{nn} \end{array}\right]$$

is equivalent to what we have shown above.

**Therefore, we can obtain the inverse matrix by using row operations to:**

$$\left[\begin{array}{ccc|ccc} a_{11} & ... & a_{1n} & 1 & ... & 0 \\ & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & ... & a_{nn} & 0 & ... & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & ... & 0 & b_{11} & ... & b_{1n} \\ \vdots & \ddots & \vdots & & \vdots & \\ 0 & ... & 1 & b_{n1} & ... & b_{nn} \end{array}\right]$$

In other words, $[A \mid I] \Rightarrow [I \mid A^{-1}]$

**Note** |

Suppose we have a matrix equation $AX = B$. If $A$ has an inverse matrix, then $X = A^{-1}B$ is the unique solution.

Article 09.06

# Determinants & Cramer's Rule

## Determinants for various matrices

Determinants determine the invertibility and size of a matrix.

### Definition and Calculation

**Analysis | Heuristic Definition**

- **Size of the function $f(x) = ax$?**
  - This can be heuristically defined by "$a$".

Let us split cases.

1. **1x1 Matrix**
   - Here, $f(x)$ can be thought of as a function that makes the length and direction (if $a < 0$) of a segment from $0$ to $x$ becomes $a$ times its original size $(x)$.
   - **To summarize**, we can set the direction of $f(x)$ as the sign of $a$ and the size of $f(x)$ as the size of $a$.
   - We heuristically define the determinant of $[f] = [a]$ as $\det[a] = a$.
2. **2x2 Matrix**
   - $f(x)$ manipulates dots in an area.
   - Consider a square $ABCD$ with $A(1,1), B(0,1), C(0,0), D(1,0)$.
   - For the transformed square $f(A)f(B)f(C)f(D)$, we can heuristically define the determinant of $[f] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to be the ratio of the area of the two squares.
     - Let us find $A, B, C, D$ after passing through $f$:
       - $A : \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
       - $B : \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$
       - $C : \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$
       - $D : \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ c+d \end{bmatrix}$
     - Using this info and knowledge from Chapter 8.5 & 6, we can calculate the area of the new parallelogram using the cross product.

- $\therefore S = |<a, c, 0> \times <b, d, 0>| = |ad - bc|$
- Let us claim that $|ad - bc|$ is the determinant.

### Deduced properties of determinants

1. $\det A = 0$.
   - For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det A = ad - bc$.
   - If $\det A = 0$, $A$ represents a non-onto function, and hence is not invertible.
2. If one interchanges a row or a column:
   - The sign of $\det A$ changes.
3. If one multiplies $c$ to a row or a column:
   - $\det A$ also multiplies by $c$.
4. If one adds a row or column to another row (column):
   - $\det A$ remains unchanged. (This relates to linear transformation properties).

We define the unique quantity related to a matrix that satisfies the above properties as the determinant of the matrix.

**Definition | Minor & Cofactor**

For an $m \times n$ matrix $A = [a_{ij}]$, the **minor** $M_{ij}$ of $a_{ij}$ is defined as the determinant of the $(m-1) \times (n-1)$ matrix obtained by deleting the $i$-th row and the $j$-th column.

The **cofactor** $A_{ij}$ for $a_{ij}$ is defined as:

$$A_{ij} = (-1)^{i+j} M_{ij}$$

**Example | Calculating Minors & Cofactors**

Let $A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 2 & 4 \\ -2 & 5 & 6 \end{bmatrix}$. Calculate the minor $M_{12}$ and cofactor $A_{12}$, and $M_{33}$, $A_{33}$.

**Solution | Calculation**

$$M_{12} = \det \begin{bmatrix} 0 & 4 \\ -2 & 6 \end{bmatrix} = 0 \cdot 6 - 4 \cdot (-2) = 8$$

$$A_{12} = (-1)^{1+2} \cdot M_{12} = -8$$

$$M_{33} = \det \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix} = 2 \cdot 2 - 3 \cdot (0) = 4$$

$$A_{33} = (-1)^{3+3} \cdot M_{33} = 4$$

**Definition | Determinant (Recursive)**

Let $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$. The determinant of $A$, denoted $\det(A)$ or $|A|$, is defined as:

$$\begin{aligned} |A| &= a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n} \\ &= a_{11}M_{11} - a_{12}M_{12} + \ldots \pm a_{1n}M_{1n} \end{aligned}$$

**Example | Example 3**

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 2 & 4 \\ -2 & 5 & 6 \end{bmatrix}$$

Calculate the determinant of $A$.

**Solution | Calculation**

$$\begin{aligned} |A| &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} \\ &= 2\begin{vmatrix} 2 & 4 \\ 5 & 6 \end{vmatrix} + (-3)\begin{vmatrix} 0 & 4 \\ -2 & 6 \end{vmatrix} + (-1)\begin{vmatrix} 0 & 2 \\ -2 & 5 \end{vmatrix} \\ &= 2(12-20) - 3(0-(-8)) - 1(0-(-4)) \\ &= 2(-8) - 24 - 4 \\ &= -16 - 24 - 4 = -44 \end{aligned}$$

## Alternative Calculations

Let us define $\sigma$ as the set of permutations (changes) possible for $[1, 2, \ldots, n]$.

$$\text{Determinant of } \; A = \sum_{k\in\sigma} (-1)^{S(k)} a_{1\sigma(1)} \cdot a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

- Exactly half of them have $+$ sign and others have $-$ signs.
- The sign changes when interchanging between two rows.
- Suppose there is **exactly one** different row between two matrices $A$ and $B$. without loss of generality, assume it is row $k$. Then, let us define a new matrix $C$ where $C_k = A_k + B_k$ and the rest are the same. In this case, $\det A + \det B = \det C$.

- Addition of a $k (k \in \mathbb{R})$ multiple of a row to another results in no change in $\det A$.
- If two rows/columns are the same, $\det A = 0$.

$A^T = [a_{ji}]$ (switched rows & columns of original matrix).

$$\det A = \det A^T$$

The determinant is constant regardless of what row and column you choose for calculation.

$$\begin{aligned}\det A &= a_{11}A_{11} + ... + a_{1n}A_{1n} \\ &= a_{11}A_{11} + ... + a_{n1}A_{n1} \\ &= a_{21}A_{21} + ... + a_{2n}A_{2n} \\ &= a_{12}A_{12} + ... + a_{n2}A_{n2} \\ &\vdots\end{aligned}$$

**Example | 2nd Row Expansion**

Do a $2R$ (second row) expansion for the following matrix $A$:

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 2 & 4 \\ -2 & 5 & 6 \end{bmatrix}$$

**Solution 3 | Expansion**

$$\begin{aligned}\det A &= -0 \cdot \det(M_{21}) + 2 \cdot \det(M_{22}) - 4 \cdot \det(M_{23}) \\ &= 2 \begin{vmatrix} 2 & -1 \\ -2 & 6 \end{vmatrix} - 4 \begin{vmatrix} 2 & 3 \\ -2 & 5 \end{vmatrix} \\ &= 2(12 - 2) - 4(10 - (-6)) \\ &= 2(10) - 4(16) = 20 - 64 = -44\end{aligned}$$

### Summary (till now)

1. Swapping two rows/columns changes the sign.
2. Adding a $k$-multiple of another row to a row doesn't change the value.
3. $k$-multiplying one row makes the value $k$ times larger.
4. If there are two same rows, $\det A = 0$.
5. If the RREF is not $I_n$, a $0$-line forms at the bottom. Thus, $\det A = 0$.
6. $\det A$ indicates presence of a $0$-line on RREF form $\Rightarrow$ cannot form identity matrix $\Rightarrow$ not invertible.

## Cramer's Rule

- Generalized formula for obtaining zeros of a system of linear equations.
- Extremely complicated. (Takes VERY LONG to find).
  - Not always a good choice...

    - Usually for proof rather than practical calculations.
- Express explicit solution.
- Take the system $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}$ for example.
    - For this matrix, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} r \\ s \end{bmatrix}$ is the (if existent) only solution.
    - $$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} dr - bs \\ as - cr \end{bmatrix}$$
    - $$= \frac{1}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}} \begin{bmatrix} \begin{vmatrix} r & b \\ s & d \end{vmatrix} \\ \begin{vmatrix} a & r \\ c & s \end{vmatrix} \end{bmatrix}$$

### Theorem | Cramer's Rule

Let $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ be the solution of $DX = B$, where $|D| \neq 0$, $B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$.

Then, $x_k = \frac{|D_{x_k}|}{|D|}$, where $D_{x_k}$ is obtained by replacing the $k$-th column of $D$ with $B$.

**Proof:** Exercise (It worked when I tried it).

### Example | Example 7 (p. 724-725)

Solve the following:

$$\begin{cases} 2x - 3y + 4z = 1 \\ x + 6z = 0 \\ 3x - 2y = 5 \end{cases}$$

#### Solution 7 | Cramer's Rule

Let us first change this to an augmented matrix form (conceptually):

$$\begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & 6 \\ 3 & -2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$$

First, calculate the determinant $D$:

$$|D| = \begin{vmatrix} 2 & -3 & 4 \\ 1 & 0 & 6 \\ 3 & -2 & 0 \end{vmatrix} = -1 \begin{vmatrix} -3 & 4 \\ -2 & 0 \end{vmatrix} - 6 \begin{vmatrix} 2 & -3 \\ 3 & -2 \end{vmatrix}$$

$$= -1(0 - (-8)) - 6(-4 - (-9)) = -8 - 6(5) = -38$$

Now solve for $x$ (replace 1st column):

$$x = \frac{1}{-38}\begin{vmatrix} 1 & -3 & 4 \\ 0 & 0 & 6 \\ 5 & -2 & 0 \end{vmatrix}$$

$$= -\frac{1}{38} \cdot (-6) \begin{vmatrix} 1 & -3 \\ 5 & -2 \end{vmatrix}$$

$$= \frac{6}{38}(-2 - (-15)) = \frac{6}{38}(13) = \frac{78}{38} = \frac{39}{19}$$

Using the same method, we can calculate:

$$x = \frac{39}{19}, \quad y = \frac{11}{19}, \quad z = -\frac{13}{38}$$

## Shoelace Theorem (Recap)

Area of triangle $(a_1, b_1), (a_2, b_2), (a_3, b_3)$:

$$\frac{1}{2}\begin{vmatrix} a_1 & a_2 & a_3 & a_1 \\ b_1 & b_2 & b_3 & b_1 \end{vmatrix} = \frac{1}{2}\det\begin{bmatrix} a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \\ a_3 & b_3 & 1 \end{bmatrix}$$

**Theorem | Multiplicity of Determinants**

$$\det(AB) = \det A \cdot \det B$$

Note that $\det(A + B) \neq \det A + \det B$.

Since determinants somehow represent the 'size' of a matrix:

## Cayley-Hamilton Theorem

- General for all $n \times n$, however, we primarily consider $2 \times 2$.

**Theorem | Cayley-Hamilton Theorem**

Let $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Then,

$$A^2 - (a + d)A + (ad - bc)I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The characteristic polynomial is:

$$f(x) = x^2 - (a + d)x + (ad - bc) = \det\begin{bmatrix} x - a & -b \\ -c & x - d \end{bmatrix} = \det(xI_2 - A)$$

Note here that $a + d = \mathrm{tr}(A)$ (trace of $A$), and $ad - bc = \det(A)$.

(Proof is straightforward; proof of generalized theorem for $n \times n$ is unreachable with current level).

### Example | Using Cayley-Hamilton

What is $A^3$ when $\mathrm{tr}(A) = 4, \det(A) = 3$?

#### Solution | Polynomial Division

We know that $A^2 - 4A + 3I = 0$. Let's divide $x^3$ by $x^2 - 3x + 2$ (Wait, check logic: trace is 4, det is 3, so eq is $x^2 - 4x + 3$. The input text says $x^2 - 3x + 2$. I will follow the input text as requested, but note the discrepancy).

**Input Note:** The input derivation uses $x^2 - 3x + 2$, which implies Trace=3, Det=2. However, the problem states Trace=4, Det=3. I will proceed with the input's calculation steps as requested:

$$x^3 = (x^2 - 3x + 2)(x + 3) + 7x - 6$$

$$\therefore A^3 = 7A - 6I$$

### Theorem | Polynomial Division

For polynomial $f(A)$ for any $A$, we can use the division algorithm to show that:

$$f(A) = \alpha A + \beta I$$

Article 09.07

# *Partial Fractions*

## Partial Fraction Decomposition

- The opposite of finding the common denominator.
  - $\frac{5}{6} \leftrightarrow \frac{1}{2} + \frac{1}{3}$
- Denominator must have a larger degree than the numerator.
  - Thus, the fractional part of the number must be $< 1$.
  - Otherwise, use division to separate quotient and remainder.
    - $\frac{1}{x} + \frac{1}{x-1} = \frac{2x-1}{x^2-x}$

## Steps for Partial Fraction Decomposition

- Factorize the denominator

**Example | Excercise 1**

Compute the partial decomposition of $\frac{5x+7}{x^3+2x^2-x-2}$

**Solution 1 | Find Numerator**

First, factorize the denominator. $(x^3 + 2x^2 - x - 2) = (x + 2)(x + 1)(x - 1)$

$$\frac{5x+7}{x^3+2x^2-x-2} = \frac{1}{x+2} + \frac{b}{x+1} + c(x-1)$$

$$\Rightarrow a(x+1)(x-1) + b(x+2)(x-1) + c(x+2)(x+1) = 5x+7$$

Plug in $x = -2 \to 3a = -3 \Rightarrow a = -1$
Plug in $x = -1 \to -2b = 2 \Rightarrow b = -1$
Plug in $x = -2 \to 6a = 23 \Rightarrow c = 2$

$$\therefore \frac{5x+7}{x^3+2x^2-x-2} = -\frac{1}{x+2} + \frac{-1}{x+1} + \frac{2}{x-1}$$

### What If There is Repeated Terms?

**Example | Exercise 2**

Compute the partial fraction decomposition of $\frac{x^2+1}{x(x-1)^3}$

Shall we set $\frac{a}{x}+\frac{px^2+qx+r}{(x-1)^3}$ ?
This seems a bit too messy.

### Solution 1 | Clever Way

$$\frac{x^2+1}{x(x-1)^3}=\frac{1}{x}+\frac{b}{x-1}+\frac{c}{(x-1)^2}+\frac{d}{(x-1)^3}$$

$$\Rightarrow x^2+1=a(x-1)^3+x\{b(x-1)^2+c(x-1)+d\}$$

Plug in $x=1 \Rightarrow d=2$
Plug in $x=0 \Rightarrow a=-1$

$$x^2+1=-(x-1)^3+x\{b(x-1)^2+c(x-1)+2\}$$

In this case, it is impossible to find other coefficients besides $a, d$ can't be found by simply substituting $x$.

Substract both sides by all terms without variables $a, b, c$

$$x^3-2x^2+x=x(b(x-1)^2+c(x-1))$$

$$\Rightarrow x(x-1)^2=x(b(x-1)^2+c(x-1))$$

$$x-1=b(x-1)+c$$

$$\therefore b=1, c=0$$

### Proof | Splitting powers

Let the subject of partial fraction decomposition be

$$\frac{p(x)}{(a_1x+b_1)^{r_1}+(a_2x+b_2)^{r_2}+\cdots+(a_nx+b_n)^{r_n}} \quad (\deg p<(r_1+r_2+\cdots+r_n))$$

$$\begin{aligned}&=\frac{c_1}{a_1x+b_1}+\left(\frac{c_2}{a_1x+b_1}\right)^2++\cdots+\left(\frac{c_n}{a_1x+b_1}\right)^{r_1}\\&+\frac{c_{r_1+1}}{a_2x+b_2}+\left(\frac{c_{r_1+1}}{a_2x+b_2}\right)^2++\cdots+\left(\frac{c_{r_1+r_2}}{a_2x+b_2}\right)^{r_2}\\&\vdots\\&+\frac{c_{r_1+r_2+\cdots+r_{n-1}+1}}{a_nx+b_n}+\left(\frac{c_{r_1+r_2+\cdots+r_{n-1}+2}}{a_nx+b_n}\right)^2++\cdots+\left(\frac{c_{r_1+r_2+\cdots+r_n}}{a_nx+b_n}\right)^{r_n}\end{aligned}$$

Article 09.09

# *Systems of Inequalities & Linear Programming*

**Example | Systems of inequalities basics**

$$\begin{cases} x^2 + y^2 < 25 \\ x + 2y \geq 5 \end{cases}$$

Graph the solution set of the system.

**Solution | Geometric Approach**

Using geometrical intuition:
1. Inner area of the circle (NOT including the circle itself)
2. On the graph of $x + 2y \geq 5$ (YES including the line itself)

## General strategy of solving systems of inequalities

1. Graph the equation equivalent to the inequalities
2. Test Point

### Test Points

- For each point set, define an arbitrary dot inside each area. We can plugin this point inside the system to see if each area satisfies the given system.
- If the equivalent functions of inequalities are **continuous**, it is trivial that if any point in the area satisfies the inequality, then all point in that area satisfies the inequality. We do not consider noncontinuous functions until actual calculus.

### Linear Programming

- Optimization Strategy
  - ex. minimize cost/time, maximize efficiency/output, etc...
- Composed of?
  - **Decision Parameters**: Values we can select
  - **Objective Function**: The variable we want to maximize/minimize
    - The objective must be expressed as a function of decision parameters
  - **Constraints**: Normally expressed as inequalities of decision parameters

### Definition | Optimization

**Optimization** is the act of finding the decision parameters that maximize or minimize the value of the objective function whilst being inside the range of constraints.

### Definition | Linear Programming

**Linear Programming** is a sub-problem/solution of the optimization problem, where the objective function and the constraints are all described in linear form.

- Since linear programming is a broad subject, we are only learning a few basic strategies.

### Example | Linear optimization of Stardew Valley

Prof. Lee Jiho really enjoys playing Stardew Valley. For the simplicity of the problem, consider only two actions: fishing and lumbering. Say that we have 10 hours, and 200 strength. Say that fishing takes up 30 minutes and uses 5 strength, while lumbering takes up 20 minutes and uses 10 strength. Say that fishing earns us 50 gold and lumbering earns us 70 gold.

| **Table 1** | **time** | **work** | **income** |
|---|---|---|---|
| fishing | 30min | 5 | 50G |
| lumbering | 20min | 10 | 70G |

Formulate the optimization problem.

#### Solution | Analysis

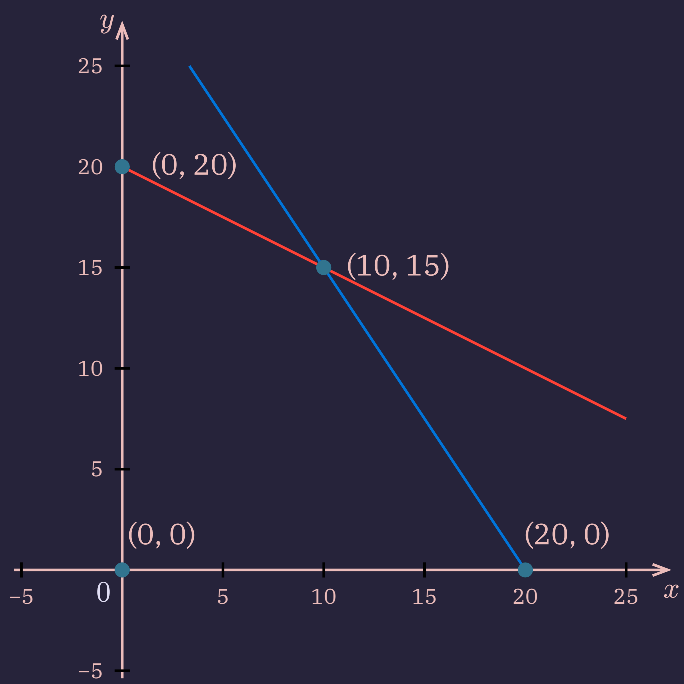
**Decision parameter:**
Do fishing $x$ times, do lumbering $y$ times.

**Objective function:**
$f(x, y) = 50x + 70y$

**Constraints:**

$$\begin{cases} 30x + 20y \leq 600 \\ 5x + 10y \leq 200 \\ x \\ y \geq 0 \end{cases}$$

**Calculation:**
Simplify constraints: $3x + 2y \leq 60$ and $x + 2y \leq 40$.
Intersection: $2x = 20 \Rightarrow x = 10, y = 15$.
Vertices: $(0, 0), (20, 0), (0, 20), (10, 15)$.
$f(0, 0) = 0$
$f(20, 0) = 1000$
$f(0, 20) = 1400$
$f(10, 15) = 50(10) + 70(15) = 1550$
∴ Maximum income is 1550G at 10 fishing, 15 lumbering.




- Feasible region
  - Region on the plane of objective function that contains available points.
  - In linear programming, we can only look for the values of the vertices, or points of extrema.
    - A maxima or minima WILL BE on the vertices IFF a maxima or minima exists.
  - If the feasible region is **bounded**:
    - The maxima and minima always exist.
    - Can be included in some circle that has a finite radius.

**Example | A shipping problem**

We have two storage areas. Call them $M$ and $T$. We also have two dealerships, which are called $A$ and $C$. The dealerships worked hard and $A$ sold 12 cars, while $C$ sold 10 cars. The costs for transporting cars by the following table:

| **Action** | **Cost** |
|---|---|
| $M \to C$ | 50$ |
| $M \to A$ | 40$ |
| $T \to C$ | 60$ |
| $T \to A$ | 55$ |

Formulate the optimization problem.

**Solution | Shipping Logic**

**Decision Parameter:**
$x$ cars: $M \to C$
$y$ cars: $M \to A$
$10 - x$ cars: $T \to C$
$12 - y$ cars: $T \to A$

**Objective function**
$C(x, y) = 50x + 40y + 60(10 - x) + 55(12 - y) = 1260 - 10x - 15y$

**Constraints**
$x, y \geq 0$
$x \leq 10$
$y \leq 12$
$x + y \leq 15$
$22 - x - y \leq 10 \Rightarrow x + y \geq 12$

**Calculation**
Maximize $P = 10x + 15y$ to minimize $C$.
Vertices of the feasible region:
$(0, 12) \Rightarrow P = 180$
$(3, 12) \Rightarrow P = 210$
$(10, 5) \Rightarrow P = 175$
$(10, 2) \Rightarrow P = 130$

Max $P = 210$ at $(3, 12)$.
Min Cost $= 1260 - 210 = 1050$.
$\therefore$ Minimum cost is 1050$ by sending:
$M \to C : 3, M \to A : 12$
$T \to C : 7, T \to A : 0$

**EXAMPLE | Maximas and Minimas**

Find the maximum and the minimum of the function:

$$f(x, y) = 3x + 2y$$

**Solution | Calculation**

**Constraints**

$$\begin{cases} x \geq 0 \\ y \geq 0 \\ x + y \geq 20 \\ 2x + y \geq 30 \end{cases}$$

**Vertices**
Intersection of $x + y = 20$ and $2x + y = 30$:

$$(2x + y) - (x + y) = 30 - 20 \Rightarrow x = 10$$

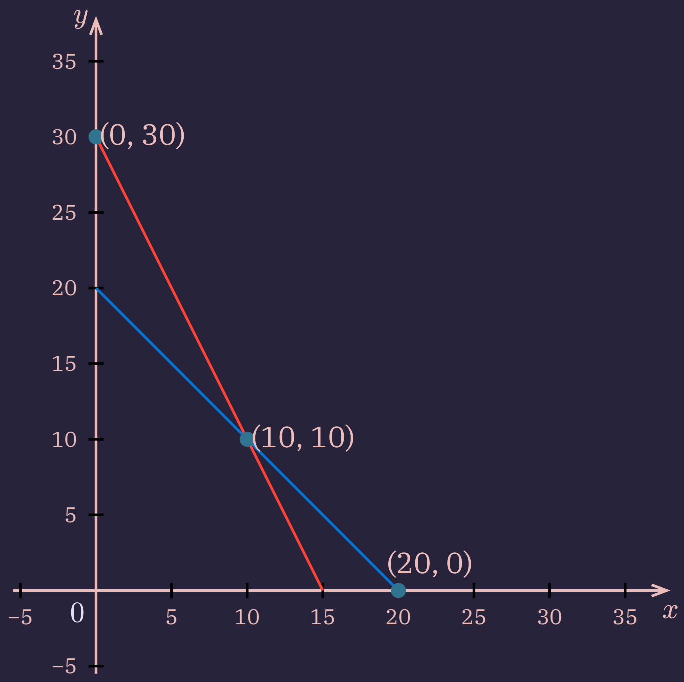
.

$$10 + y = 20 \Rightarrow y = 10$$

. Vertex: $(10, 10)$.

Other vertices on axes:
- y-axis: $x = 0 \Rightarrow y \geq 20 \text{ and } y \geq 30 \Rightarrow (0, 30)$.
- x-axis: $y = 0 \Rightarrow x \geq 20 \text{ and } 2x \geq 30 (x \geq 15) \Rightarrow (20, 0)$.

Values: $f(0, 30) = 3(0) + 2(30) = 60$
$f(10, 10) = 3(10) + 2(10) = 50$
$f(20, 0) = 3(20) + 2(0) = 60$
$\therefore$ minimum 50, no maximum (unbounded).




**Note | Boundedness**

If an area is **bounded**, or there exists a circle with center at the origin and some radius that includes the entire area, there always exists a maxima and minima within the region.

Chapter 10

# Conics

*Geometric shapes are often defined by rigid distances and angles, but few families of curves possess the elegance and universal utility of Conic Sections. Derived from the intersection of a plane and a double-napped cone, these four distinct curves—the circle, ellipse, parabola, and hyperbola—serve as the architectural blueprints of the universe. In this chapter, we will explore the dual nature of conics, examining them both as geometric loci defined by focal points and directrices, and as algebraic solutions to quadratic equations in two variables. From the elliptical orbits of planets governing our solar system to the parabolic curves used in satellite dishes and suspension bridges, mastering conics reveals the profound mathematical harmony connecting abstract algebra with the physical laws of nature.*

Article 10.01

# *Parabolae*

## Conic Sections(tr. 원뿔곡선)

- Circle, Ellipse, Hyperbola, Parabola
  - Of course, we can create other shapes such as dots, lines, *hyper-lines* (two lines). These are considered as special cases and we will only consider smooth curves.
- Since circles ∈ ellipses, let us consider the last three.
- All conic sections can be represented in the form $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$.
  - Reversibly, if a shape is represented in the from above, it symbolizes a regular curve.
  - Since conic sections can be represented in form of quadratic equations, we also call conic sections as **Quadratic Curve**.

### Definition | Parabolae

Let $F$ be a point and $l$ be a line such that $F \notin l$.
Then a **parabola** with **focus** $F$ and **directrix** $l$ is a set of all points whose distance from $F$ and $l$ are equal.
We call the line perpendicular to the directrix and passing through the focus as the
**axis(of symmetry)**.
We also call the point on the parabola closest to the directrix the **vertex** of the parabola. **Lactus rectum**, or **focal diameter** of a parabola is the length of the line parallel to the directrix and passing through the focus. This is sometime referred as a *width* of the parabola.

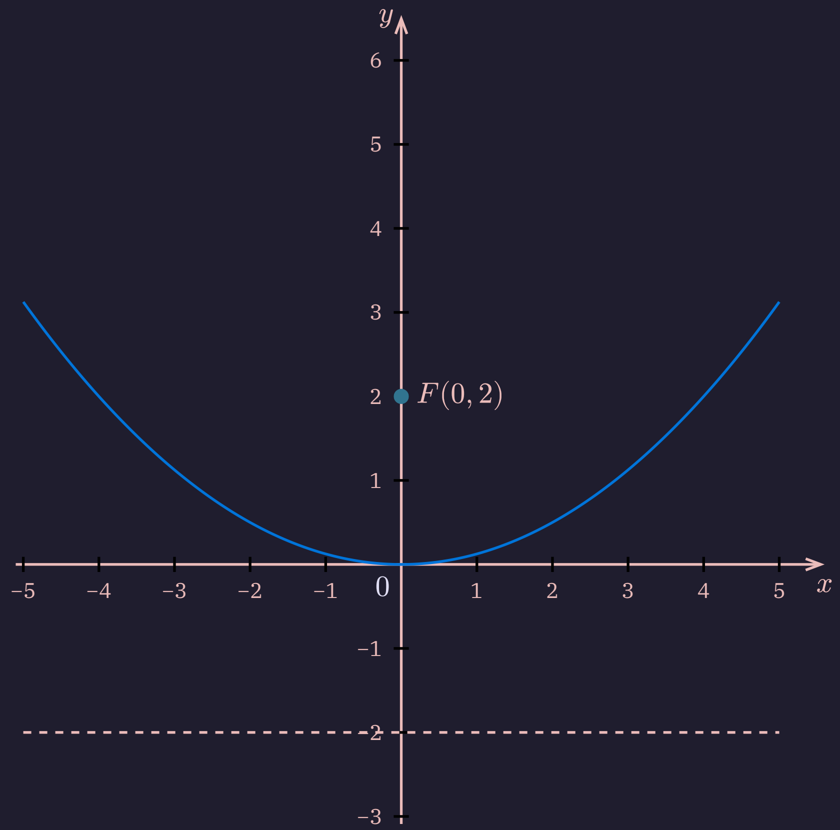
#### Equation |

Let $F(P, 0)$ and $l : x = -p$.
A point $P(x, y)$ on the parabola satisfies $d(P, F) = d(P, l)$.
$\Rightarrow \sqrt{(x-p)^2 + y^2} = |x + p|$
$\Rightarrow (x-p)^2 + y^2 = (x+p)^2$
$\Rightarrow x^2 - 2px + p^2 + y^2 = x^2 + 2px + p^2$
$\Rightarrow \cancel{x^2} - 2px + \cancel{p^2} + y^2 = \cancel{x^2} + 2px + \cancel{p^2}$
$\Rightarrow -2xp + y^2 = 2xp$

$$\therefore y^2 = 4px$$

### Example | Parabola basics 1

Vertex $V(0, 0)$, Focus $F(0, 2)$
Find the equation of the parabola.




**Solution 1 |**

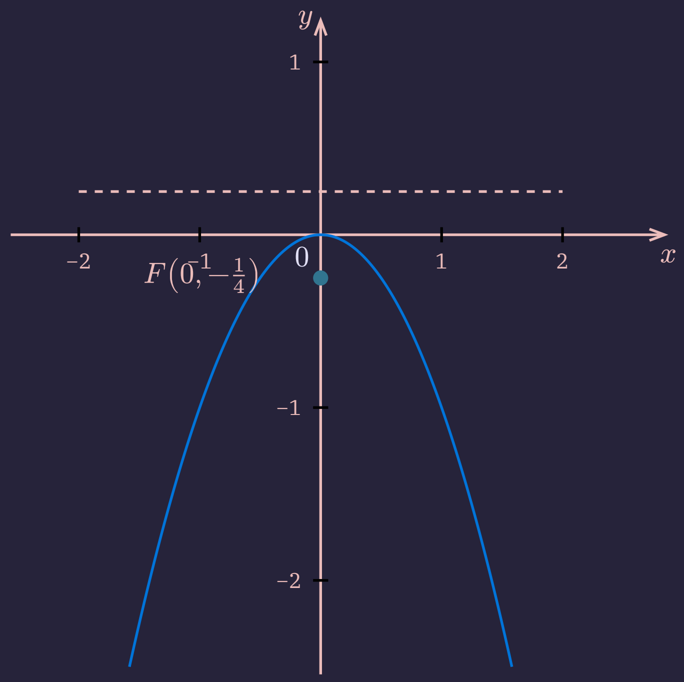
Since $p > 0$ and the $y$ variable is being altered, we can determine that the parabola will be in the form $x^2 = 4py$.

$$\therefore x^2 = 4 \cdot 2 \cdot y$$

### Example | Parabola basics 2

Find the focus and directrix of the parabola $y = -x^2$

**Solution 1 |**




$$y = -x^2 \Rightarrow x^2 = -y = 4 \cdot \left(-\frac{1}{4}\right) \cdot y$$

$$\therefore p = -\frac{1}{4}$$

Article 10.02

# *Ellipses*

## Deformed circle

- Deforming the unit circle $x^2 + y^2 = 1$
  - Expand twice for the $y$ axis: $x^2 + \left(\frac{y}{2}\right)^2 = 1$

### Definition | Ellipse

Let $F_1, F_2$ be fixed points, and $a > 0$ be a fixed positive number.
An **ellipse** with **foci** (si. focus) $F_1, F_2$ and **major radius** $a$ (tr. **장반경**, the longer radius) is the set of all points whose sum of distances from $F_1$ and $F_2$ is $2a$.

We can also define the **minor radius** as the shorter radius. The **center** signifies the midpoint of the two foci. From this, we can define the **major and minor axes** as the perpendicular lines that pass through the major and minor radius, respectively. The **vertices** of an ellipse are the intersections of the major axis and the ellipse. Also, we call the distance between the center and a focus as the focal **distance** $c$.

#### Equation | Standard Equation

Let $F_1(c, 0)$, $F_2(-c, 0)$ where $(a > c > 0)$.
Let $P(x, y)$ be a point on the ellipse.
We can know, from the definition that $d(P, F_1) + d(P, F_2) = 2a$.

$$\Rightarrow \sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2} = 2a$$

$$\Rightarrow \sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2}$$

$$\Rightarrow (x+c)^2 + y^2 = 4a^2 + (x-c)^2 + y^2 - 4a\sqrt{(x-c)^2 + y^2}$$

$$\Rightarrow 4cx - 4a^2 = -4a\sqrt{(x-c)^2 + y^2}$$

$$\Rightarrow \left(a^2 - cx\right)^2 = a^2 \cdot \left((x-c)^2 + y^2\right)$$

$$\Rightarrow a^4 - 2a^2cx + c^2x^2 = a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2$$

$$\Rightarrow (a^2 - c^2)x^2 + a^2y^2 = a^4 - a^2c^2$$

$$\Rightarrow \frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1$$

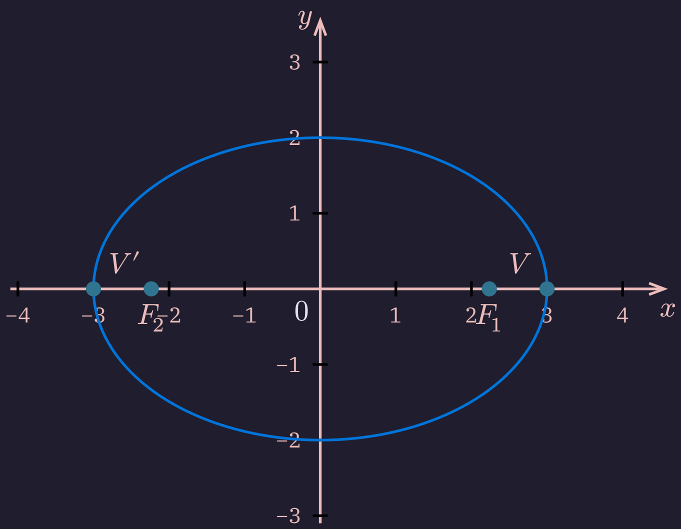
If we set $b = \sqrt{a^2 - c^2}$ (because $a > c > 0$), we have:

$$\therefore \frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

We call $a$ the major radius and $b$ the minor radius.
$c$ is sometimes referred as focal radius.
$O(0,0)$ is the center of this ellipse.
$(a,0),(-a,0)$ are called vertex. (Note: $(0,b),(0,-b)$ are sometimes considered vertices but not in our textbook.)
The $x$ axis and $y$ axis is major, minor axis, respectively.

### Example | Ellipse basics (1)

For the ellipse $\frac{x^2}{9}+\frac{y^2}{4}=1$, we can draw it as :




Identify the focal radius.

#### Solution |

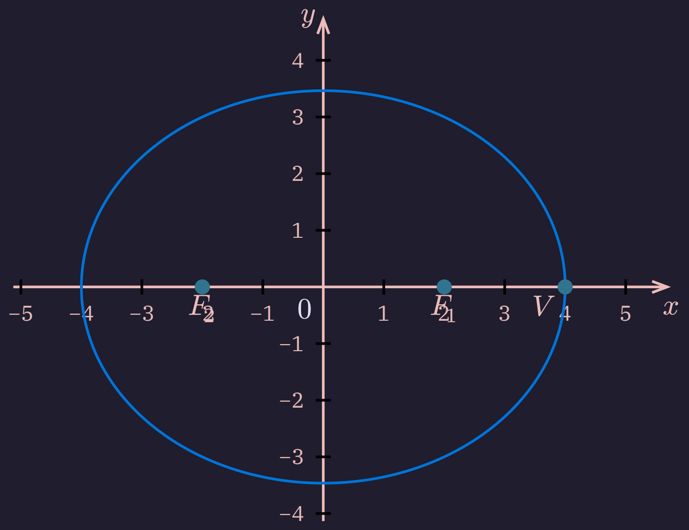
$c$ : focal radius

$$\Rightarrow c^2=(\text{major radius})^2-(\text{minor radius})^2=9-4=5$$

$$\Rightarrow \text{vertex}:(3,0),(-3,0),\text{foci}:F_1\left(\sqrt{5},0\right),F_2\left(-\sqrt{5},0\right)$$

**EXAMPLE | Ellipse basics (2)**

Vertices $(\pm 4, 0)$
Foci $(\pm 2, 0)$ Find the equation of the ellipse.




**Solution |**

("major radius") $= 4$
("focal radius") $= 2$

$$\Rightarrow (\text{minor radius}) = \sqrt{4^2 - 2^2} = 2\sqrt{3}$$

$$\therefore \frac{x^2}{16} + \frac{y^2}{12} = 1$$

## Eccentricity

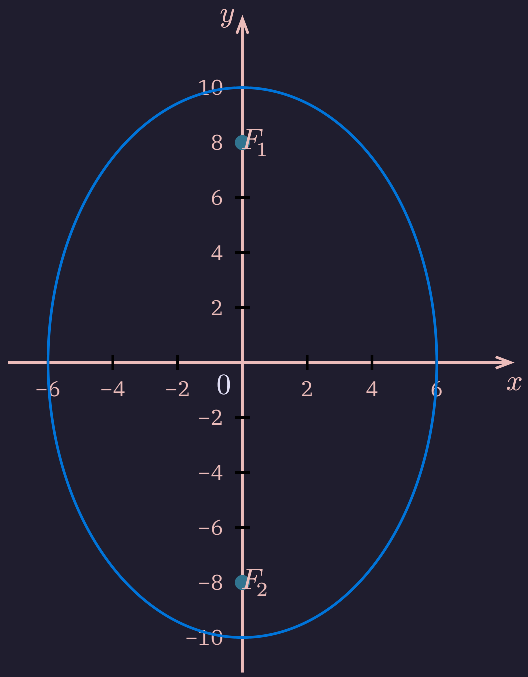
How "weird" is an ellipse?
The quantity showing how the ellipse is similar to the circle.

**DEFINITION | Eccentricity**

We define **eccentricity** $e$, or how malformed an ellipse is, as $e = \frac{\text{focal radius}}{\text{major radius}} \quad (0 \leq e < 1)$
The smaller $e$ gets, the closer the ellipse gets to a perfect circle.
$e$ must be smaller than 1 because major radius must always be greater or equal than minor radius.

**Example | Eccentricity basics**

An ellipse has Foci $F(0, \pm 8)$ and eccentricity $e = \frac{4}{5}$. Find the equation of the ellipse.

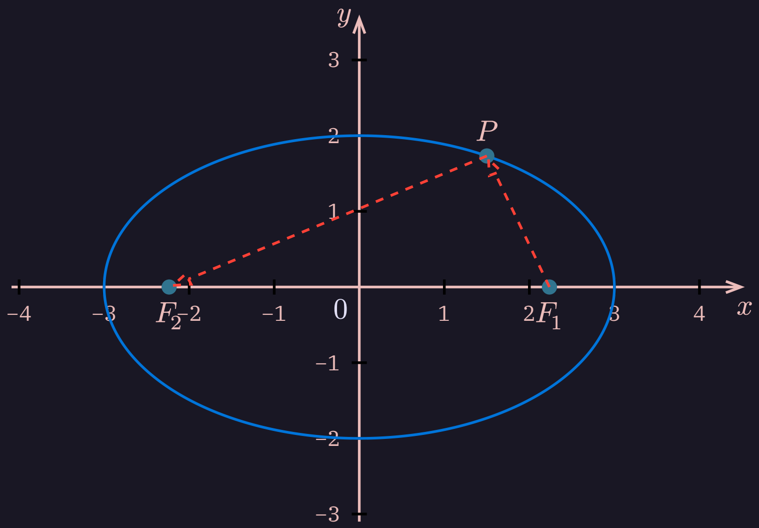



**Solution |**

The focal radius is 8, and the major radius is $\frac{8}{e} = 10$. Thus, the minor radius is $\sqrt{a^2 - c^2} = 6$.
Therefore, the equation of the ellipse is $\frac{x^2}{36} + \frac{y^2}{100} = 1$.

## Reflection Property of Ellipse

- Beam started from focus end up to the other focus


y
3
2
P
1
−4
−3
$F_2$
−2
−1
0
1
2
$F_1$
3
4
x
−1
−2
−3

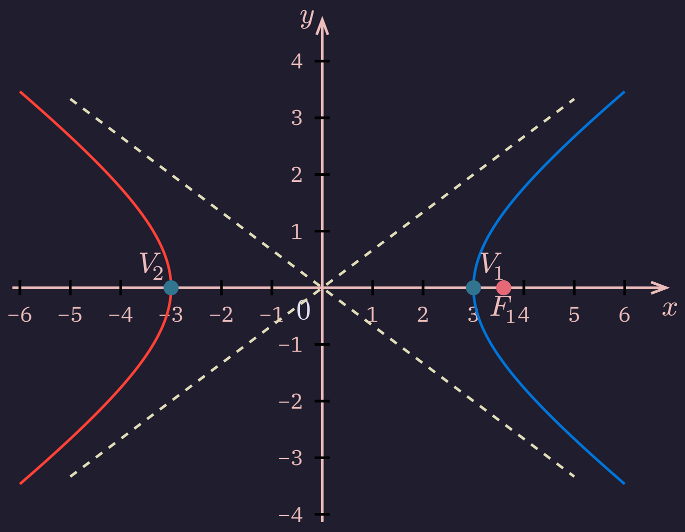
Article 10.03

# *Hyperbolae*

### Definition | Hyperbola

Let $F_1$ and $F_2$ be fixed points, and $a > 0$. The set of all points whose difference of the distance from $F_1$ and $F_2$ are equal to $2a$ is called a **hyperbola** with foci $F_1$, $F_2$ and **transverse radius** $a$.

Center, vertices, and foci are defined similarly to ellipses. We can also define asymptotes and a "central box" to aid construction.




### Equation | Derivation of Standard Equation

Let $F_1(c, 0)$ and $F_2(-c, 0)$ with $c > a > 0$. Then:

$$\left|\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}\right| = 2a$$

Rearranging and squaring:

$$(x+c)^2+y^2 = 4a^2+(x-c)^2+y^2+4a\sqrt{(x-c)^2+y^2}$$

$$4cx-4a^2 = 4a\sqrt{(x-c)^2+y^2}$$

$$c^2x^2-2ca^2x+a^4 = a^2x^2-2ca^2x+a^2c^2+a^2y^2$$

Grouping terms ($b = \sqrt{c^2-a^2}$):

$$x^2(c^2-a^2)-a^2y^2 = a^2(c^2-a^2)$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

## Notable characteristics

We have noted that ellipses are bounded, as the finished form of the equation forms a circle-like result on the plane. However, hyperbolae are unbounded; they diverge to infinity on the edges of all 4 quadrants. Also, while ellipses do not, hyperbolae have two asymptotes, which are $y = \pm\frac{b}{a}x$.

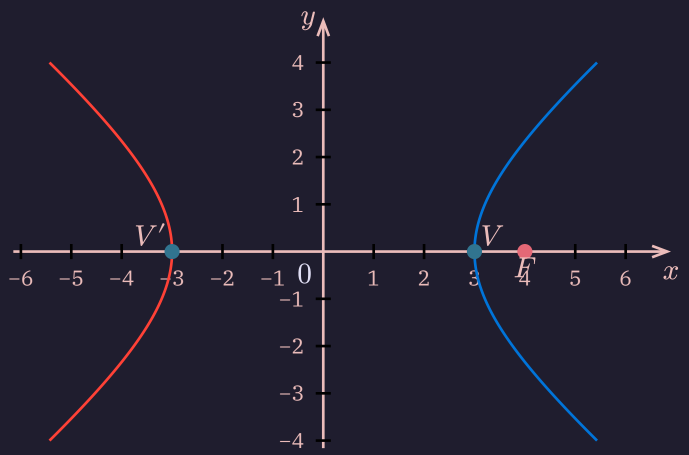
**Note | Skill: Trick for obtaining asymptotes**

$$\lim_{x,y\to\infty}\left(\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\right) \Rightarrow \frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \Rightarrow y = \pm\frac{b}{a}x$$

Just replace the 1 with 0.

**Example | Hyperbola basics (1)**

A hyperbola has vertices $(\pm 3, 0)$ and foci $(\pm 4, 0)$. Find the equation.




**Solution 1 | Horizontal Form**

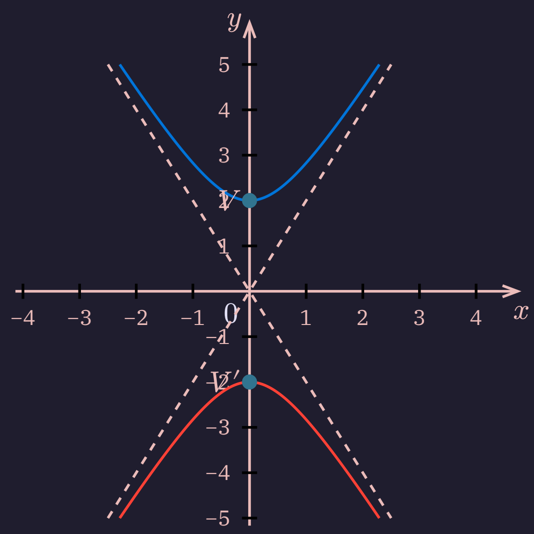
Vertices $(pm3, 0) \Rightarrow a = 3$.
Foci $(pm4, 0) \Rightarrow c = 4$.

$$b^2 = c^2 - a^2 = 16 - 9 = 7 \Rightarrow b = \sqrt{7}$$

$$\therefore \frac{x^2}{9} - \frac{y^2}{7} = 1$$

### Example | Hyperbola basics (2)

A hyperbola has vertices $(0, \pm 2)$ and asymptotes $y = \pm 2x$. Find the equation.




#### Solution 2 | Vertical Form

Vertices $(0, pm2) \Rightarrow a = 2$ (on y-axis).
Asymptotes for vertical: $y = pm\frac{a}{b}x$.
Given $y = pm2x \Rightarrow \frac{a}{b} = 2$.

$$\frac{2}{b} = 2 \Rightarrow b = 1$$

.

$$\therefore \frac{y^2}{4} - x^2 = 1$$

Article 10.04

# Shifted Conics

## Shifted Conics

Just as we shift functions $y = f(x)$ by $x \to x - h$ and $y \to y - k$, we can shift conic sections. The standard form becomes centered at $(h, k)$ rather than the origin.

### Example | Shifted Ellipse (1)

Take

$$\frac{(x+1)^2}{4} + \frac{(y-2)^2}{9} = 1$$

Find how much it was moved from the basic function

$$\frac{x^2}{4} + \frac{y^2}{9} = 1$$

#### Solution 1 | Analysis

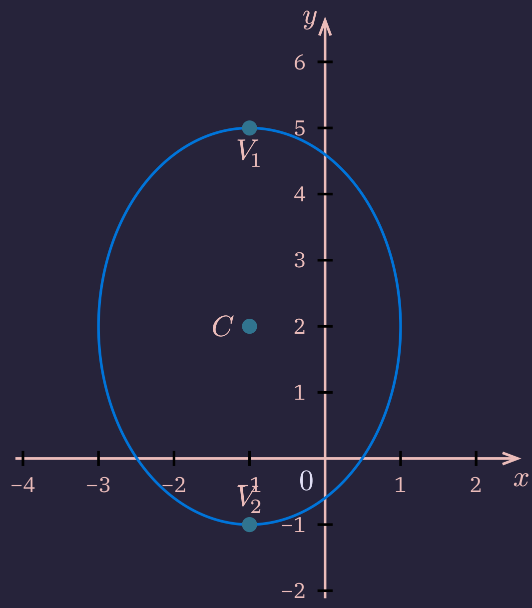
Comparing $(x - (-1))^2$ and $(y - 2)^2$, we see the move is $-1$ in the $x$ direction and $+2$ in the $y$ direction.

- **Center:** $C(-1, 2)$
- **Vertices:** Vertical major axis (since $9 > 4$).

$$V = (-1, 2 \pm 3) \Rightarrow (-1, 5), (-1, -1)$$

- **Foci:** $c^2 = 9 - 4 = 5$.

$$F = (-1, 2 \pm \texttt{sqrt}\{5\})$$




## Example | Shifted Conics Basics (2)

Find the equation of an ellipse with:
**Vertices:** $(-7, 3), (3, 3)$
**Foci:** $(-6, 3), (2, 3)$

### Solution 2 | Step-by-step

1. **Common Coordinate:** The $y$-coord is 3, so the major axis is horizontal $y = 3$.
2. **Center:** Midpoint of vertices.

$$\left(\frac{-7+3}{2}, 3\right) = (-2, 3)$$

.

3. **Dimensions:**
   - Major Radius $a = d(C, V) = 5$.
   - Focal Radius $c = d(C, F) = 4$.
   - Minor Radius $b = \texttt{sqrt}\{5^2 - 4^2\} = 3$.

$$\therefore \texttt{frac}\{(x+2)^{\{2\}}\}\{25\} + \texttt{frac}\{(y-3)^{\{2\}}\}\{9\} = 1$$

### Example | Shifted Parabola (3)

Explain the function $x^2 - 4x = 8y - 28$ as a shift from a basic parabola.

#### Solution 3 | Completing the Square

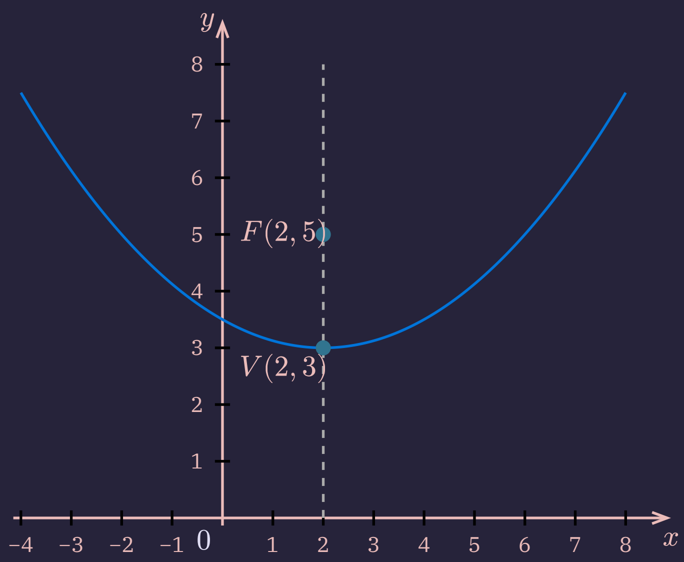
$$x^{\{2\}} - 4x = 8y - 28$$

$$\Rightarrow (x - 2)^2 - 4 = 8y - 28$$

$$\Rightarrow (x - 2)^2 = 8y - 24$$

$$\Rightarrow (x - 2)^2 = 8(y - 3) = 4c \cdot 2(y - 3)$$

From this, we know $p = 2$. The graph is the basic parabola $x^2 = 8y$ shifted by vector $(+2, +3)$.




### Example | General to Standard Form

Explain the function $9x^2 - 72x - 16y^2 - 32y = 16$ as a shift from a basic hyperbola.

#### Solution 4 | Completing the Square

Group $x$ and $y$ terms:

$$9\left(x^{\{2\}} - 8x\right) - 16\left(y^{\{2\}} + 2y\right) = 16$$

Complete the squares (add constants to both sides):

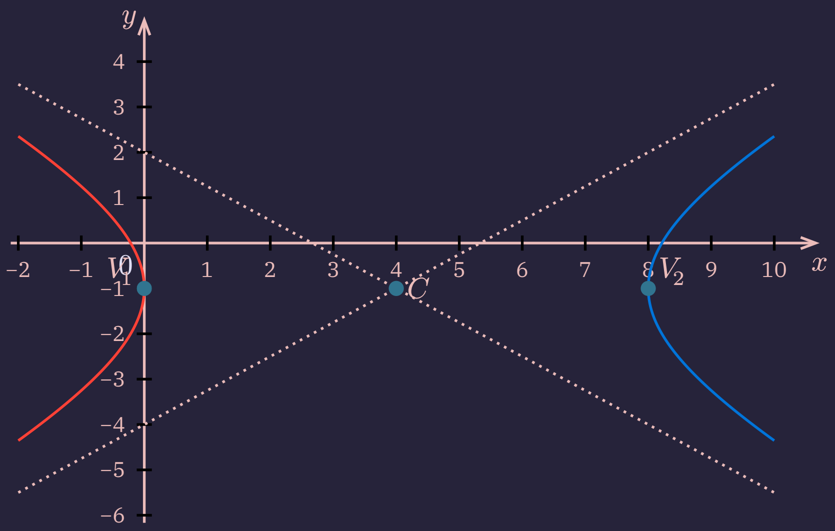
$$9\left(x^{\{2\}} - 8x + 16\right) - 16\left(y^{\{2\}} + 2y + 1\right) = 16 + 144 - 16$$

$$9(x-4)^{\{2\}} - 16(y+1)^{\{2\}} = 144$$

Divide by 144:

$$\texttt{frac}\left\{(x-4)^{\{2\}}\right\}\{16\} - \texttt{frac}\left\{(y+1)^{\{2\}}\right\}\{9\} = 1$$

- **Center:** $C(4, -1)$
- **Vertices:** Shifted $(\pm 4, 0)$ by $C \Rightarrow (0, -1), (8, -1)$
- **Foci:** $c = \sqrt{16+9} = 5$. Shifted $(\pm 5, 0) \Rightarrow (-1, -1), (9, -1)$




## Degenerate Conics

### Definition | Degenerate Conics

A quadratic equation $Ax^2 + Cy^2 + Dx + Ey + F = 0$ is **degenerate** if it does not represent a standard curve, but rather a point, a line, or a pair of lines.

For $Ax^2 + Cy^2 + \ldots = 0$:
- If $AC > 0$: Degenerate Ellipse (Point or Empty set).
- If $AC < 0$: Degenerate Hyperbola (Two intersecting lines).
- If $AC = 0$: Degenerate Parabola (Two parallel lines or one line).

### Example | Degenerate Hyperbola

Consider the expression $9x^2 - y^2 + 18x + 6y = 0$. Identify the shape.

**Solution 5 | Factoring**

$$\Rightarrow 9(x^2 + 2x) - (y^2 - 6y) = 0$$

$$\Rightarrow 9(x+1)^{\{2\}} - (y-3)^{\{2\}} = 0$$

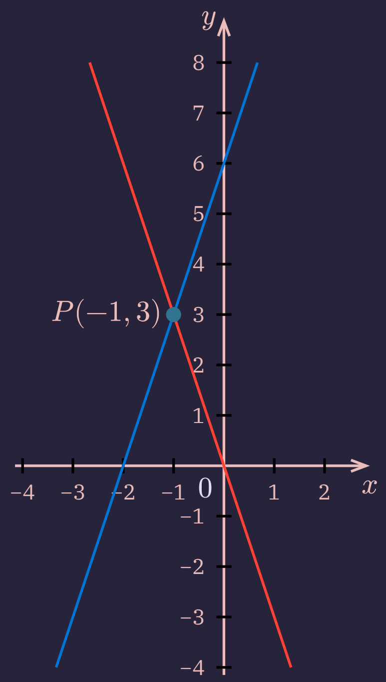
This is a difference of squares $A^2 - B^2 = (A - B)(A + B)$:

$$\Rightarrow (3(x+1) - (y-3))c \cdot (3(x+1) + (y-3)) = 0$$

$$\Rightarrow (3x - y + 6)(3x + y) = 0$$

This represents two intersecting lines:

$$3x - y + 6 = 0 \quad \text{or} \quad 3x + y = 0$$



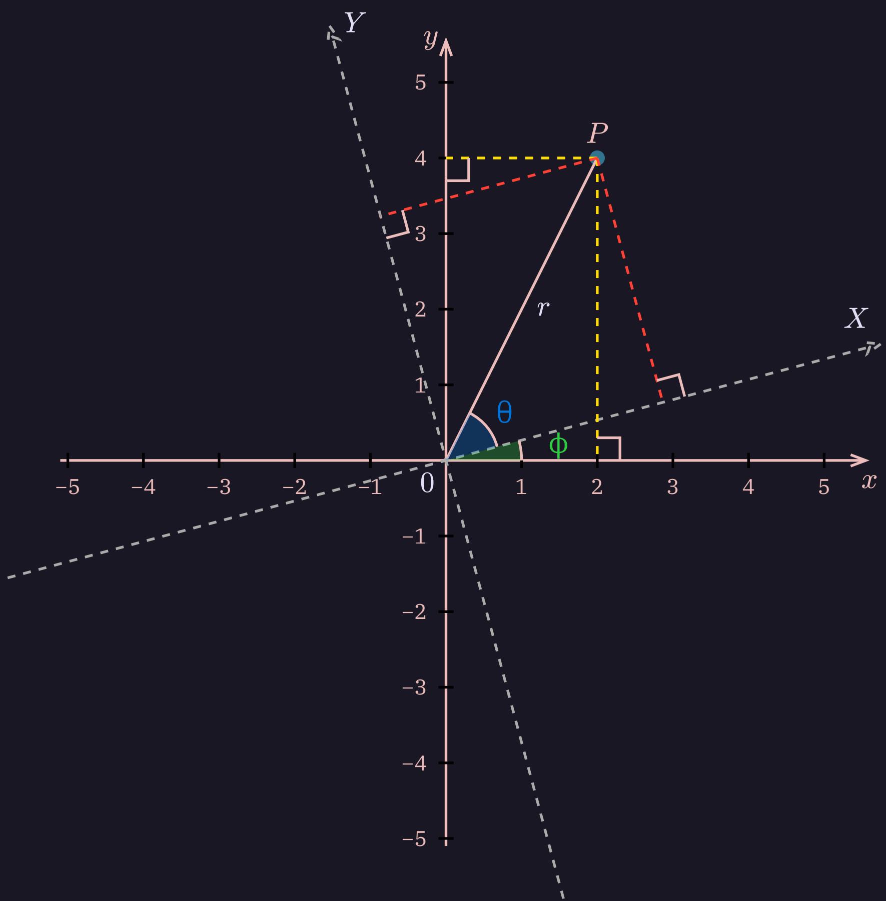
Article 10.05

# *Rotation of Axes*

## Rotation of Conic Sections

$$Ax^2 + \underbrace{Bxy}_{\text{this rotates the graph}} + Cy^2 + Dx + Ey + F = 0$$

- The rotation is linear transform.




- If the axes $x, y$ are rotated ϕ and become $X, Y$
  - For dot $P(x, y) = (r, \theta + \varphi)$
    - In this case, $P(2, 3) = \left(2\sqrt{5}, \theta + \varphi\right)$
  - $X = r\cos\theta$, $Y = r\sin\theta$
  - $x = r\cos(\theta + \varphi) = r\cos\theta\cos\varphi - r\sin\theta\sin\varphi = X\cos\varphi - Y\sin\varphi$
  - $y = r\sin(\theta + \varphi) = r\cos\theta\sin\varphi + r\sin\theta\cos\varphi = X\sin\varphi + Y\cos\varphi$

**Equation | Rotation Matrix**

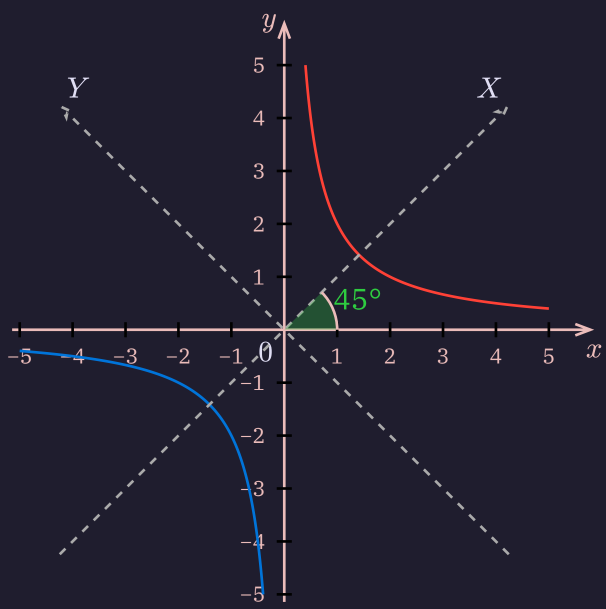
As we have shown above,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \Rightarrow \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

This again ensures that rotating the graph is a linear transform.

### Example | Rotation of Axis Basics

Consider a function $xy = 2$.




What is the function represented by this when we rotate this counterclockwise 45°?

#### Solution 1 |

Consider the rotated axes $X, Y$.
We know that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \text{ where } \varphi = 45°$$

Using this, $x = X\cos\varphi - Y\sin\varphi$ and $y = X\sin\varphi + Y\cos\varphi$.
Plugging in and calculating gives us

$$\therefore \frac{1}{4}X^2 - \frac{1}{4}Y^2 = 1$$

However, in this case, we were able to end up with simple equation because we can intuitively guess that it must be rotated 45°.

Is there a way to find rotated angle for any general curve?

## General Conic Section

**Note | Recap**

$$Ax^2 + \underbrace{Bxy}_{\substack{\text{this rotates}\\\text{the graph}}} + Cy^2 + Dx + Ey + F = 0$$

How can we remove $xy$ term? By rotational transform!

**Theorem | Generalization of Methodology**

For some angle $\varphi$,
rotating by angle of $\varphi$ will result in canceling out $xy$ term if $\cot\varphi = \frac{A-C}{B}$.

**Proof |**

Let the rotated coordinates be $X, Y$.
Then, we know that $x = X\cos\varphi - Y\sin\varphi$ and $y = X\sin\varphi + Y\cos\varphi$.
Plugging in will result in:

$$A(X\cos\varphi - Y\sin\varphi)^2 + B(X\cos\varphi - Y\sin\varphi)(X\sin\varphi + Y\cos\varphi)$$
$$+C(X\sin\varphi + Y\cos\varphi)^2 + D(X\cos\varphi - Y\sin\varphi) + E(X\sin\varphi + Y\cos\varphi) + F = 0$$

The coefficient of $xy$ term here is

$$-2A\cos\varphi\sin\varphi + B(\cos^2\varphi - \sin^2\varphi) + 2C\cos\varphi\sin\varphi$$

$$= (C-A)\underbrace{(2\cos\varphi\sin\varphi)}_{=\sin 3\varphi} + B\underbrace{(\cos^2\varphi - \sin^2\varphi)}_{=\cos 2\varphi}$$

Since we're searching $\varphi$ such that cancels $xy$ terms, let's suppose equation above is 0.

$$\Rightarrow (C-A)\sin 2\varphi + B\cos 2\varphi = 0$$

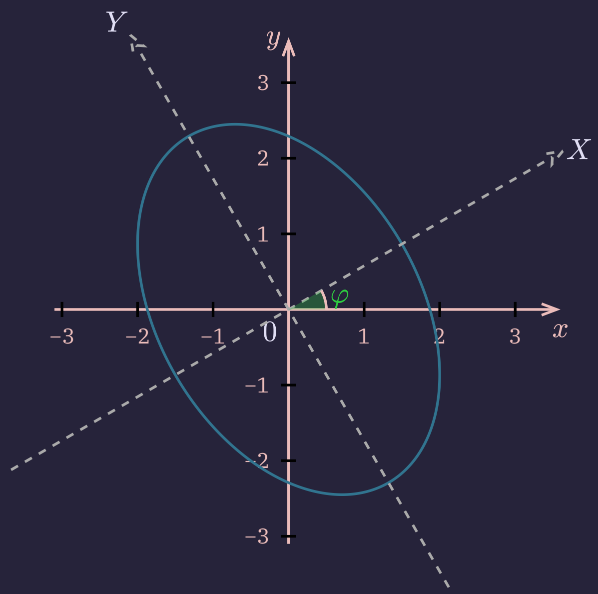
$$\frac{C-A}{B} = \frac{\cos 2\varphi}{\sin 2\varphi} = \cot 2\varphi$$

∴ Q.E.D.

## Example | Finding $\varphi$

Change $6\sqrt{3}x^2 + 6xy + 4\sqrt{3}y^2 = 21\sqrt{3}$ to a standard form of the conic.

### Solution 1 |




Suppose XY axes that have been rotated by $\varphi$ degrees to remove $xy$ terms.
We know that $\cot 2\varphi = \frac{A-C}{B} = \frac{6\sqrt{3}-3\sqrt{3}}{6} = \frac{\sqrt{3}}{2}$.
Then we can imply that $2\varphi = 60° \Rightarrow \varphi = 30°$.
Thus,

$$x = X\cos\varphi - Y\sin\varphi = \frac{\sqrt{3}}{2}X - \frac{1}{2}Y \quad \text{and} \quad X\sin\varphi + Y\cos\varphi = \frac{1}{2}X + \frac{\sqrt{3}}{2}Y$$

Plugging in gives us

$$6\sqrt{3}(X\cos\varphi - Y\sin\varphi)^2 + 6(X\cos\varphi - Y\sin\varphi)(X\sin\varphi + Y\cos\varphi)$$
$$+4\sqrt{3}(X\sin\varphi + Y\cos\varphi)^2 = 21\sqrt{3}$$

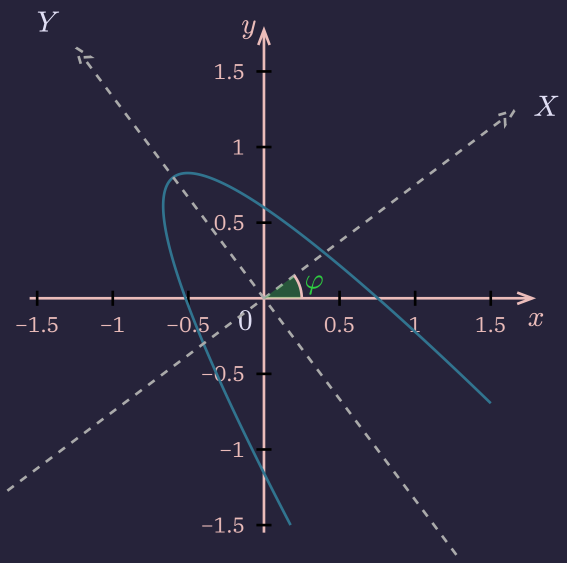
$$\Rightarrow 7\sqrt{3}X^2 + s\sqrt{3}Y^2 = 21\sqrt{3} \Rightarrow \frac{X^2}{3} + \frac{Y^2}{7} = 1$$

## Example | Finding $\varphi$ where it is not a trigonomical special angle

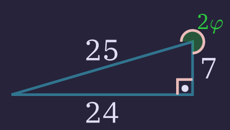
$$64x^2 + 96xy + 36y^2 - 15x + 20y - 25 = 0$$

Remove the $xy$ term by rotating the graph.

**Solution 1 |**




Suppose XY axes that have been rotated by $\varphi$ degree to remove $xy$ term.
Then we know that $\cot 2\varphi = \frac{A-C}{B} = \frac{64-36}{96} = \frac{7}{24}$.




$$\because \sqrt{24^2 + 7^2} = 25$$

By drawing the triangle, we can obtain $\cos 2\varphi = \frac{7}{25}$.
Using the half angle formulae from Mathematics I,

$$\cos\varphi = \sqrt{\frac{1+\cos 2\varphi}{2}} = \sqrt{\frac{1+\frac{7}{25}}{2}} = \sqrt{\frac{\frac{25+7}{25}}{2}} = \sqrt{\frac{16}{25}} = \frac{4}{5}$$

$$\sin\varphi = \sqrt{\frac{1-\cos 2\varphi}{2}} = \sqrt{\frac{1-\frac{7}{25}}{2}} = \sqrt{\frac{\frac{25-7}{25}}{2}} = \sqrt{\frac{9}{25}} = \frac{3}{5}$$

Thus,

$$x = X\cos\varphi - Y\sin\varphi = \frac{4}{5}X - \frac{3}{5}Y \text{ and } y = X\sin\varphi + Y\cos\varphi = \frac{3}{5}X + \frac{4}{5}Y$$

Plugging in,

$$64\left(\frac{4}{5}X-\frac{3}{5}Y\right)^2+96\left(\frac{4}{5}X-\frac{3}{5}Y\right)\left(\frac{3}{5}X+\frac{4}{5}Y\right)+36\left(\frac{3}{5}X+\frac{4}{5}Y\right)^2$$
$$-15\left(\frac{4}{5}X-\frac{3}{5}Y\right)+20\left(\frac{3}{5}X+\frac{4}{5}Y\right)-25=0$$

After more calculations,

$$\therefore 100X^2+25Y-25=0$$

## Determining the Type of Conic Section via General Formula

**Note | Recap**

$$Ax^2+Bxy+y^2+Dx+Ey+F=0$$

**Analysis | Seeking Approach**

- Since we know the translation, we know that $\exists p, q \in \mathbb{R}$ such that $A(x-p)^2+B(x-p)(x-q)+C(y-q)^2+F'$ can be obtained using translation without change in the type of conic section.
- Thus, we can imply that we can only focus at $A, B, C$.

**Theorem | Identification of conics**

Let $Ax^2+Bxy+y^2+Dx+Ey+F=0$ be a non-degenerate conic section.
Then, the type of the conics can be determined by discriminant $\Delta = B^2-4AC$:

$$\begin{cases} B^2-4AC<0: & \text{Ellipse} \\ B^2-4AC=0: & \text{Parabola} \\ B^2-4AC<0: & \text{Hyperbola} \end{cases}$$

**Proof | Textbook Approach**

Let an equation of a graph rotated $\varphi$ degree be

$$A'X^2+B'XY+C'Y^2+D'X+E'Y+F'=0$$

Knowing that $x = X\cos\varphi - Y\sin\varphi$ and $y = X\sin\varphi + Y\cos\varphi$,
We can show via calculation (comparing coefficient of each sides) that

$$B^2 - 4AC = B'^2 - 4A'C'$$

Let's suppose $\varphi$ that removes the XY term.
That case, The discriminant will be $\Delta' = -4A'C'$.
We already know how to determine the type when there are no XY term.
Using that methodology, we can easily show that the theorem holds for $\Delta'$.
Since $\Delta = \Delta'$, we can conclude that the theorem is true. $\therefore$ Q.E.D.

## Proof | Linear Algebra Aproach

Like we have shown on the analysis, first degree terms aren't important, so let's omit them. Let's utilize matrices. For rotated equation $A'X^2 + C'Y^2$ that lost the XY term,

$$Ax^2 + Bxy + Cy^2 = [x \;\; y]\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = A'X^2 + C'Y^2 = [X \;\; Y]\begin{bmatrix} A' & 0 \\ 0 & C' \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} \cdots (*)$$

### Notation | Rotation Matrix

Let the rotation matrix that rotates the graph with $\varphi$ amount be

$$R_\varphi = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

$$\text{So that } \begin{bmatrix} x \\ y \end{bmatrix} = R_\varphi \begin{bmatrix} X \\ Y \end{bmatrix}$$

### Analysis | Properties of Rotation Matirx

Let's think the rotation as a linear transform $f_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$.
Then, obviously $[f_\varphi] = R_\varphi$.
Intuitively, rotating $\varphi$ amount then $-\varphi$ amount will result in original state.

$$f_\varphi\left(f_{-\varphi}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)\right) = f_{-\varphi}\left(f_\varphi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Rightarrow R_\varphi R_{-\varphi}\begin{bmatrix} x \\ y \end{bmatrix} = R_{-\varphi}R_\varphi\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\therefore {R_\varphi}^{-1} = R_{-\varphi}$$

Let's think of a transpose of $R_\varphi$:

$$R_\varphi^T = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}^T = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{bmatrix} = R_{-\varphi}$$

$$\therefore {R_\varphi}^{-1} = R_\varphi^T = R_{-\varphi}$$

We know, from above notation and analysis,

$$\begin{bmatrix} x \\ y \end{bmatrix} = R_\varphi \begin{bmatrix} X \\ Y \end{bmatrix} \text{ and } [x \;\; y] = [X \;\; Y]R_\varphi^T = [X \;\; Y]{R_\varphi}^{-1}$$

$$\Rightarrow (*) = [X \;\; Y]{R_\varphi}^{-1}\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} R_\varphi \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow {R_\varphi}^{-1}\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} R_\varphi = \begin{bmatrix} A' & 0 \\ 0 & C' \end{bmatrix}$$

$$\Rightarrow \det\left({R_\varphi}^{-1}\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} R_\varphi\right) = \det\cancel{\left(R_\varphi {R_\varphi}^{-1}\right)}\det\left(\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}\right) = \det\left(\begin{bmatrix} A' & 0 \\ 0 & C' \end{bmatrix}\right)$$

$$\Rightarrow AC - \frac{B^2}{4} = A'C' \Rightarrow B^2 - 4AC = -4A'C'$$

$$\therefore \Delta = \Delta'$$

Next procedures are same as the textbook proof.

### Note | Trick for finding A',C'

Using the concepts introduced on linear algebra proof, we can find A',C' of rotated equation without XY term at ease without using trigonometries.

$$\begin{bmatrix} A - \lambda & \frac{B}{2} \\ \frac{B}{2} & A - \lambda \end{bmatrix}$$

Two $\lambda$ that makes the determinant of above matrix 0 is each A' and C'.

The proof exceeds the bound of precalculus, involving linear algebra concepts.

Article 10.06

# *Polar Equations of Conics*

- When the focus is on the origin

**THEOREM | Redefinition of conics**

Let $F$ be a fixed point, $l$ a fixed line, and $e > 0$ be a fixed positive number.
Then, the set of all points $P$ satisfying $d(P, F) = e \times d(P, l)$ is a conic section.
Moreover,
- if $e = 1$, then the conic section is a parabola.
- if $e < 1$, then the conic section is an ellipse.
  $\mapsto$ If then, $e$ is the eccentricity of the ellips
- When the focus is on the origin

**THEOREM | Redefinition of conics**

Let $F$ be a fixed point, $l$ a fixed line, and $e > 0$ be a fixed positive number.
Then, the set of all points $P$ satisfying $d(P, F) = e \times d(P, l)$ is a conic section.
Moreover,
- if $e = 1$, then the conic section is a parabola.
- if $e < 1$, then the conic section is an ellipse.
  $\mapsto$ If then, $e$ is the eccentricity of the ellipse.
- if $e > 1$, then the conic section is a hyperbola.

Also, $F$ is a focus of the conic.
In general, we call
- $F$ as a focus,
- $l$ as a directrix,
- $e$ as the eccentricity of the conics.

**Proof |**

Assume that $F(0)$, and let $l : x = d(d > 0)$.
For polar axis, $r = e(d - r\cos\theta)$.
The cartesian version is $x^2 + y^2 = e^2(d - x)^2$.
If $e = 1$, the curve is obviously a parabola. So lets assume that $e \neq 1$.

$$x^2 + y^2 = e^2x^2 + e^2d^2 - 2de^2x$$

$$\Rightarrow (1-e^2)x^2 + 2de^2x + y^2 = e^2d^2$$

$$\Rightarrow (1-e^2)\left(x^2 + \frac{2de^2}{1-e^2}x + \frac{d^2e^4}{(1-e^2)^2}\right) + y^2 = e^2d^2 + \frac{d^2e^4}{1-e^2} = \frac{e^2d^2}{1-e^2}$$

$$\Rightarrow \frac{\left(x + \frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{(1-e^2)^2}} + \frac{y^2}{\frac{e^2d^2}{1-e^2}} = 1$$

$$\Rightarrow a^2 = \frac{e^2d^2}{(1-e^2)^2} \text{ and } b^2 = \frac{e^2d^2}{1-e^2}$$

As a simplification, we have $\frac{(x-h)^2}{a} + \frac{y^2}{b^2} = 1$. Judging from the values of $b$, $e$ determines the shape of this figure.

### Analysis | Classification : $e < 1$

We have $\frac{(x-h)^2}{a^2} + \frac{y^2}{b^2} = 1$, where

$$h = -\frac{de^2}{1-e^2}, \quad a = \frac{ed}{1-e^2}, \quad b = \frac{ed}{\sqrt{1-e^2}}$$

and $a > b$ (wide horizontally).
The focal radius $c$ satisfies

$$c^2 = a^2 - b^2 = \frac{e^2d^2}{(1-e^2)^2} - \frac{e^2d^2}{1-e^2} = -h$$

$$c = \frac{e^2d}{1-e^2} = -h$$

Finally, the eccentricity is

$$\frac{c}{a} = \frac{\frac{e^2d}{1-e^2}}{\frac{ed}{1-e^2}} = e$$

### Analysis | Classification : $e > 1$

$$\frac{\left(x + \frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{(1-e^2)^2}} + \frac{y^2}{\frac{e^2d^2}{1-e^2}} = 1$$

Since $1-e^2<0$, substitute $1-e^2$ to $-(e^2-1)$ for simplification

$$\frac{\left(x-\frac{de^2}{e^2-1}\right)^2}{\frac{e^2d^2}{(e^2-1)^2}}-\frac{y^2}{\frac{e^2d^2}{e^2-1}}=1$$

Let $a=\frac{ed}{e^2-1}$, $b=f$

**Analysis | Classification : $e>1$**

$$\frac{\left(x+\frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{(1-e^2)^2}}+\frac{y^2}{\frac{e^2d^2}{1-e^2}}=1$$

Since $1-e^2<0$, substitute $1-e^2$ to $-(e^2-1)$ for simplification:

$$\frac{\left(x-\frac{de^2}{e^2-1}\right)^2}{\frac{e^2d^2}{(e^2-1)^2}}-\frac{y^2}{\frac{e^2d^2}{e^2-1}}=1$$

Let $a=\frac{ed}{e^2-1}$, $b=\frac{ed}{\sqrt{e^2-1}}$, $h=\frac{de^2}{e^2-1}$.
Then, $\frac{(x-h)^2}{a^2}-\frac{y^2}{b^2}=1$.

The focal radius $c$ satisfies:

$$c^2=a^2+b^2=\left(\frac{ed}{e^2-1}\right)^2+\left(\frac{ed}{\sqrt{e^2-1}}\right)^2$$

$$=\frac{e^2d^2}{(e^2-1)^2}+\frac{e^2d^2}{e^2-1}=\frac{e^2d^2+e^2d^2\cdot(e^2-1)}{(e^2-1)^2}=\frac{e^4d^2}{(e^2-1)^2}$$

$$\Rightarrow c=\frac{de^2}{e^2-1}=h$$

$$\frac{c}{a}=\frac{\frac{de^2}{e^2-1}}{\frac{ed}{e^2-1}}=\frac{\cancel{d}\frac{\cancel{e^2}}{\cancel{e^2-1}}}{\frac{\cancel{ed}}{\cancel{e^2-1}}}=e$$

## Polar Equations of Conics

**Theorem | Polar form of conics**

Let us return to the original parametric equation.

$$r=e(d-r\cos\theta)$$

$$\Rightarrow r + re\cos\theta = ed$$

$$r = \frac{ed}{1 + e\cos\theta}$$

Consequently,

$$r = \frac{ed}{1 + e\cos\theta}$$

is a polar equation of a conic with focus $(0, 0)$, directrix $x = d$, and eccentricity $e$.

For when $\sin$ and $\cos$ is substituted ($r = \frac{ed}{1+e\sin\theta}$), directrix becomes $y = d$.

### Example | Polar equation basics for conic sections (1)

Parabola, focus $(0, 0)$, directrix $y = -6$.
Thus $e = 1$. Find the polar equation.

**Solution | Solution**

Since the directrix is horizontal ($y = -6$), we use the form involving sine. The directrix is below the focus ($y = -6 < 0$), so the form is:

$$r = \frac{ed}{1 - e\sin\theta}$$

Here $e = 1$ and $d = 6$ (distance from focus to directrix).

$$\therefore r = \frac{1 \cdot 6}{1 - 1\sin\theta} = \frac{6}{1 - \sin\theta}$$

### Example | Polar equation basics for conic sections (2)

What is $r = \frac{10}{3-2\cos\theta}$?

**Solution | Solution**

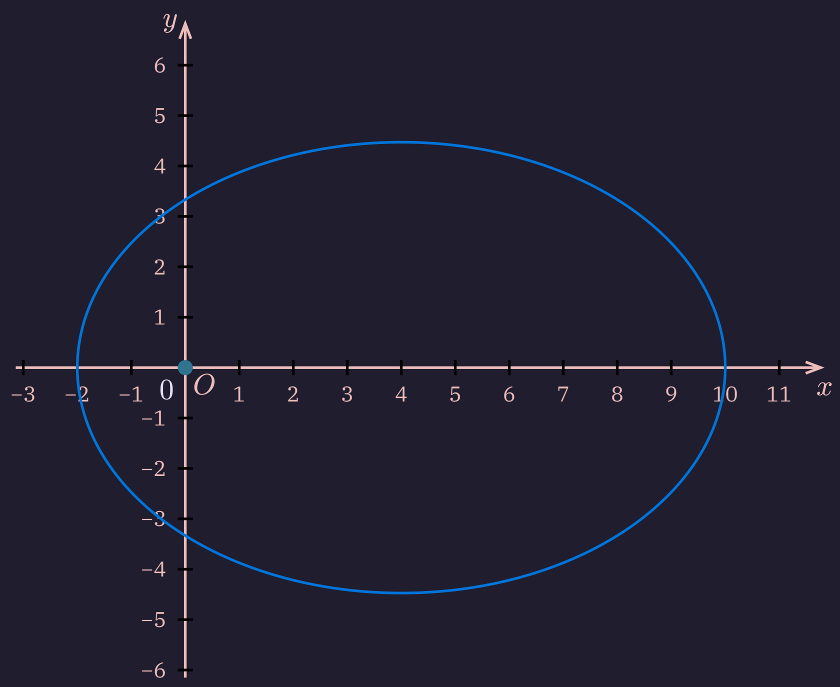
$$r = \frac{10}{3 - 2\cos\theta} = \frac{\frac{10}{3}}{1 - \frac{2}{3}\cos\theta} = \frac{\frac{2}{3}\cdot 5}{1 - \frac{2}{3}\cos\theta}$$

$$\Rightarrow e = \frac{2}{3}, \quad d = 5,$$

directrix $x = -5$. Since $e < 1$, this is an ellipse.

**Note | TIP**

Find $x, y$ intercepts before drawing. In this case, $r = 10, \frac{10}{3}, 2, \frac{10}{3}$ for $\theta = 0, \frac{\pi}{2}, \pi, 3\frac{\pi}{2}$, respectively.




## Example | Polar equation basics for conic sections (3)

$$r = \frac{12}{2 + 4\sin\theta} = \frac{2\cdot 3}{1 + 2\sin\theta}$$

Identify the conic section and find its properties.

### Solution | Solution

Rewrite in standard form:

$$r = \frac{12}{2(1 + 2\sin\theta)} = \frac{6}{1 + 2\sin\theta}$$

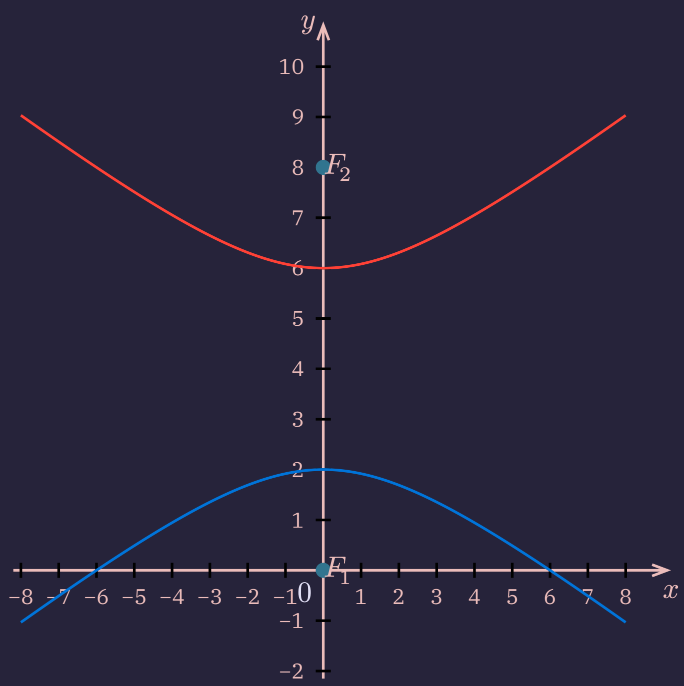
Comparing with $r = \frac{ed}{1+e\sin\theta}$:

$$e = 2, \quad ed = 6 \Rightarrow d = 3.$$

Since $e = 2 > 1$, this is a hyperbola. Directrix is $y = 3$ (horizontal, above focus).

Properties:

- Eccentricity $e = 2$.
- Directrix $y = 3$.
- Vertices:
  - $\theta = \frac{\pi}{2} \Rightarrow r = \frac{6}{1+2} = 2$. Vertex at $(2, \frac{\pi}{2})$ (polar) or $(0, 2)$ (Cartesian).
  - $\theta = 3\frac{\pi}{2} \Rightarrow r = \frac{6}{1-2} = -6$. Vertex at $(-6, 3\frac{\pi}{2})$ (polar) or $(0, 6)$ (Cartesian).
- Center is midpoint of vertices: $(0, 4)$.
- $a = 2$ (distance from center to vertex).
- $c = ae = 2(2) = 4$.
- $b = \sqrt{c^2 - a^2} = \sqrt{16 - 4} = \sqrt{12} = 2\sqrt{3}$.



$$r = \frac{ed}{1 + e\cos\theta}$$

equals the points that suffice $d(P, F) = e \cdot d(P, l)$. Here,

$$\begin{cases} e = 1 \text{ parabola} \\ e > 1 \text{ hyperbola} \\ e < 1 \text{ ellipse} \end{cases}$$

, where they are all confocal.
In the end, if $e \gg 1$, the parabola approaches a line.

### Example | Polar equation basics for conic sections (4)

What shape does the following expression represent?

$$r = \frac{10}{3 - 2\cos(\theta - \frac{\pi}{4})}$$

#### Solution | Solution

The term $\cos(\theta - \frac{\pi}{4})$ represents a rotation of the polar graph by $\frac{\pi}{4}$ (45 degrees) counterclockwise. Let's analyze the base equation without rotation:

$$r' = \frac{10}{3 - 2\cos\theta} = \frac{\frac{10}{3}}{1 - \frac{2}{3}\cos\theta}$$

Here, $e = \frac{2}{3} < 1$, so it is an ellipse. $ed = \frac{10}{3} \Rightarrow \left(\frac{2}{3}\right)d = \frac{10}{3} \Rightarrow d = 5$. Directrix $x = -5$.

Thus, the given equation represents an **ellipse** with eccentricity $\frac{2}{3}$, rotated by $45^{\circ}$ counterclockwise about the focus (origin).

**Confocal**
e.
- if $e > 1$, then the conic section is a hyperbola.

Also, $F$ is a focus of the conic.
In general, we call
- $F$ as a focus,
- $l$ as a directrix,
- $e$ as the eccentricity of the conics.

**Proof |**

Assume that $F(0)$, and let $l : x = d(d > 0)$.
For polar axis, $r = e(d - r\cos\theta)$.
The cartesian version is $x^2 + y^2 = e^2(d - x)^2$.
If $e = 1$, the curve is obviously a parabola. So lets assume that $e \neq 1$.

$$x^2 + y^2 = e^2x^2 + e^2d^2 - 2de^2x$$

$$\Rightarrow (1 - e^2)x^2 + 2de^2x + y^2 = e^2d^2$$

$$\Rightarrow (1 - e^2)\left(x^2 + \frac{2de^2}{1 - e^2}x + \frac{d^2e^4}{(1 - e^2)^2}\right) + y^2 = e^2d^2 + \frac{d^2e^4}{1 - e^2} = \frac{e^2d^2}{1 - e^2}$$

$$\Rightarrow \frac{\left(x + \frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{(1-e^2)^2}} + \frac{y^2}{\frac{e^2d^2}{1-e^2}} = 1$$

$$\Rightarrow a^2 = \frac{e^2d^2}{(1 - e^2)^2} \text{ and } b^2 = \frac{e^2d^2}{1 - e^2}$$

As a simplification, we have $\frac{(x-h)^2}{a} + \frac{y^2}{b^2} = 1$. Judging from the values of $b$, $e$ determines the shape of this figure.

### Analysis | Classification : $e < 1$

We have $\frac{(x-h)^2}{a^2} + \frac{y^2}{b^2} = 1$, where

$$h = -\frac{de^2}{1 - e^2}, \quad a = \frac{ed}{1 - e^2}, \quad b = \frac{ed}{\sqrt{1 - e^2}}$$

and $a > b$ (wide horizontally).
The focal radius $c$ satisfies

$$c^2 = a^2 - b^2 = \frac{e^2d^2}{(1 - e^2)^2} - \frac{e^2d^2}{1 - e^2} = -h$$

$$c = \frac{e^2d}{1 - e^2} = -h$$

Finally, the eccentricity is

$$\frac{c}{a} = \frac{\frac{e^2d}{1-e^2}}{\frac{ed}{1-e^2}} = e$$

### Analysis | Classification : $e > 1$

$$\frac{\left(x+\frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{\left(1-e^2\right)^2}}+\frac{y^2}{\frac{e^2d^2}{1-e^2}}=1$$

Since $1-e^2<0$, substitute $1-e^2$ to $-\left(e^2-1\right)$ for simplification

$$\frac{\left(x-\frac{de^2}{e^2-1}\right)^2}{\frac{e^2d^2}{\left(e^2-1\right)^2}}-\frac{y^2}{\frac{e^2d^2}{e^2-1}}=1$$

Let $a=\frac{ed}{e^2-1}$, $b=f$

**Analysis | Classification : $e>1$**

$$\frac{\left(x+\frac{de^2}{1-e^2}\right)^2}{\frac{e^2d^2}{\left(1-e^2\right)^2}}+\frac{y^2}{\frac{e^2d^2}{1-e^2}}=1$$

Since $1-e^2<0$, substitute $1-e^2$ to $-\left(e^2-1\right)$ for simplification:

$$\frac{\left(x-\frac{de^2}{e^2-1}\right)^2}{\frac{e^2d^2}{\left(e^2-1\right)^2}}-\frac{y^2}{\frac{e^2d^2}{e^2-1}}=1$$

Let $a=\frac{ed}{e^2-1}$, $b=\frac{ed}{\sqrt{e^2-1}}$, $h=\frac{de^2}{e^2-1}$.
Then, $\frac{(x-h)^2}{a^2}-\frac{y^2}{b^2}=1$.

The focal radius $c$ satisfies:

$$c^2=a^2+b^2=\left(\frac{ed}{e^2-1}\right)^2+\left(\frac{ed}{\sqrt{e^2-1}}\right)^2$$

$$=\frac{e^2d^2}{\left(e^2-1\right)^2}+\frac{e^2d^2}{e^2-1}=\frac{e^2d^2+e^2d^2\cdot\left(e^2-1\right)}{\left(e^2-1\right)^2}=\frac{e^4d^2}{\left(e^2-1\right)^2}$$

$$\Rightarrow c=\frac{de^2}{e^2-1}=h$$

$$\frac{c}{a}=\frac{\frac{de^2}{e^2-1}}{\frac{ed}{e^2-1}}=\frac{\cancel{d}\frac{e\cancel{^2}}{\cancel{e^2-1}}}{\frac{\cancel{ed}}{\cancel{e^2-1}}}=e$$

## Polar Equations of Conics

**Theorem | Polar form of conics**

Let us return to the original parametric equation.

$$r = e(d - r\cos\theta)$$

$$\Rightarrow r + re\cos\theta = ed$$

$$r = \frac{ed}{1 + e\cos\theta}$$

Consequently,

$$r = \frac{ed}{1 + e\cos\theta}$$

is a polar equation of a conic with focus $(0, 0)$, directrix $x = d$, and eccentricity $e$.

For when $\sin$ and $\cos$ is substituted ($r = \frac{ed}{1+e\sin\theta}$), directrix becomes $y = d$.

### Example | Polar equation basics for conic sections (1)

Parabola, focus $(0, 0)$, directrix $y = -6$.
Thus $e = 1$.

$$\therefore r = \frac{e \cdot d}{1 - e\sin\theta} = \frac{6}{1 - \sin\theta}$$

### Example | Polar equation basics for conic sections (2)

What is $r = \frac{10}{3-2\cos\theta}$?

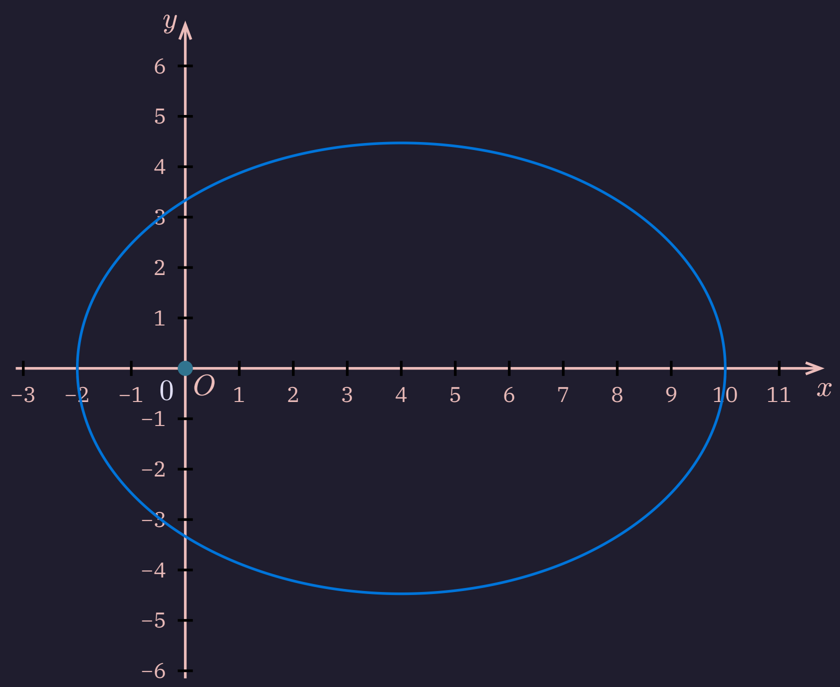
$$r = \frac{10}{3 - 2\cos\theta} = \frac{\frac{10}{3}}{1 - \frac{2}{3}\cos\theta} = \frac{\frac{2}{3} \cdot 5}{1 - \frac{2}{3}\cos\theta}$$

$$\Rightarrow e = \frac{2}{3}, \quad d = 5,$$

directrix $x = -5$. Since $e < 1$, this is an ellipse.

#### Note | TIP

Find $x, y$ intercepts before drawing. In this case, $r = 10, \frac{10}{3}, 2, \frac{10}{3}$ for $\theta = 0, \frac{\pi}{2}, \pi, 3\frac{\pi}{2}$, respectively.




**Example | Polar equation basics for conic sections (3)**

$$r = \frac{12}{2 + 4\sin\theta} = \frac{2 \cdot 3}{1 + 2\sin\theta}$$

$$\Rightarrow e = 2, d = 3,$$

directrix $y = 3$. Since $e > 1$, this is a hyperbola.

$$a = 2, \quad c = 4, \quad b = 2\sqrt{3}$$

($e = \frac{c}{a}$ also stands for all conics.)

Asymptotes when the denominator approaches to 0. > The slope of asymptote is $\tan\theta$ of $\theta$ for when $1 \pm e\sin\theta$ or $\cos\theta = 0$.

$$r = \frac{ed}{1 + e\cos\theta}$$

equals the points that suffice $d(P, F) = e \cdot d(P, l)$. Here,

$$\begin{cases} e = 1 \text{ parabola} \\ e > 1 \text{ hyperbola} \\ e < 1 \text{ ellipse} \end{cases}$$

, where they are all confocal.
In the end, if $e \gg 1$, the parabola approaches a line.

**Example | Polar equation basics for conic sections (4)**

What shape does the following expression represent?

$$r = \frac{10}{3 - 2\cos(\theta - \frac{\pi}{4})}$$

**Solution | Solution**

Counterclockwise 45° to $r = \frac{10}{3 - 2\cos\theta}$

$$\Rightarrow r = \frac{\frac{2}{3} \cdot 5}{1 - \frac{2}{3}\cos\theta}$$

**Confocal**

Chapter 14

# Probabilities & Statistics

*In our exploration of mathematical models thus far, we have primarily dealt with deterministic systems—where a given input inevitably yields a specific, calculated output. However, the world we inhabit is rarely so predictable; it is governed by uncertainty, variability, and chance. This chapter transitions us from the absolute certainty of algebraic functions to the stochastic nature of Probability and Statistics. Here, we will develop the analytical tools necessary to quantify uncertainty, calculate the likelihood of future events, and interpret complex datasets to uncover hidden patterns. By mastering concepts such as combinatorics, probability distributions, and statistical analysis, we bridge the gap between theoretical mathematics and practical application, empowering us to make data-driven decisions in fields ranging from quantum mechanics and engineering to economics and social policy.*

Article 14.01

# *Counting*

## Some background info

Where is counting included?

- Discrete Mathematics (tr. **이산수학**)
  - A branch of mathematics dealing with finite sets & calculations
- Combinatorics (tr. **조합론**)

**Notation | [n]**

$[15] = \{1, 2, 3, ..., 15\}$
Or, to generalize:

$$[n] = \{1, 2, ..., n\}$$

**Definition | Identical Size**

We can define two sets $A$ and $B$ have **identical size** iff there exists a bijective function $F : A \to B$.

Thus, the number of elements on set $S$ is $n$ means that there exists a bijective function $F : S \to [n]$ between the set $S$ and $[n]$.

**Note | Dice**

Set of possible outcomes by throwing 2 dice have **identical size** with:

$$[6] \times [6] = \{(i, j) \mid i, j \in [6]\}$$

## The Fundamental Counting Principle

**Definition | Principle**

Suppose that two events occur in order. If the first event can occur in $m$ ways and the second in $n$ ways, then two events can occur in $m \times n$ ways.

It can be easily shown that for the $k$-th event $E_k$ and its number of possible outcomes $n_k$, the entire events have total outcomes:

$$\prod_{i=1}^{k} n_k$$

**Example | Isn't it a bit too hot?**

The ice cream vendor has 3 types of cones and 30 types of flavors. How many types of finished ice cream cones can the vendor serve?

**Solution 1 |**

We apply the Fundamental Counting Principle.
- Step 1: Choose a cone. There are 3 options.
- Step 2: Choose a flavor. There are 30 options.

Since these choices are independent and made in sequence, the total number of combinations is the product of the number of options for each step:

$$3 \times 30 = 90 \text{ types.}$$

**Example | Rationals**

For $i, j \in [6]$, find the number of possible rationals of the form $\frac{i}{j}$.

**Note | Hint**

Simply trying $6 \times 6$ won't work here, as the function $[6] \times [6] \to S$ is not bijective.

**Note | Important note: Fools**

If your counting problem reaches a problem, consider the following:
- You messed up too much (one to one but not onto)
- You messed up too less (not one to one but onto)

**Definition | Permutation (tr. 순열)**

A permutation of a set is an ordering of the elements of the set.

**Equation |**

For $X$ that $|X| = n$:

- Selecting $1^{\text{st}}$ element: $n$ possibilities
- Selecting $2^{\text{nd}}$ element: $n-1$ possibilities
- Selecting $3^{\text{rd}}$ element: $n-2$ possibilities

$$\vdots$$

By the Fundamental counting principle, there are in total $n(n-1)(n-2)\cdots 2 \cdot 1 = n!$ ways of orderings.

**Definition | Alternative Definition**

1:1 onto function that $[r] \to [n]$

**Definition | Permutation taken r at a time**

A permutation for $[n]$ taken $r$ at a time. Selecting $r$ and give them order.

In the same way, total ways:

$$\frac{n!}{(n-r)!}$$

**Notation | Permutations**

We note permutations as $nPr$ or $P(n, r)$ or ${}_nP_r$.

**Definition | Combination taken r at a time**

A combination for $[n]$ taken $r$ at a time. Selecting $r$ WITHOUT order.

In the same way, total ways:

$$\frac{n!}{r!(n-r)!}$$

This equals the number of subsets of $[n]$ such that have $r$ elements.

**Notation | Combinations**

We note combinations as $nCr$, $C(n,r)$ or $\binom{n}{r}$.

**Example | Car number plate**

3 alphabets (distinct), 3 digits. Possible number of license plates?

**Solution 1 |**

Choose 3 distinct alphabets out of 26 with order: $P(26,3) = 26 \times 25 \times 24$
Choose 3 digits out of 0 9: $10 \times 10 \times 10$
Therefore, the final count becomes $62 \times 25 \times 24 \times 10 \times 10 \times 10 = 1560000$.

**Example | Committee**

Choose a chairman, a vice chairman, a secretary (named, order matters) and four other committee members (no order). Find the number of ways to choose the committee members.

**Solution 1 |**

Selecting 3 named committee members: $P(20,3)$
Result: $P(20,3) \times C(17,4)$

**Sol 2:** Pick 7 with order and divide by $4!$.

## Other Counting Methods

### Permutation with repetition

- No. of permutation with repetition of $n$ objects taken $r$ at a time:
- $$\underbrace{n\times n \times \cdots \times n}_{r} = n^r$$
- Number of **every** function $[r] \to [n]$
- (tr. **중복순열**)

### Combination with repetition

- Number of combinations with repetition of $n$ objects taken $r$ at a time.
- (tr. **중복조합**)

- Set (no order) where it can have multiple identical elements: **multiset**.
  - Combination with repetition have a one-to-one relationship with multiset.

**Notation | Combination with Repetitions**

$$H(n, r) = {}_nH_r = C(n + r - 1, r)$$

**Equation | Stars and Bars**

Put $r$ objects to $n$ boxes. Objects aren't distinguishable whereas boxes are. Let's think of boxes divided by $|$ and objects represented as $*$

$$|\underbrace{*}_{\text{object}}|*|**|\underbrace{\ }_{\text{box}}|$$

Then, the methods to put objects to boxes have one-to-one relationship with ordering $|$ and $*$.

$\therefore H(n, r) =$ number of ordering $n - 1$ of $|$ and $r$ of $*$
$= C(n + r - 1, r)$

**Example | 3 boxes, 5 balls**

Find the number of ways to distribute 5 balls into 3 boxes.

**Solution | Case 1: Distinguishable balls, Distinguishable boxes**

Each of the 5 balls can be placed into any of the 3 boxes.
- Ball 1: 3 choices
- Ball 2: 3 choices
- ...
- Ball 5: 3 choices

Total ways: $3 \times 3 \times 3 \times 3 \times 3 = 3^5 = 243$. (This is permutation with repetition).

**Solution | Case 2: Indistinguishable balls, Distinguishable boxes**

Since the balls are identical, only the number of balls in each box matters. Let $x_1, x_2, x_3$ be the number of balls in box 1, 2, and 3 respectively. We need to find the number of non-negative integer solutions to:

$$x_1 + x_2 + x_3 = 5$$

Using the Stars and Bars formula (Combination with repetition):

$$H(3,5) = C(3+5-1,5) = C(7,5) = C(7,2) = \frac{7 \times 6}{2} = 21.$$

**Solution | Case 3: Indistinguishable balls, Indistinguishable boxes**

Since both balls and boxes are identical, we are looking for the number of ways to partition the integer 5 into at most 3 parts (order of boxes doesn't matter). We list the possible partitions (formatted as counts in the 3 boxes):

1. (5, 0, 0) - All in one group
2. (4, 1, 0)
3. (3, 2, 0)
4. (3, 1, 1)
5. (2, 2, 1)

Total ways: 5.

## Distinguishable Permutations

- $\{1\}\ \{2\}\ \{3\}\ \{4\}\ \{5\}$
  - $5! = 120$ ways
- Imagine all balls were painted (now all identical)
  - 1 way
- What if some were painted and some were not?

**Equation |**

$n$ objects, $k$ types.
$n_i$ objects for type $i$ $(1 \leq i \leq k, i \in \mathbb{N})$

$$\sum_{i=1}^{k} n_i = n$$

---

**Solution 1 | F.C.P.**

Number of ways to decide the location of objects in type $i$:

$$C\left(n-\sum_{j=1}^{i} n_j, n_i\right)$$

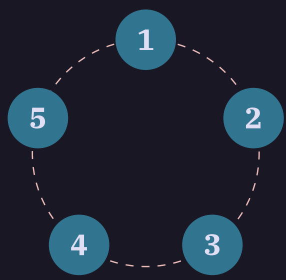
$$\Rightarrow \text{total ways}: \frac{n!}{n_{1!}(n-n_1)!}\cdot\frac{(n-n_1)!}{n_{2!}(n-n_1-n_2)!}\cdots\frac{n_{k!}}{n_{k!}}$$

$$=\frac{n!}{n_{1!}n_{2!}n_{3!}\cdots n_{k!}}$$

**Solution 2 |**

Ways when omitting indistinguishability: $n!$
Ways to change location between same types: $n_{1!}n_{2!}n_{3!}\cdots n_{k!}$

### Circular Permutation




- Ways of 4 people sitting around the round table: $4!$
- If we are only considering relative position of people, $\frac{4!}{4} = 6$
- Circular permutation of $n$: $\frac{n!}{n} = (n-1)!$

**Free circular permutation**

- Can be flipped (think of it as a necklace)
- (tr. **염주순열**)
- Free circular permutation of $n$: $\frac{n!}{n}\cdot\frac{1}{2} = \frac{(n-1)!}{2}$

## Ex) Burnside

- $X$: ordering linearly
- $G$: ordering as free circular permutation (In fact, group)

**Definition | Orbit**

Orbit of $x \in X$, denoted $O_x$. Set of every $y$ such that can be obtained by operation of $G$ to $x$.

- If all distinguishable: orbit = $2 \times \text{number of bids}$
- If all identical: orbit = 1
- More the elements that can be obtained by operation of $G$, less the orbit become.
  - $|O_x| = \frac{|G|}{|S_x|}$
  - $S_x$: number of $g \in G$ such that $x$ become $x$ when applied. (return to initial condition)
  - *(How many $g$ needed to return to original? IDK I don't think this is right)*
- $$\sum |F_g|$$
  - $F_g$: number of $a \in X$ such that don't change by $g$
  - $$\sum_{x \in X} |S_x| = \sum_{g \in G} |S_x| = |G| \sum_{x \in X} \frac{1}{|O_x|} = |G| \cdot \text{number of orbits}$$
  - Number of orbits = $\frac{1}{|G|} \sum_{g \in G} |F_g|$

Article 14.02
# *Probability*

## What is probability?

- Ratio of certain result when event was repeated for large number?
  - Coin throwing: what if 4987 heads for 10000 throw?
  - We can't ever prove it.

**NOTE | Learn More**

Fun fact, the universe is composed up of only uncertain knowledge; as the total sum of knowledge never increases. Only deductive reasoning proves 100% correct numbers, but it is always a subset of known information. Induction or analogy creates seemingly new information, but it has the chance of being uncertain.

## Three elements of probabilistic models

- Model to explain phenomenon that can't be predicted
  - Limit of measurement
  - **Heisenberg** Uncertainty Principle

### 1. Sample Space

**DEFINITION | Sample Space**

**Sample Space** $S$ is defined as the set including all **possibilities**. For instance, the sample Space of throwing a coin will be:

$$S = \left\{ \underbrace{H}_{\text{heads}}, \underbrace{T}_{\text{tails}} \right\}$$

- Throwing the coin; **experiment** can be known as dissipation of uncertainty.
  - This can also be said as selecting one element of $S$.

### 2. Events

**DEFINITION | Event**

An **event** is a case where the probability function breaks down due to an **experiment**, resulting in a deterministic event which is an element of the sample space.

It can be said that the **event is a subset of the sample space**.

Take a die for example.
- **Event where 1 comes out :** $\{1\}$
- **Event where even comes out :** $\{2, 4, 6\}$
- **Event where number larger than 4 comes out :** $\{5, 6\}$
- **Event where positive integer comes out :** $\{1, 2, 3, 4, 5, 6\}$ **(Certain Event)**
- **Event where 7 comes out :** $\varphi$ **(Null Event)**

From this, we can note that an event has a one-to-one correspondence with the set of elements within $S$ that trigger the event. Thus, we can define event as a subset of the sample space. To say otherwise, when a $s \in S$ is selected, for event $E$:

$$\begin{cases} s \in E : E & \text{has occured} \\ s \notin E : E & \text{has not occured} \end{cases}$$

## 3. Probability

- There must be a properties that probability must have.
  - The sum must be 1.

### Definition | Probability

Probability $P(E) \in [0, 1]$ is a number that somehow(...) represents the **likeliness** of event $E$ and satisfies the essential properties.

### Equation | Identical likeliness

If every element of $S$ is equally likely and $S$ is a finite set, for an event $E$,

$$P(E) = \frac{n(E)}{n(S)} = \frac{|E|}{|S|}$$

### Example | Throwing 2 coins

Benjamin has two coins. For simplicity, call them $A$ and $B$. What is the sample space for all possible faces? What is the sample space for the number of heads? Calculate the probability from our equation in the definition of probability. Is there a difference? Why? Explain your answer.

**Solution 1 |**

$S_1 = \{HH, HT, TH, TT\}$

$$P(E)_1 = \frac{n(E)}{n(S_1)} = \frac{1}{4}$$

On the other hand,

$$S_2 = \{0, 1, 2\}$$

(number of heads)

$$P(E)_2 = \frac{n(E)}{n(S_2)} = \frac{1}{3}$$

Why is there a difference? $P(E)_2$ is **wrong** because in $S_2$, the probabilities of each element have different "**likeliness**".

### Example | Poker flush

A flush is a five-card hand with the same suit. What is the probability that we have a flush?

**Solution 1 |**

$S =$ Set of every possible five-card hand
$E =$ Event of having same suit

$$P(E) = \frac{n(E)}{n(S)} = \frac{4 \times C(13, 5)}{C(52, 5)} \approx 0.20\%$$

### Note | Dice Dice

Throwing 2 dice
$S = [6] \times [6] = \{(i, j) \mid i, j \in [6]\}$
Event $E$: sum of 2 dice is $5$
$E = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$

> **e.g.**
> $(2,3) \in E$: $E$ occurred
> $(3,4) \notin E$: $E$ not occurred.

### The complement of an event $E$

- (tr. 여사건)

**Definition | Complement**

**The complement of an event $E$** is defined as $E^c = E' = S - E$.

Written in set builder form, it becomes

$$E^c = \{c \mid c \in S \land c \notin E\}$$

It also can be said that $E^c$ is an event where $E$ does not occur.

- $P(E^c) = 1 - P(E)$
  - $\because n(E^c) = n(S) - n(E)$

### Unions and Intersections of Probability

**Definition | Unions and Intersections**

The **union** (tr. **합사건**) of $E$ and $F$ is written as $E \cup F$ and is defined as the event where $E$ or $F$ occur.
The **intersection** (tr. **곱사건**) of $E$ and $F$ is written as $E \cap F$ and is defined as the event where $E$ and $F$ occur.

We can mathematically define $P(E \cup F) = P(E) + P(F) - P(E \cap F)$. Here, if $E \cap F = \varphi$, then $P(E) + P(F)$. This shows a case where **$E$ and $F$ cannot occur mutually**, and is therefore called a **mutually exclusive event**.

**Definition | Mutually Exclusive Events**

If $E \cap F = \varphi$, we say that $EF$ are **mutually exclusive** (tr. **배반사건**).

### Conditional Probability

- When we observed information partially, it results in a change in probability.
  - The quantity of this change in probability is called **Conditional Probability**.

### Analysis | Two dice

$A$: Event where sum of two dice is greater or equal than $8$.
$B$: Event where $2$ is observed for the first dice.

If $B$, possible elements of $S$ that can be selected have reduced to $B$. The possible outcome that occur $A$ is $(2, 6)$ only.

Probability where $A$ occur after $B$:

$$\frac{\text{number of possibilities that triggers the event}}{\text{number of every possibilities}} = \frac{n(A \cap B)}{n(B)} = \frac{P(A \cap B)}{P(B)}$$

### Definition | Conditional Probability

Let $A, B$ be events. Then, the conditional probability of $A$ given that $B$ has occurred is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Here, we must note that $P(B) \neq 0$. Since conditional probability **assumes $B$ has occurred**, a probability of $0$ is paradoxical.

- For events $A, B$, if probability of $A$ does not change even after knowing $B$, it can be said that $A$ and $B$ are independent.

### Definition | Independent

$A, B$ are independent when:

$$P(A|B) = P(A)$$

or

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(B) \Rightarrow P(A \cap B) = P(A)P(B)$$

### Note | Two dice (Independence vs Dependence)

$A_k$: first dice is $k$
$B_k$: second dice is $k$
$C_k$: sum of two dice is $k$

**Solution | Independence**

$P(A_k) = \frac{1}{6}$
$P(B_k) = \frac{1}{6}$
$P(A_k \cap B_{k'}) = \frac{1}{36} = P(A_k)P(B_{k'})$
Thus $A_k$ and $B_{k'}$ are **independent**.

**Solution | Dependence**

$P(C_5) = \frac{1}{9}$
$P(A_3 \cap C_5) = \frac{1}{36}$
$P(A_3 \mid C_5) = \frac{P(A_3 \cap C_5)}{P(C_5)} = \frac{\frac{1}{36}}{\frac{1}{9}} = \frac{1}{4}$

### The Law of Total Probability

**Definition | Law of Total Probability**

Let $F_1, \ldots, F_n$ be events such that $F_i \cap F_j = \varphi$ and $\bigcup_{k=1}^{n} F_k = S$.

Then, for an event $E$,

$$P(E) = \sum_{k=1}^{n} P(E|F_k)P(F_k)$$

**Proof |**

$$\sum_{k=1}^{n} P(E \mid F_k)P(F_K)$$

$$= \sum_{k=1}^{n} P(E \cap F_k)$$

$$= P\left(\bigcup_{k=1}^{n} (E \cap F_k)\right)$$

$$= P\left(E \cap \left(\bigcup_{k=1}^{n} F_k\right)\right)$$

$$= P(E \cap S) = P(E)$$

$\therefore$ Q.E.D.

**EXAMPLE | Vaccines**

There is a pandemic! We are going to vaccinate people. People have received different doses of vaccines. 10% of the population has received no vaccines, with an infection rate of 5%. 40% of the population received one dose of the vaccine and had an infection rate of 0.5%. 50% of the population had two doses and had a 0.1% infection rate.

Q : What is the total infection rate?

**Solution 1 |**

Let $E$ be the event that someone is infected. let $F_0$ be the event where somebody has 0 doses of vaccination. Define similarly for $F_1$ and $F_2$.
$\therefore P(E) = P(E|F_0)P(F_0) + P(E|F_1)P(F_1) + P(E|F_2)P(F_2) = 0.75\%$

Q: Given that a person is infected, what is the (conditional) probability that they were not vaccinated?

**Solution 2 |**

We want to calculate $P(F_1 \mid E)$. This is equal to :

$$\frac{P(E \cap F_1)}{P(E)} = \frac{P(E|F_1)P(F_1)}{P(E)} = \frac{P(E|F_1)P(F_1)}{\sum_{k=1}^{n} P(E|F_K)P(F_K)}$$

Plugging in the values, we can find that $\frac{2}{3}$ of the total infected was not vaccinated.

The formula we have derived in this example is directly linked with the **Bayes' formula** below.

**Bayes' Formulae**

**EQUATION |**

$$P(F_1|E) = \frac{P(E \cap F_1)}{P(E)} = \frac{P(E|F_1)P(F_1)}{P(E)} = P(E|F_1)\frac{P(F_1)}{\sum_{k=1}^{n} P(E|F_k)P(F_k)}$$

$$\therefore P(F_1|E) = \frac{P(E|F_1)P(F_1)}{\sum_{k=1}^{n} P(E|F_k)P(F_k)}$$

### Example | Drug tests

We are testing for drugs. Every drug test has an error.
There are two types :

- False Positive(Shows true when the tester did not take drugs, **1%**)
- False Negative(Shows false when the tester did take drugs, **1%**).

The actual drug usage rate is 0.5%. When my test results show positive, under what possibility have I actually taken drugs?

#### Solution 1 |

Let $E$ be the event where my test shows positive and $F$ be the event where I actually took drugs.
Using the Bayes' formula,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^c)P(F^c)}$$

$$= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} \approx \frac{1}{3}$$

### Example | The Monti-Hall dilemma

The Monti-Hall Dilemma was a very famous TV show from Canada. The rules are the following :

1. There are three doors. One with a car and two with goats.
2. Participant choose one of the three doors.
3. The instructor revels one of the remaining two doors which have goat behind it.
4. Participant can change his/her choice after seeing.

Find the probability of winning if the participant switches.

#### Solution |

Let

- $A$ := There is a car behind the door participant had selected
- $B_k$ := There is a car behind $\text{k}^{\text{th}}$ door.
- $C_k$ := $\text{k}^{\text{th}}$ door is opened by rule 4

$$P(A|C_1) = \frac{P(C_1|A)P(A)}{P(C_1|A)P(A) + P(C_1|B_1)P(B_1) + P(C_1|B_2)P(B_2)}$$

## Exercises

**Example | Problem 1**

A digital thermometer is placed in the classroom. It shows the temperature between 10 40 in integer. We read the temp at a particular moment.
The class dismisses if the temperature is higher than 35 or lower than 15.

1. What is the sample space?

**Solution 1 |**

The sample space $S$ consists of all possible integer temperature readings from 10 to 40 inclusive.

$$S = \{10, 11, 12, ..., 40\}$$

The size of the sample space is $40 - 10 + 1 = 31$.

1. What is the event that the class dismisses (as a mathematical object)?

**Solution 2 |**

The event $E$ corresponds to temperatures $T$ such that $T < 15$ or $T > 35$.

$$E = \{10, 11, 12, 13, 14, 36, 37, 38, 39, 40\}$$

The size of the event is $|E| = 5 + 5 = 10$.

1. Is the probability that the class dismisses equal to $\frac{10}{31}$? Why or why not?

**Solution 3 |**

No, not necessarily. The calculation $P(E) = |E|S| = \frac{10}{31}$ assumes that every temperature outcome is **equally likely** (uniform distribution). In reality, temperatures like 20-25 are much more likely than 10 or 40. Without knowing the probability distribution of the temperature, we cannot simply count outcomes.

## Example | Problem 2

A **straight** is a hand of five cards consisting of five consecutive numbers. (e.g. 4,5,6,7,8) The Jacks are counted as 11, Queens 12, and Kings 13. The Aces are considered as 14 or 1, but not both. Thus, A,K,Q,J,10 and A,2,3,4,5 are straights, but Q,J,K,A,2 is not a straight because the Ace cannot be both 14 and 1 at the same time.

Find the probability that a five-card hand is a straight. (Include straight flushes in your count).

### Solution 1 |

1. **Count total hands:** $C(52, 5)$.
2. **Count straights:**
    - There are 10 possible sequences of ranks: (A,2,3,4,5), (2,3,4,5,6), ..., (10,J,Q,K,A).
    - For each rank in a sequence, there are 4 possible suits.
    - Since there are 5 cards, there are $4 \times 4 \times 4 \times 4 \times 4 = 4^5$ ways to choose suits for a given rank sequence.
    - Total number of straights = $10 \times 4^5$.

$$P(\text{Straight}) = \frac{10 \times 4^5}{C(52, 5)} = \frac{10 \times 1024}{2598960} \approx 0.00394 \approx 0.4\%$$

## Example | Problem 3

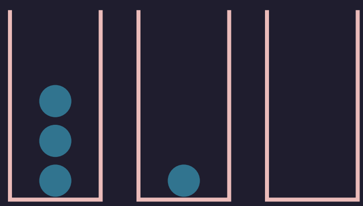
You are given four balls and three boxes. For each ball, choose a box at random and put the ball into that box.
What is the probability that no box is empty?

**Solution 1 |**

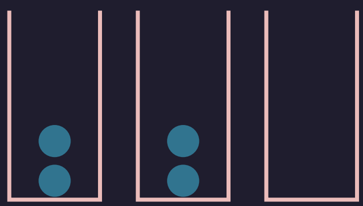
Balls and boxes must be both distinctive.

**Proof | Justification**

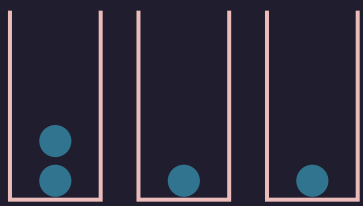
Configuration A



Configuration B



Configuration C



Configuration D

Thus, the total probability is $3^4 = 81$
It is equivalent of finding all onto function $[3] \to [4]$

$$= 3^4 - 2^4 \binom{3}{2} + 1^4 \binom{3}{1} = 81 - 48 + 3 = 36$$

$$\therefore \frac{36}{81} = \frac{4}{9}$$

## Example | Problem 4

Let A and B be independent events with P(A) = 0.4, P(B) = 0.7. Find the probability that precisely one of A or B occurs.

**Solution 1 |**

Since the probability is $P(A \cup B - A \cap B)$, $P(A \cup B - A \cap B) = P(A) + P(B) - 2 \times P(A \cap B) = 0.4 + 0.7 - 2 \times 0.4 \times 0.7 = 0.54$

**Example | Problem 5**

In winter, when people leave their houses for a long time, they open their water taps to make water flow so that the pipe does not freeze.

Before I leave for a vacation, I decide to open the water tap.
The probability that the pipe freezes is 5% when I open the tap.
The probability that the pipe freezes is 25% when I forget to open the tap.
The probability that I forget is 30%.

**Solution 1 | What is the probability that the tap actually freezes?**

Let $E$ be event where pipe freezes.
Let $F$ be event where I forget.
By the total probability law,

$$P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$$

$$= 5\% \times 70\% + 25\% \times 30\%$$
$$= 3.5\% + 7.5\% = 11\%$$

$$\therefore 11\%$$

**Solution 2 | Given that the tap froze, what is the probability that I had forgotten to open the tap?**

Finding $P(F^c|E)$

$$= \frac{P(E|F^c) \cdot P(F^c)}{P(E)} = \frac{7.5\%}{3.5\% + 7.5\%} = \frac{15}{22}$$

**Example | Problem 6**

In a computer RPG game, your hero passes through an area where a goblin or an orc randomly appear independently. The probability of encountering a goblin is $P(G) = 0.3$. The probability of encountering an orc is $P(O) = 0.4$. Find the probability that you encounter at least one monster.

**Solution 1 |**

Let $G$ be the event of encountering a goblin, and $O$ be the event of encountering an orc. We are looking for $P(G \cup O)$. Since the events are independent, $P(G \cap O) = P(G)P(O)$.

$$\begin{aligned} P(G \cup O) &= P(G) + P(O) - P(G \cap O) \\ &= 0.3 + 0.4 - (0.3 \times 0.4) \\ &= 0.7 - 0.12 = 0.58. \end{aligned}$$

Alternatively, using the complement (encountering no monsters):

$$P(G^c) = 1 - 0.3 = 0.7$$

$$P(O^c) = 1 - 0.4 = 0.6$$

$$P((G \cup O)^c) = P(G^c \cap O^c) = P(G^c)P(O^c) = 0.7 \times 0.6 = 0.42$$

$$P(G \cup O) = 1 - 0.42 = 0.58.$$

Article 14.03

# *Binomial Probability*

## Random Variable

**DEFINITION | Random Variable(tr. 확률 변수)**

A **Random Variable** is the variable that the value changes.
- Due to probabilistic uncertainty, we cannot estimate the variable's number
- After the uncertainty has disappeared (after element of $S$ has been selected), we can know its value.
- In the end, the random variable is a **function** that takes a sample space as the domain.

**NOTATION | Variables**

(Variables) : $x, y, z$ (Random Variables) : $X(s), Y(s), Z(s)$ where $s \in S.S$
- Capitalize to empathize that it is random(...)

- We can use random variable to express events:
  - $\{s \in S \mid X(s) = 8\}$ : Event (subset of $S$) where the element of $S$ is 8

**NOTATION | Event**

$$\{s \in S \mid X(s) = 8\} = \{X = 8\}$$

$$\{s \in S \mid X(s) > 0\} = \{X > 0\}$$

$$\{s \in S \mid 1 \leq X(s) \leq 3\} = \{1 \leq X \leq 3\}$$

**EXAMPLE | Two dice toss**

Find the probability of sum of two dice being 8

**Solution 1 |**

$S = [6] \times [6]$
$X(i, j) = i + j$
$\{X = 8\} = \{(2.6), (3, 5), (4, 4), (5, 3), (6, 2)\}$
$\Rightarrow P\{X = 8\} = \frac{5}{36}$

### Note | A closed thermodynamic system

In a thermodynamic system, where the variables $T(s), P(s), V(s)$ are defined by physics, These variables are random variables. Here, the sample space is the set of all thermodynamical states.
In this case, we can't simply find $\{T(s) = 293K\}$ because sample set have continuous, infinite elements; making it 0.

### Note | Discrete & Continuous Random Variables

A **discrete random variable** is a variable that only takes discrete values as input. For instance, a random variable with the sample space $\{8, 9, 10\}$ is a discrete random variable.

However, in a continuous region, we cannot easily say that P(some value) will be some number. It will be close to 0. Instead, for **continuous random variables**, we define a region to view its probability. For instance, a random variable with the sample space $\{x \mid x \in [8, 10]\}$ where $x \in \mathbb{R}$. We use a **probability mass function** to find its probability in a region.

### Definition | Probability mass function

Let $X$ be a discrete random variable.
Then, the **probability mass function (p.m.f)** is defined as

$$p(a) = P[X = a]$$

### Note | Extra

The deliminator around $X = a$ (e.g. $\{X = a\}$, $(X = a)$, $[X = a]$) doesn't really matter.

**Note | Remark**

$p(a) = 0$ for $a$, except for the values that $X$ can take,
Domain of $p$: $[0, 1]$, $\sum_a = 1$ is the criteria of being valid probability mass function.

## Binomial Random Variables

**Definition | Binomial Random Variables**

A **binomial random variable** is the number of successes among *repeated independent experiments* of $n$ times each experiment has two outcomes “success” and “fail”.

**Notation | Binomial Random Variables**

A binomial random variable has two parameters

- $n$ : number of repeated experiments
- $p$ : (the probability of “success” of each experiment)

We write $X$ : a binomial random variable with parameters $n$, $p$. To simplify, we write

$$X \sim B(n, p)$$

**Note |**

If $X \sim B(n, p)$, $Y \sim B(n, p)$, their p.m.f. are same.
However, $X = Y$ is not necessary.
Beware that even if the probability functions are the same, the actual random variables may not.

## Probability Mass Functions of Binomial Random Variables

Let $X \sim B(n, p)$,

$$\underbrace{P(X = k)}_{\text{k success in n}} = \begin{cases} \binom{n}{k} p^k (1-p)^k \text{ , k= 0,...,n} \\ 0 \text{, otherwise.} \end{cases}$$

**Example | dice dice dice**

$X$: the number of outcomes $\geq 5$, among 7 dice tosses. Find the probability mass function.

**Solution 1 |**

$$X \sim B\left(\underbrace{7}_{\text{trials}}, \underbrace{\frac{1}{3}}_{\substack{\text{probability}\\ \text{of having 5,4,6}}}\right)$$

Count for $X$ to be satisfied 3 times : $P(X=3) = \binom{7}{3} \times \left(\frac{1}{3}\right)^3 \times \left(\frac{2}{3}\right)^4 = \frac{35*2^4}{3^7}$
Count for $X$ to be satisfied $\leq 2$ times :

$$P(X \leq t. = 2) = P(X=0) + P(X=1) = \binom{7}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^7 + \binom{7}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^6$$
$$= \frac{35 \times 2^4}{3^7}$$

## Example | Playing Cards

Choose 5 cards out of a deck without replacement. $X$: the number of Aces. Determine if $X$ is a binomial random variable.

**Solution 1 |**

A binomial random variable requires:
1. A fixed number of trials ($n$).
2. Each trial has only two outcomes (Success/Failure).
3. The probability of success ($p$) is constant for each trial.
4. The trials are independent.

Here, we are choosing cards **without replacement**.
- The probability of the first card being an Ace is $\frac{4}{52}$.
- If the first card was an Ace, the probability of the second card being an Ace is $\frac{3}{51}$.
- If the first card was not an Ace, the probability of the second card being an Ace is $\frac{4}{51}$.

Since the probability changes depending on previous outcomes, the trials are **not independent**. Therefore, $X$ is **not** a binomial random variable. (Note: $X$ follows a **Hypergeometric distribution**).

### Note | Population Research

We consider population research a binomial event. Why is this?

#### Analysis |

Since populations are very large in size, one person being selected does not have a large effect on the next person's probability.

### Example | on-base percentage

Normally, an on-base percentage defines the percentage of a player leaving home base. Since around 2-3 leaves can lead to a score, it is very dangerous to let a player to leave base continuously.

Q. What is the probability of outing three hitters in a row? Calculate for on-base percentage $0.3$ and $0.4$.

#### Solution 1 |

i) $P = 0.3$

$$P(X=3) = \binom{3}{3} \times 0.7^3 \times 0.3^0$$
$$\mathbf{= 0.342}$$

ii) $P = 0.4$

$$P(X=3) = \binom{3}{3} \times 0.6^3 \times 0.4^0$$
$$\mathbf{= 0.216}$$

Q. What is the probability of allowing two people in once (no double play)?

**Solution 2 |**

Y: Number of outed hitters when fighting aginst 4 hitters
$Y \sim B(4, 0.7 \text{ or } 0.6)$
For when 0.7
$P(Y \geq 3) = P(Y = 3) + P(Y = 4) = \binom{4}{3}0.7^3 0.3^1 + \binom{4}{4}0.7^4$
$\approx 0.44$
For when 0.6

**Example | Two-sided 4 dice toss**

$X =$ the absolute difference between two outcomes.
Find the probability mass function of the system.

**Solution 1 |**

Since discrete, we can obtain a probability mass function.
Probability mass function for $X$:
$p(0) = P(X = 0) = P\{(0, 1), (2, 2), (3, 3), (4, 4)\} = \frac{1}{4}$

$p(1) = P(X = 1) = P\{(1, 2), (2, 3), (3, 4), (2, 1), (3, 2), (4, 3)\} = \frac{3}{8}$

$p(2) = P(X = 2) = \frac{1}{4}$

$p(3) = P(X = 3) = \frac{1}{8}$

Article 14.04

# *Expected Value*

## Expectation & Variance

- Expectation $\approx$ Mean

**Notation | Means**

- A sane human would calculate the mean like the following :

$$M = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

- However, some insane mathematicians calculate the mean like this :

$$M + \frac{1 \times c_1 + 2 \times c_2 + \cdots + k \times c_k}{n}$$

- Here, $c_{1 \ldots k}$ stand for the count for each number, with

$$\sum_{k=1}^{k} c_k = n$$

- These means require actual data, so this is a *calculated* mean.

**Note | Another means?**

Let X be the number of head on 100 times of coin throw.

$$X \sim B\left(100, \frac{1}{2}\right)$$

Intuitively, the mean of $X$ is 50.
But is this the *same means* as the one above?
This was not calculated by real value.

**Definition | Expectation**

Let $X$ be a discrete random variable, with probability mass function $f$.

Then, the expectation of $X$ is

$$E(X) = \sum_x x \times f(x)$$

**Notation | Expectation**

$$E(X) = \sum_{x\in\Omega} x \times P(X = x) = \sum_{x\in\Omega} x \times p(x)$$

- Here, $p(x)$ is the probability mass function, and $\Omega$ is the sample space.

**Example | Dice will be back**

$X =$ the outcome of a dice. Calculate $E(X)$.

**Solution 1 |**

Probability mass function of $X$ is:

$$p(x) = P(X = x) = \frac{1}{6},\ \ x = 1, 2, 3, 4, 5, 6$$

$$E(X) = \sum_{x} x \times p(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = \frac{7}{2}$$

**Equation |**

$X \sim B(n, p)$
Guess: $E(X) = np$

$$E(X) = \sum_{k=0}^{n} k(X = k)$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^k$$

$$= \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-k)!} p^k (1-p)^k$$

$$= n \sum_{k=1}^{n} \binom{x-1}{k-1} P^k (1-p)^{x-k}$$

let $j = k - 1$

$$= n \sum_{j=0}^{n-1} \binom{n-1}{j} p^{j+1} (1-p)^{n-1-j}$$

$$= np\sum_{j=0}^{n-1}\binom{n-1}{j}p^{j(1-p)^{n-1-j}}$$

$$= np$$

$$\therefore \text{ If } X \sim B(n,p), \text{ then } E(X) = np$$

**NOTE | The expected value of functions using random variables**

In general, to find the expected value of $f(X)$ where $X$ is the random variable, we can do :

$$E(f(X)) = \sum_x f(x) \times p(x)$$

**NOTE | The linear decomposition of random variables**

$$E(a + bX) = \sum_x (a + bx)p(x)$$

$$= a\sum_x p(x) + b\sum_x xp(x) = a + bE(x)$$

Similarly,

$$E(X + Y) = E(X) + E(Y)$$

However,

$$E(XY) \neq E(X)E(Y)$$

Article 14.05

# *Variance & Deviance*

The variance that we have learned in middle school is defined as:

$$(\text{variance}) = (\text{average of sqared deviation})$$

As we have discussed before, in context of probability, we can applicate this like

$$\text{Var}(X) = E[(X - E(X))^{2\text{`}}]$$

**Equation |**

Simplifying the expression above, If $m = E(X)$,

$$Var(X) = \sum_x (x - m)^2 p(x)$$

$$= \sum_x (x^2 - 2mx + m^2) p(x)$$

$$= \sum_x x^2 p(x) - \sum_x 2mxp(x) + \sum_x m^2 p(x)$$

$$= E(X^2) - 2m^2 + m^2 = E(X^2) - E(X)^2$$

**Notation | Variance & Standard Deviation**

$$\text{Var}(X) = E[(X - E(X))^2] = E(X^2) - E(X)^2$$

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

**Example |**

$X$: the outcome of a dice. $E(X) = \frac{7}{2}$. Find the variance of $X$.

**Solution 1 |**

$E(X^2) = \sum x^2 p(x)$

$= \frac{1}{6}(1^2 + 2^2 \cdots + 6^2)$

$= \frac{91}{6}$

$$\therefore \mathrm{Var}(X) = E(X^2) - E(X)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

**Example | Squared Means in B(n,p)**

From the previous chapter, we know $E(X) = np$ when $X \sim B(n, p)$. Now, find $\mathrm{Var}(X)$

**Solution 1 |**

$$E(X^2) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}$$

Article 14.06

# *Statistics*

## Introduction : What is statistics?

Descriptive statistics shows the correlation of a model with the actual world. In contrast to the probability we learned earlier, which "predicts" the possibility of an event happening, statistics "analyzes" an happened event, organizing and categorizing the data to find meaningful trends.

**Definition | Descriptive Statistics**

**Descriptive statistics** is the field of statistics that analyzes and organizes data.

**Definition | Inferential Statistics**

**Analytical statistics** is the field of statistics that looks at an organized dataset to find meaningful trends.

## Basic Terminology

**Definition | Mean**

An **average/mean** from an sample is the sum of the elements of the sample divided by the number of samples.

**Definition | Variance**

A **variance** is the number representing how "spread" the data is. It is calculated as the average of the square of the difference of each sample to the mean.

* A **standard deviation** is calculated by the root of variance.

**Definition | Median**

A **median** is the value in the middle of the sorted data. It can either be the $\frac{n+1}{2}$th value in $n = 2k$, or the average of the $\frac{n}{2}$th element and the $\frac{n}{2} + 1$th element if n=2k+1.

**DEFINITION | Mode**

A **mode** shows the element with the most shows inside of a dataset.

**DEFINITION | Quartiles**

- **First Quartile(Q1)**
  - the median of the lower 50%.
- **Third Quartile(Q3)**
  - The median of the upper 50%.
- The second quartile is equal to the global median.

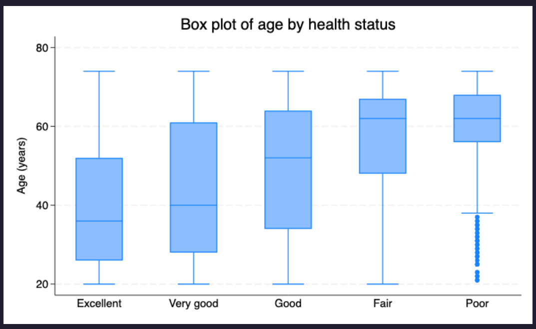
**DEFINITION | Five-number Summary**

$$\text{min} - —— \text{Q1} - —— \text{median(Q2)} - —— \text{Q3} - —— \text{max}$$

Here, the difference between each region(———) is called the **IQR(Interquartile range)**

The larger the IQR, the more widespread the data is.

Using this, we can define a thing called a box plot.




Example of a box plot.

## The Central Limit Theorem

Say we nave $n$ "random generators" which spit out a number in the range $[0, 1]$. Here, if we add multiple vaiables together, the probability of each number to appear will slowly take on a bell-shaped curve. Why is this?

**THEOREM | The Central Limit Theorem**

If $X \sim B(n, p)$, and $n$ is sufficiently large, $P(a \leq X \leq b)$ is equal to the area under the graph :

$$y = f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

The **central limit theorem** states that the more random variables are merged, the closer the probability above approaches the area under the graph.

### Definition | Normal Random Variable

If a random variable $Y$ satisfies the property

$$P(a \leq Y \leq b) = \left(\text{the area under } y = f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right)$$

then $Y$ is called a **normal random variable** or a **Gaussian** variable whith mean $\mu$ and variance $\sigma^2$.

Then, we write $Y$ as

$$Y \sim N(\mu, \sigma^2)$$

### Note | Remark

If $Y \sim N(\mu, \sigma^2)$ then

$$aY + b \sim N(a\mu + b, a^2\sigma^2)$$

### Definition | Normalization

If we subtract from $Y$ $\mu$ and divide by $\sigma$, we get a normal distribution of $N(0, 1)$. We call this normalized distribution a **standard normal distribution**.

### Example | Coin toss

Let $X$ be the number of heads among 80 coin tosses.

Q1. Express X as a normal distribution.

**Solution 1 |**

$$X \sim B\left(80, \frac{1}{2}\right) \approx N\left(80 \times \frac{1}{2}, 80 \times \frac{1}{2} \times \left(1 - \frac{1}{2}\right)\right) = N(40, 20)$$

Q2. Find the probability where $X \leq 35$.

**Solution 2 |**

$$P(X \leq 35) = P\left(\frac{X - 40}{\sqrt{20}}\right) \leq \frac{35 - 40}{\sqrt{20}}$$

Here,we can see that $Z = \frac{X-40}{\sqrt{20}}$ is approximately a standard normal distribution.

$$\approx P(Z \leq -1.118)$$

**Note | cf**

$$P(Z \leq n) = \Phi(n)$$

This function is called the **Gauss culmitative distribution function**, represented by the equation

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Some notable characteristics include:

- $\Phi(0) = \frac{1}{2}$
- $\Phi(x) + \Phi(1 - x) = 1$

We can easily calculate this using the $\text{norm.s.dist}(x)$ function in google spreadsheets.

Continuing from above using this new knowledge,

$$\approx \Phi(-1.118) \approx 13.2\%$$

Q3. Find the probability where $X < 36$.

**Solution 3 |**

$$P(X < 36) = P\left(\frac{X-40}{\sqrt{20}} < \frac{36-40}{\sqrt{20}}\right) \approx P(Z <\sim\sim) \approx 18.6\%$$

### Note | limitations of the normal distribution

Binomial distributions are fundamentally discrete random variables. However, since normal distributions are continuous, they will have a few disrepancies similar to the one shown above. Thus, mathematicians use the middle value(35.5 in this case).

### Example | Dice x300

We are going to throw a dice 300 times. Find the probability that we get 61 64 threes.

#### Solution 1 |

$$X \sim B\left(300, \frac{1}{6}\right) \approx N\left(300 \times \frac{1}{6}, 300 \times \frac{1}{6} \times \frac{5}{6}\right) = N(50, 6.455^2)$$

Therefore,

$$P(61 \leq X \leq 64)$$

$$= P\left(\frac{X-50}{6.455} \leq \frac{64-50}{6.455}\right) - P\left(\frac{X-50}{6.455} \leq \frac{60-50}{6.455}\right)$$

Now that we have reached normality, we can calculate using $\Phi$.

$$\approx \Phi\left(\frac{14}{6.455}\right) - \Phi\left(\frac{10}{6.455}\right)$$

### Theorem | Empirical Rule

$$\Phi(1.282) = 0.9$$

$$\Phi(1.645) = 0.95$$

$$\Phi(1.960) = 0.975$$

From this, we can know the following :
- Within $\pm 1.282\sigma$, 80% of the data exists.
- Within$\pm 1.645\sigma$, 90% of the data exists.
- Within $\pm 1.960\sigma$, 95% of the data exists.

**Notation | $Z$**

We write for a $Z_\alpha \in \mathbb{R}$ such that $\Phi(-Z_\alpha) = 1 - \alpha$. We can note that $-Z_\alpha = Z_{1-\alpha}$

**Example | The defection rate of a 4nm semiconductor**

The defection rate of a 4nm semiconductor at a s***ung factory is 40%.

Q1. If we choose a 100 random chips from a pile, describe the distribution of defective chips.

**Solution 1 |**

Let $X$ be the number of defective chips.

$$X \sim B(100, 0.4) \overset{\text{CLT}}{\approx} \mathcal{N}(40, 100 \cdot 0.4 \cdot 0.6)$$
$$= N(40, 24)$$

Q2. Find $n$ such that $X$ exceeds $n$ with probability 5%.

**Solution 2 |**

$$P(X > n) = 0.05$$
$$P\left(\frac{X-40}{\sqrt{24}} > \frac{n-40}{\sqrt{24}}\right)$$
$$= 1 - \Phi\left(\frac{n-40}{\sqrt{24}}\right) = Z_{0.05} \approx 1.645$$
$$\therefore n \approx 48$$

Q3. Find integers $a$ and $b$ sich that $P(a. \leq X \leq b) = 99\%$

### Solution 3-1 | Center is the mean

The region equals $\mu \pm \sigma \times Z_{0.005}$. Thus, the region is 27 ~ 53.

### Solution 3-2 | A is minimum

The region equals 0 to $\mu + \sigma \times Z_{0.01}$. Thus, the region is 0 ~ 51.

### Solution 3-3 | A is minimum

The region equals $\mu - \sigma \times z_{0.001}$ to 100. Thus, the region is 29 ~ 100.