Mathematica Compendium

Miguel Antonio Méndez Hernández

April 17, 2025

Contents

1 Introduction **8**

2 Propositional Logic **9**
2.1 Logic Operators and Truth Tables . . . 9
2.2 Tautology and Contradiction . . . 9
2.3 Logical Equivalences . . . 10
2.4 Truth Tables . . . 10
2.4.1 Filling a truth table . . . 11
2.4.2 General Procedure: . . . 11
2.4.3 Example with 3 Variables (A, B, and C) . . . 11
2.4.4 The pattern for filling the input truth values . . . 11
2.4.5 Truth Table for the Expression $(A \land B) \lor C$ . . . 11
2.5 Disjunctive Normal Form (DNF) . . . 12
2.6 Conjunctive Normal Form (CNF) . . . 12
2.7 Karnaugh Maps . . . 12
2.7.1 Karnaugh Map for Two Variables . . . 12
2.7.2 Karnaugh Map for Three Variables . . . 13
2.7.3 Solving a Karnaugh Map (K-Map) . . . 13
2.8 Mathematical Quantifiers with Negations and Examples . . . 14
2.9 Common Symbols Used in Mathematical Expressions . . . 14

3 Set Theory **16**
3.1 Basics . . . 16
3.1.1 Visuals . . . 16
3.2 Axioms of Set Theory (Zermelo Fraenkel) . . . 17
3.3 The Cartesian Product . . . 17
3.4 Laws of Set Algebra . . . 18
3.4.1 Proof of De Morgans's Law for sets and logic . . . 18
3.5 Indexed Sets . . . 19
3.5.1 More Partitions Laws . . . 19
3.6 Cardinality . . . 19
3.6.1 General Formula for the cardinality of the union of sets . . . 20
3.7 The Power Set . . . 20
3.8 Family of Subsets . . . 20
3.9 Partition . . . 20
3.10 Family of Subsets Operations . . . 20

4 Relations, Maps and Functions **21**
4.1 Types of relations . . . 21
4.2 Equivalence relation . . . 21
4.3 The Graph . . . 21
4.4 The identity . . . 21
4.5 Image and Domain . . . 21
4.5.1 Image . . . 21
4.5.2 Domain . . . 22
4.6 Equivalence Class . . . 22
4.7 Quotient Space . . . 22
4.8 Definition of a Map . . . 22
4.9 Composition of Maps . . . 22
4.10 Types of Functions . . . 23
4.10.1 Injective Functions . . . 23
4.10.2 Surjective Functions . . . 23
4.10.3 Bijective Functions . . . 23
4.11 Propositions on Images and Preimages under Set Operations . . . 23
4.11.1 Union and Cut Sets . . . 23
4.11.2 Union and Cut of the whole Domain and Range . . . 24
4.12 Inverse of a Function . . . 24

4.12.1 Steps to Find the Inverse of a Function . . . 24
4.12.2 Example: Finding the Inverse of $f(x) = 2x + 3$ . . . 25
4.12.3 Properties of the Inverse Function . . . 25
4.13 Transformations of a Function . . . 26

5 Mathematical Proofs **27**
5.1 Proof by Direct Argument . . . 27
5.2 Proof by Contradiction . . . 27
5.3 Proof by Induction . . . 27
5.4 Proof by Exhaustion . . . 28
5.5 Proof by Cases . . . 28
5.6 Proof by Construction . . . 28
5.7 Proof by Counterexample . . . 28
5.8 Proof by Contrapositive . . . 29
5.9 Proof by Reduction to Absurdity . . . 29
5.10 Proof by Analogy . . . 29

6 The Natural Numbers **30**
6.1 Propositions and Proofs . . . 30
6.1.1 Proposition 1: $n \neq m \implies S(n) \neq S(m)$ . . . 30
6.1.2 Proposition 2: For any $n \in \mathbb{N},\ n \neq S(n)$ . . . 30
6.1.3 Proposition 3: $n \neq 1\ \exists m \in \mathbb{N} \mid n = S(m)$ . . . 30
6.2 Definition of Addition in $\mathbb{N}$ . . . 31

7 The Archimedean Principle **32**
7.1 Statement . . . 32
7.2 Proof . . . 32
7.3 Equivalent Formulations . . . 32
7.4 Applications . . . 32

8 Fundamental Theorem of Arithmetic **33**
8.1 Lemmas . . . 33
8.2 Proof of the Fundamental Theorem . . . 33
8.3 Least Common Multiple . . . 34

9 Real Numbers **35**
9.1 Definition . . . 35

10 Complex Numbers **36**
10.1 What are Complex Numbers? . . . 36
10.2 The Complex Plane . . . 36
10.3 Conjugate of a Complex Number . . . 36
10.4 Operations in Cartesian Coordinates . . . 36
10.5 Polar Coordinates . . . 37
10.6 Multiplication and Division in Polar Coordinates . . . 37
10.7 Exponentiation and Roots (De Moivre's Theorem) . . . 37
10.7.1 Exponentiation . . . 37
10.7.2 Roots of Complex Numbers . . . 37
10.8 Example: Solve $z^4 = 1 + \sqrt{3}i$ . . . 38
10.9 Solving Equations with Complex Numbers . . . 38
10.9.1 Example 1: . . . 39
10.9.2 Example 2: . . . 39
10.10 The Complex Logarithm . . . 39
10.11 Complex Exponents . . . 40
10.12 Euler's Formula Proof . . . 40

11 Topology **41**
11.1 Introduction to topological nomenclature . . . 41

12 Fractions, Roots, and Exponents **42**
12.1 Fractions . . . 42
12.2 Roots and Exponents . . . 42

13 Logarithms **43**
13.1 Properties of Logarithms . . . 43
13.2 Fundamental Identity of Logarithms . . . 43
13.3 Change of Base Formula . . . 43
13.4 The Chain Rule . . . 43
13.5 The derivative of the Natural Logarithm . . . 43

14 Sum and Product Notation **45**
14.1 Sum Notation $\sum$ . . . 45
14.2 Product Notation $\prod$ . . . 46

15 Means and Proofs **48**
15.1 Proof of the Arithmetic Mean-Geometric Mean Inequality . . . 48
15.2 Proof of the Harmonic Mean Geometric Mean Inequality . . . 50

16 Solving Polynomial Equations **51**
16.1 The PQ Formula . . . 51
16.2 The ABC Formula (Quadratic Formula) . . . 51
16.3 The Cubic Formula . . . 52

17 The Fundamental Theorem of Algebra **53**
17.1 Theorem 1 . . . 53
17.2 Theorem 2 . . . 54

18 The Binomial Coefficient **55**
18.1 Properties . . . 55
18.2 The binomial Theoreme . . . 55

19 Proportionality and the Rule of Three **56**
19.1 Proportionality . . . 56
19.2 The Rule of Three (Simple) . . . 56
19.3 The Rule of Three (Compound) . . . 56

20 Factorization Techniques **58**
20.1 Common Factor . . . 58
20.2 Difference of Squares . . . 58
20.3 Sum/Difference of Cubes . . . 58
20.4 Trinomial: General Form $ax^2 + bx + c$ . . . 58
20.5 Trinomial: Special Case $x^2 + bx + c$ . . . 58
20.6 Perfect Square Trinomials . . . 58
20.7 Substitution . . . 58
20.8 Rationalization of Radicals . . . 59
20.9 Ruffini's Rule (Horner's Method) . . . 59
20.10Long Division of Polynomials . . . 59

21 Partial Fractions **60**
21.1 The Simplest Case . . . 60
21.1.1 Lemma 1 . . . 60
21.1.2 Recursive Decomposition . . . 60
21.2 Lemma 2 . . . 61

22 Solving Inequalities **62**
22.1 Basic Principles . . . 62
22.2 Inequalities Involving Absolute Value . . . 62
22.2.1 Example: $|x + 1| \leq 10 + |2x - 6|$ . . . 62

23 Geometry **64**
23.1 Formulas for Area and Perimeter of Geometric Figures . . . 64
23.2 Formulas for Area and Volume of 3D Geometric Figures . . . 64
23.3 Thales' Theorem . . . 65
23.4 Conversion Between Radians and Degrees . . . 65
23.5 3D Geometric Figures with Formulas . . . 66
23.5.1 Cube and Cylinder . . . 66
23.5.2 Cone and Sphere . . . 67
23.5.3 Pyramid and Prism . . . 67
23.6 Intercept Theorems . . . 67
23.6.1 First Intercept Theorem . . . 67
23.6.2 Second Intercept Theorem (General Form) . . . 68

24 Trigonometry **69**
24.1 Trigonometric Functions . . . 69
24.2 SOH-CAH-TOA . . . 69
24.3 Pythagorean Identities . . . 69
24.4 Sum and Difference Formulas . . . 69
24.5 Double Angle Formulas . . . 69
24.6 Half Angle Formulas . . . 70
24.7 Product to Sum Formulas . . . 70
24.8 Power Reducing Formulas . . . 70
24.9 Even and Odd Functions . . . 70
24.10 Graph Identity . . . 70
24.11 Trigonometric Values . . . 71

25 Linear Systems of Equations **73**
25.1 Matrix Representation . . . 73
25.2 Gaussian Elimination . . . 73
25.3 Gauss-Jordan Elimination . . . 74
25.4 Homogeneous Linear Equations . . . 74
25.5 Particular Solution . . . 75
25.6 General = Particular + Homogeneous . . . 75
25.7 Linear Combination Lemma . . . 75
25.8 Example: Gaussian Elimination with 3 Equations and 4 Unknowns . . . 75
25.9 The Determinant, the cross product and the solutions of linear Systems of Equations . . . 76

26 Analytical Geometry **78**
26.1 Vectors and Points . . . 78
26.2 Vector Addition and Scalar Multiplication . . . 78
26.3 Equation of a Line . . . 78
26.4 Equation of a Plane . . . 78
26.5 Scalar (Dot) Product . . . 78
26.6 Length (Norm) of a Vector . . . 79
26.7 Angle Relations . . . 79
26.8 Line Relations . . . 80
26.9 Normalization of a vector . . . 80
26.10 Orthogonal Vectors and the Orthogonal Projection . . . 80
26.11 The Cross Product . . . 81
26.12 Orthogonal vectors in $\mathbb{R}^2$ $\mathbb{R}^3$ . . . 82
26.12.1 Orthogonal vectors in $\mathbb{R}^2$ . . . 82
26.12.2 Orthogonal vectors in $\mathbb{R}^3$ . . . 82
26.13 Hessian Normal Form . . . 82
26.14 Converting from the parametric form to the Hessian normal form . . . 83
26.15 Converting from the Hessian normal form to the parametric form . . . 83
26.16 Properties of lines and planes . . . 83
26.17 Convert Normal Vector in Two Direction Vectors . . . 83
26.18 Intersection between Line and Plane . . . 84
26.19 Distances between points, lines and planes . . . 85

26.19.1 Example: Distance Between Two Skew Lines . . . 86
26.20 Foot of the Perpendicular and Mirror Point . . . 87
26.20.1 Foot of the Perpendicular . . . 87
26.20.2 Mirror Point . . . 87

27 Algebraic Structures **88**
27.1 Introduction . . . 88
27.2 Operations: Internal and External . . . 88
27.3 Properties of Operations . . . 88
27.4 Homomorphisms and Isomorphisms . . . 88
27.5 Common Algebraic Structures . . . 88

28 Vector Spaces **90**
28.1 Definition . . . 90
28.2 Examples . . . 90
28.3 Subspaces . . . 90
28.4 Linear Combinations . . . 90
28.5 Properties of the subspaces . . . 91
28.6 Linear Independence . . . 91
28.6.1 Properties of the linear independence . . . 91
28.7 Base . . . 91
28.8 Dimension . . . 92
28.8.1 How to find the base of a set vector . . . 92
28.9 Basis Extension Theorem . . . 92
28.10 Exchange Lemma . . . 93
28.11 Dimension of a sum of subspaces . . . 93
28.12 Linear Independence of polynomials . . . 93
28.13 Interpolation Polynomial . . . 94
28.13.1 Example of the interpolation polynomial . . . 94

29 Dot Product, Euclidean and Unitary Space **95**
29.1 Scalar Product . . . 95
29.2 Standard Scalar Product for Complex Number . . . 95
29.3 Scalar Product on $C[a, b]$ . . . 95
29.4 Euclidean and Unitary Vector Spaces . . . 95
29.5 Norms in Vector Spaces . . . 95
29.5.1 Induced Norm by a Scalar Product . . . 96
29.6 Cauchy-Schwarz Inequality in Unitary Vector Spaces . . . 96

30 Orthogonality **97**
30.1 Orthogonality and Projection . . . 97
30.2 Orthogonal Projection and Orthogonal Complement . . . 97
30.3 Orthogonality of the basis to the Complement . . . 97
30.4 How to find the orthogonal complement . . . 97
30.5 Orthogonal and Orthonormal Systems . . . 98
30.6 Writing a vector in terms of the Orthogonal Basis . . . 98
30.7 Linear Independence of Orthogonal Systems . . . 98
30.8 Representation with Respect to an Orthogonal Basis . . . 98
30.9 Orthogonal Projection and Direct Sum Decomposition . . . 99
30.10 The Gram-Schmidt Process . . . 99
30.11 Existence of Orthonormal Bases and Orthogonal Projections . . . 99
30.12 Best Approximation . . . 100

31 Linear Maps **101**
31.1 Types of Linear Maps . . . 101
31.2 Properties of Linear Maps . . . 101
31.3 The Kernel of a Linear Map . . . 101
31.4 The Image of a Linear Map . . . 101
31.5 The Rank of a Linear Map . . . 102

31.6 The Nullity of a Linear Map . . . 102
31.7 The Rank-Nullity Theorem . . . 102
31.8 Proof of the injectivity of the Kernel . . . 102
31.9 Dimension Formula for Linear Mappings . . . 102
31.10 Identifying the type of linear map . . . 103
31.11 Image of the basis . . . 103
31.12 Linear Maps and Matrices . . . 103

32 Matrices **104**
32.1 Definition of a Matrix . . . 104
32.2 Matrix Addition and Subtraction . . . 104
32.2.1 Addition . . . 104
32.2.2 Subtraction . . . 104
32.3 Matrix Multiplication . . . 104
32.4 The Transpose of a Matrix . . . 105
32.5 The Equivalence of Matrices . . . 105
32.5.1 Row Echelon Form (REF) . . . 105
32.5.2 Reduced Row Echelon Form (RREF) . . . 105
32.6 The Inverse of a Matrix and Its Properties . . . 106
32.6.1 Properties of the Inverse . . . 106
32.6.2 Finding the Inverse . . . 106
32.7 The Rank of a Matrix and How to Find It . . . 106
32.7.1 Finding the Rank . . . 107
32.8 The Definitions of Column Space, Row Space, and Null Space . . . 107
32.8.1 Column Space . . . 107
32.8.2 Row Space . . . 107
32.8.3 Null Space . . . 107
32.8.4 Left Null Space . . . 108
32.9 Examples of Matrix Operations . . . 108
32.9.1 Example of Matrix Multiplication . . . 108
32.9.2 Example of Gauß-Jordan Elimination . . . 109
32.9.3 Example of Finding the Inverse of a Matrix using Gauß-Jordan Elimination . . . 110
32.10 The Determinant of a Matrix . . . 111
32.10.1 Example of Determinant Calculation . . . 113
32.11 Linear Maps as matrices and Their Properties . . . 114
32.12 Example of an exercise . . . 114
32.13 Eigenvectors and Eigenvalues . . . 115
32.13.1 How to find the Eigenvectors and Eigenvalues . . . 115
32.13.2 How to diagonalize a matrix . . . 115

1 Introduction

Hello my name is Miguel, this compendium is not mean to be a complete guide/book about all of mathematics, but a collection of theorems, definitions, and proofs that I have found useful in my studies. This compendium is not a replacement for a the books or the lectures at your university, but a complement to it. On the one side I will try to keep it as simple as possible, but sometimes there are going to topics that will be quite complex and on the other side some proofs will be skiped or not included at all, because I think that they are not necessary for the understanding of the topic and may lead to confusion.

My main idea while wirting this was to take as musch as I could from the books I have read, the lectures I have attended, the videos I have watched and the notes I have taken. This is also the reason why the Order may seem a bit strange, but I think that it is the best way to give an overiew about a lot of topics and also via the table of contents you can easily find the topic you are looking for.

The whole compendium is written in LaTeX, so if you find any mistake or you want to add something, please feel free to do so in your own computer. If you see mistakes or you want to add something to the online version, please let me know and I will try to fix it as soon as possible. I will try to keep the compendium updated as much as I can, but I am not a professional writer nor a LaTeX veteran, so please be patient with me.

If you find this useful please consider maybe donating to this project, but dont worry I this document to be free for ever.

2 Propositional Logic

Propositional logic is also called Boolean logic as it works on 0 and 1. In propositional logic, we use symbolic variables to represent the logic, and we can use any symbol for a representing a proposition, such *A, B, C, P, Q, R*, etc. Propositions can be either *true* or *false*, but it cannot be both.

2.1 Logic Operators and Truth Tables

NOT ($\neg$, $\sim$)

A	$\neg A$
0	1
1	0

AND ($\wedge$)

A	B	$A \wedge B$
0	0	0
0	1	0
1	0	0
1	1	1

OR ($\vee$)

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

IMPLIES ($\implies$)

A	B	$A => B$
0	0	1
0	1	1
1	0	0
1	1	1

IFF ($\iff$)

A	B	$A <=> B$
0	0	1
0	1	0
1	0	0
1	1	1

XOR ($\oplus$)

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

NOR ($\downarrow$)

A	B	$A \downarrow B$
0	0	1
0	1	0
1	0	0
1	1	0

NAND ($\uparrow$)

A	B	$A \uparrow B$
0	0	1
0	1	1
1	0	1
1	1	0

2.2 Tautology and Contradiction

- *Tautology*: A logical formula that is always true.
    - *Contradiction*: A formula that is always false.

2.3 Logical Equivalences

Commutative Laws

$$p \wedge q \Leftrightarrow q \wedge p \qquad p \vee q \Leftrightarrow q \vee p$$

Associative Laws

$$(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r) \qquad (p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$$

Distributive Laws

$$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r) \qquad p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$$

Identity Laws

$$p \wedge \mathrm{T} \Leftrightarrow p \qquad p \vee \mathrm{F} \Leftrightarrow p$$

Negation Laws

$$p \vee \sim p \Leftrightarrow \mathrm{T} \qquad p \wedge \sim p \Leftrightarrow \mathrm{F}$$

Double Negation Law

$$\sim (\sim p) \Leftrightarrow p$$

Idempotent Laws

$$p \wedge p \Leftrightarrow p \qquad p \vee p \Leftrightarrow p$$

Universal Bound Laws

$$p \vee \mathrm{T} \Leftrightarrow \mathrm{T} \qquad p \wedge \mathrm{F} \Leftrightarrow \mathrm{F}$$

De Morgan's Laws

$$\sim (p \wedge q) \Leftrightarrow (\sim p) \vee (\sim q) \qquad \sim (p \vee q) \Leftrightarrow (\sim p) \wedge (\sim q)$$

Absorption Laws

$$p \vee (p \wedge q) \Leftrightarrow p \qquad p \wedge (p \vee q) \Leftrightarrow p$$

Conditional Laws

$$(p \Rightarrow q) \Leftrightarrow (\sim p \vee q) \qquad \sim (p \Rightarrow q) \Leftrightarrow (p \wedge \sim q)$$

Complement Law

$$p \vee \neg p \Leftrightarrow \mathrm{T} \qquad p \wedge \neg p \Leftrightarrow \mathrm{F}$$

Biconditional

$$p \Leftrightarrow q \Leftrightarrow (p \Rightarrow q) \wedge (q \Rightarrow p)$$

Transitivity

$$(p \Rightarrow q) \wedge (q \Rightarrow r) \Rightarrow (p \Rightarrow r)$$

Indirect Proof (Contrapositive)

$$(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$$

Disjunctive Syllogism (Disjunctive Exclusion)

$$p \nabla q \equiv (p \vee q) \wedge \neg p \Rightarrow q$$

$$p \nabla q \equiv (p \wedge q) \vee \neg (p \wedge q)$$

2.4 Truth Tables

Truth tables are a fundamental tool in logic that systematically show the truth value (true or false) of a compound statement for every possible combination of the truth values of its individual component statements.

Essentially, they lay out all the scenarios and the resulting truth of the overall logical expression. This helps determine if an argument is valid, if statements are logically equivalent, or the circumstances under which a complex statement is true or false.

Example:

p	q	$\neg p$	$\neg q$	$\neg p \Rightarrow q$	$(\neg p \Rightarrow q) \wedge \neg p$	$[(\neg p \Rightarrow q) \wedge \neg p] \Rightarrow q$
T	T	F	F	T	F	T
T	F	F	T	T	F	T
F	T	T	F	T	T	T
F	F	T	T	F	F	T

2.4.1 Filling a truth table

To fill a truth table for a logical expression with truth values (True or False), you follow a specific order for the input variables. This order ensures that all possible combinations of truth values for the variables are covered.

2.4.2 General Procedure:

1. **List all possible combinations of truth values for the input variables**: If you have n variables, the number of rows in the truth table will be 2^n. Each variable can be either True (T) or False (F).

2. **Order of the input variables**:

   - Start by filling in the truth values for the first variable. It alternates between True and False every 2^{n-1} rows.
   - Then for the second variable, it alternates every 2^{n-2} rows, and so on.
   - In short: the first variable alternates every other row, the second variable every two rows, the third every four rows, and so on.

2.4.3 Example with 3 Variables (A, B, and C)

For 3 variables, there are $2^3 = 8$ possible combinations of truth values. The truth values are filled in the following order:

A	B	C	Expression Result
T	T	T	
T	T	F	
T	F	T	
T	F	F	
F	T	T	
F	T	F	
F	F	T	
F	F	F	

2.4.4 The pattern for filling the input truth values

- The first column (A) alternates every 4 rows: 'T, T, F, F, T, T, F, F'.
- The second column (B) alternates every 2 rows: 'T, T, F, F, T, T, F, F'.
- The third column (C) alternates every row: 'T, F, T, F, T, F, T, F'.

This ensures that all combinations of A, B, and C are covered, and you can then evaluate the logical expression for each combination.

2.4.5 Truth Table for the Expression $(A \wedge B) \vee C$

A	B	C	$(A \wedge B) \vee C$
T	T	T	T
T	T	F	T
T	F	T	T
T	F	F	F
F	T	T	T
F	T	F	F
F	F	T	T
F	F	F	F

2.5 Disjunctive Normal Form (DNF)

Disjunctive Normal Form (DNF) is a standard way of writing a logical expression as a disjunction (OR) of conjunctions (ANDs). A DNF expression consists of a series of conjunctions of literals, where each conjunction is connected by disjunctions.

To find the DNF in a truth table take the rows of the final result where there are true statements and bind the propositions that generated it with an AND inside parenthesis. Repeat it with each of ¡¡true¿¿ rows and connect all parenthesis with OR's

Example of DNF:

Consider the logical expression:

$$(A \wedge B) \vee (\neg A \wedge C) \vee (B \wedge \neg C)$$

This is in DNF because it is a disjunction (OR) of conjunctions (ANDs) of literals.

2.6 Conjunctive Normal Form (CNF)

Conjunctive Normal Form (CNF) is a standard way of writing a logical expression as a conjunction (AND) of disjunctions (ORs). A CNF expression consists of a series of disjunctions of literals, where each disjunction is connected by conjunctions.

To find the CNF proceed just as the DNF but with the "false rows and instead of ANDs inside the parenthesis use OR and connect the terms with OR. Also add a negation before each parenthesis.

Example of CNF:

Consider the logical expression:

$$\neg(A \vee B) \wedge \neg(\neg A \vee C) \wedge \neg(B \vee \neg C)$$

This is in CNF because it is a conjunction (AND) of disjunctions (ORs) of literals.

2.7 Karnaugh Maps

Karnaugh Maps (K-Maps) are a graphical method used to simplify Boolean expressions. The main goal of a K-map is to group adjacent cells that contain 1's in order to simplify the expression. A K-map helps identify common terms, allowing the Boolean expression to be reduced to its simplest form.

2.7.1 Karnaugh Map for Two Variables

Consider the Boolean expression $(A \vee (B \wedge \neg A \wedge \neg B))$.

We first construct a K-map for two variables, A and B. The truth table for this expression gives the following values:

A	B	$(A \vee (B \wedge \neg A \wedge \neg B))$
0	0	0
0	1	1
1	0	1
1	1	1

The corresponding K-map is:

AB	00	01	11	10
Value	0	1	1	1

Here, we group the ones together to simplify the Boolean expression. The simplified expression is:

$$A \vee B$$

2.7.2 Karnaugh Map for Three Variables

Now, let's consider the expression $\neg C$. This expression only depends on one variable, but for illustration, we will use a 3-variable K-map with variables A, B, and C.

The truth table for $\neg C$ is as follows:

A	B	C	$\neg C$
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

The corresponding K-map for three variables A, B, and C is:

$AB \backslash C$	0	1
00	1	0
01	1	0
11	1	0
10	1	0

We see that the ones are grouped in a column, leading to the simplified Boolean expression:

$$\neg C$$

2.7.3 Solving a Karnaugh Map (K-Map)

To solve a Karnaugh Map (K-Map) and simplify a Boolean expression, follow these steps:

1. **Determine the Number of Variables:**
   Decide how many variables are in the Boolean function. This determines the size of the K-Map:
   - 2 variables: 2×2
   - 3 variables: 2×4
   - 4 variables: 4×4
   - etc.
2. **Fill in the K-Map:**
   Place 1's in the cells that correspond to the minterms (where the function outputs 1). You may also include don't-care conditions (usually denoted as X).
3. **Group the 1's:**
   Form groups (called *implicants*) of 1's. The groups must follow these rules:
   - Each group must contain $1, 2, 4, 8, \ldots$ (powers of 2) 1's.
   - Groups must be rectangular (e.g., 1×2, 2×2).
   - Groups can wrap around the edges of the K-Map.
   - Try to form the largest groups possible to simplify the expression.
   - Each 1 should be included in at least one group.
4. **Write the Simplified Expression:**
   For each group:
   - Identify the variables that are constant (either always 0 or always 1) across the group.
   - Write a product term (AND) using only the constant variables.
   - Combine all product terms with OR operations to get the final simplified SOP (Sum of Products) expression.

Example:
Given minterms $F(A, B, C) = \sum m(1, 3, 5, 7)$:

K-Map:

AB \C	0	1
00	0	1
01	0	1
11	0	1
10	0	1

Simplified expression:

$$F = B \lor C$$

2.8 Mathematical Quantifiers with Negations and Examples

Universal Quantifier: $\forall$ Means **"for all"** or **"for every"**.

- **Example:** $\forall x \in \mathbb{R},\ x^2 \geq 0$
  (For all real numbers, the square is greater than or equal to zero.)
- **Negation:** $\neg(\forall x)\, P(x) \equiv (\exists x)\, \neg P(x)$
  ("Not all" is the same as "There exists one that does not".)
- **Negated Example:** $\exists x \in \mathbb{R},\ x^2 < 0$
  (There exists a real number whose square is less than zero — this is false.)

Existential Quantifier: $\exists$ Means **"there exists at least one"**.

- **Example:** $\exists x \in \mathbb{N},\ x > 10$
  (There exists a natural number greater than 10.)
- **Negation:** $\neg(\exists x)\, P(x) \equiv (\forall x)\, \neg P(x)$
  ("There does not exist" is the same as "For all, not".)
- **Negated Example:** $\forall x \in \mathbb{N},\ x \leq 10$
  (All natural numbers are less than or equal to 10 — this is false.)

Unique Existential Quantifier: $\exists!$ Means **"there exists exactly one"**.

- **Example:** $\exists! x \in \mathbb{R},\ x + 5 = 0$
  (There exists exactly one real number such that $x + 5 = 0$.)
- **Negation:** "Not exactly one" means:

$$\neg(\exists! x)\, P(x) \equiv (\forall x)\, \neg P(x)\ \lor\ (\exists x_1 \neq x_2)\, P(x_1) \land P(x_2)$$

  (Either no such x exists, or more than one does.)
- **Negated Example:** $\exists x_1 \neq x_2 \in \mathbb{R},\ x_1^2 = 4 \land x_2^2 = 4$
  (There are multiple solutions to $x^2 = 4$.)

2.9 Common Symbols Used in Mathematical Expressions

- $>$ (greater than)
- $<$ (less than)
- $\geq$ (greater than or equal to)
- $\leq$ (less than or equal to)
- $=$ (equals)
- $\neq$ (not equal)
- $\in$ (element of a set)
- $\notin$ (not an element of)
- $\subset$ (proper subset)
- $\subseteq$ (subset)

- $\supset$ (proper superset)
- $\supseteq$ (superset)
- $\wedge$ (logical AND)
- $\vee$ (logical OR)
- $\Rightarrow$ (implies)
- $\Leftrightarrow$ (if and only if)

3 Set Theory

Set theory is a foundational branch of mathematics that studies sets, which are simply collections of objects. At its core, it deals with the fundamental concepts of membership (whether an object belongs to a set), equality (when two sets are the same), and relationships between sets (like subsets, intersections, and unions).
It might seem simple, but set theory provides the basic language and tools to define and reason about almost all mathematical objects, from numbers and functions to more complex structures. It helps us understand the concept of infinity, organize mathematical ideas logically, and resolve paradoxes that arise from dealing with collections.
In essence, set theory provides the building blocks upon which much of modern mathematics is constructed.

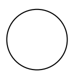
3.1 Basics

To be concise, a set is a collection of mathematical objects that can also be other sets.
Here is a list of common symbols for set theory.

- **Empty set ($\emptyset$)**: The set that contains no elements. It is the unique set with zero elements.
- **Example set with two elements**: A set that contains exactly two elements, such as $\{1, 2\}$.
- $A \subseteq B$: Set A is a subset of B. This means every element of A is also in B.
- $A \subseteq B$ **or** $A = B$: This notation already includes the possibility that A equals B since a set is always a subset of itself.
- $A \cup B$: The union of sets A and B. It includes all elements that are in A, in B, or in both.
- $A \cap B$: The intersection of sets A and B. It includes only the elements that are in both sets.
- $A \setminus B$: The difference of sets. Elements in A that are not in B.
- A^c **or** $\overline{A}$: The complement of set A. All elements not in A, relative to a universal set.
- $|A|$: The cardinality of set A, which is the number of elements in the set.

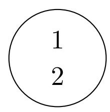
3.1.1 Visuals

1. $\emptyset$



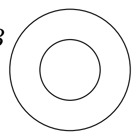
1. Empty Set

2. $\{1, 2\}$



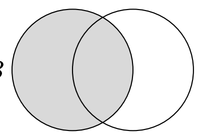
2. Set with Two Elements

3. $A \subseteq B$

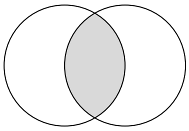


3. A is a subset of B

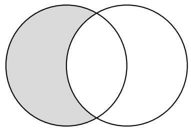
5. $A \cup B$



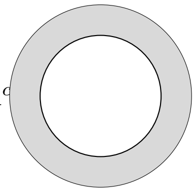
5. Union of A and B

6. $A \cap B$ 

6. Intersection of A and B

7. $A \setminus B$ 

7. A minus B

8. A^c 

8. Complement of A

3.2 Axioms of Set Theory (Zermelo Fraenkel)

This are Zermelo Fraenkel axioms of set theory including *The Axiom of Choice*.

Axiom of Extensionality: $\forall A, B : A = B \Rightarrow (\forall C : C \in A \Leftrightarrow C \in B)$

Empty-Set Axiom: $\exists \emptyset : \forall X : X \notin \emptyset$

Axiom of Pairing: $\forall A, B : \exists C : \forall D : D \in C \Leftrightarrow (D = A \vee D = B)$

Axiom of Union: $\forall A : \exists B : \forall C : C \in B \Leftrightarrow (\exists D : C \in D \wedge D \in A)$

Axiom of Infinity: $\exists N : \emptyset \in N \wedge (\forall x : x \in N \Rightarrow x \cup \{x\} \in N)$

Axiom Schema of Specification: $\forall A : \exists B : \forall C : C \in B \Leftrightarrow (C \in A \wedge P(C))$

Axiom Schema of Replacement: $\forall A : \exists B : \forall y : y \in B \Rightarrow \exists x \in A : y = F(x)$

Powerset Axiom: $\forall A : \exists B : \forall C : C \subseteq A \Rightarrow C \in B$

Foundation Axiom: $\forall A \neq \emptyset : \exists B \in A : A \cap B = \emptyset$

Axiom of Choice: $\forall X : ([\forall A \in X : A \neq \emptyset] \wedge [\forall B, C \in X : B \neq C \Rightarrow B \cap C = \emptyset])$
$\Rightarrow \exists Y : \forall I \in X : \exists! J \in Y : J \in I$

3.3 The Cartesian Product

The Cartesian product of two sets A and B, written $A \times B$, is the set of all ordered pairs in which the first element belongs to A and the second belongs to B:

$$A \times B = \{(a, b) : a \in A, \wedge\ b \in B\}.$$

Example:

Table 1: Cartesian Product of $A = \{1, 2, 3\}$ and $B = \{4, 5, 6\}$

	$b \in B$		
$A \times B$	4	5	6
$a \in A$			
1	$(1, 4)$	$(1, 5)$	$(1, 6)$
2	$(2, 4)$	$(2, 5)$	$(2, 6)$
3	$(3, 4)$	$(3, 5)$	$(3, 6)$

The general cartesian product of n sets can be written as:

$$X_{i=1}^{n+1} A_i = (X_{i=1}^{n} A_i) \times A_{n+1} \quad \text{with} \quad X_{i=1}^{1} A_i = A_1$$

When $A_i = M$ for all i :

$$M^n := M \times M \times \cdots \times M = X_{i=1}^{n} M \quad \text{with} \quad M^1 = M$$

3.4 Laws of Set Algebra

Let X be the universal set and $A, B, C \subseteq X$.

- $\emptyset \subseteq A$
- $A \subseteq B \iff A \cap B = A \iff A \cup B = B \iff X \setminus B \subseteq X \setminus A \iff B \subseteq A$
- $A \cup B = B \cup A$ (Commutative Law)
- $A \cap B = B \cap A$ (Commutative Law)
- $(A \cup B) \cup C = A \cup (B \cup C)$ (Associative Law)
- $(A \cap B) \cap C = A \cap (B \cap C)$ (Associative Law)
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributive Law)
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributive Law)
- $A \cup A = A$ and $A \cap A = A$ (Idempotent Law)
- $A \setminus B = A \cap (X \setminus B) = A \cap \overline{B}$
- $B = \overline{A} \iff (A \cup B = X \wedge A \cap B = \emptyset)$ (Disjoint Partition of X)
- $\overline{A} \cap \overline{B} = \overline{A \cup B}$ (De Morgan's Law)
- $\overline{A} \cup \overline{B} = \overline{A \cap B}$ (De Morgan's Law)
- $\overline{\overline{A}} = A$ (Double Negation)

3.4.1 Proof of De Morgans's Law for sets and logic

The complement of $A \cup B$ is $\overline{(A \cup B)}$, and Law (11) on disjoint decomposition states:

$$B = \overline{A} \iff (A \cup B = X) \wedge (A \cap B = \emptyset)$$

So define $\overline{C} := A \cup B$ and $D := \overline{A} \cap \overline{B}$, and use Law (11) to show the disjoint decomposition:

$$D = C \iff A \cap B = A \cup B$$

a To show:

$$D \cup C = X \iff (\overline{A} \cap \overline{B}) \cup (A \cup B) = X$$

$$\begin{aligned}
(\overline{A} \cap \overline{B}) \cup (A \cup B) &= (\overline{A} \cup A \cup B) \cap (\overline{B} \cup A \cup B) \quad \text{(Law (8))} \\
&= (X \cup B) \cap (X \cup A) \\
&= X \cap X \\
&= X
\end{aligned}$$

b To show:

$$D \cap C = \emptyset \iff (\overline{A} \cap \overline{B}) \cap (A \cup B) = \emptyset$$

$$\begin{aligned}
(\overline{A} \cap \overline{B}) \cap (A \cup B) &= (A \cap B \cap A) \cup (A \cap B \cap B) \quad \text{(Law (7))} \\
&= (\overline{A} \cap A \cap \overline{B}) \cup (\overline{A} \cap \overline{B} \cap B) \\
&= (\emptyset \cap \overline{B}) \cup (\overline{A} \cap \emptyset) \\
&= \emptyset \cup \emptyset \\
&= \emptyset
\end{aligned}$$

3.5 Indexed Sets

Let X be a set, and $A_i \subseteq X$ for all $i \in J$, where J is the index set.

a If $J = \{1, 2, \ldots, n\}$**:**

$$\bigcup_{i=1}^{n} A_i := A_1 \cup A_2 \cup \cdots \cup A_n = \{x \mid \exists i \in J \ (x \in A_i)\}$$

$$\bigcap_{i=1}^{n} A_i := A_1 \cap A_2 \cap \cdots \cap A_n = \{x \mid \forall i \in J \ (x \in A_i)\}$$

$$X_{i=1}^{n} A_i = \{(a_1, \ldots, a_n) \mid a_i \in A_i\}$$

b If J **is any set:**

$$\bigcup_{i \in J} A_i := \{x \mid \exists i \in J \ (x \in A_i)\}$$

$$\bigcap_{i \in J} A_i := \{x \mid \forall i \in J \ (x \in A_i)\}$$

c If J **is any set, then** $(A_i)_{i \in J}$ **are pairwise disjoint if and only if:**

$$\forall i_1, i_2 \in J, \ i_1 \neq i_2 \Rightarrow A_{i_1} \cap A_{i_2} = \emptyset$$

d If J **is any set, then** $(A_i)_{i \in J}$ **forms a (disjoint) decomposition of** X **if and only if:**

$$(A_i)_{i \in J} \text{ are pairwise disjoint and } \bigcup_{i \in J} A_i = X$$

3.5.1 More Partitions Laws

Let $A_i, B_j \subseteq X$ for $i \in I$ and $j \in J$. Then the following holds:

- **De Morgan's Laws:**

$$\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \overline{A_i} \quad \text{and} \quad \overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \overline{A_i}$$

-

$$\bigcap_{i \in I} A_i \cup \bigcap_{j \in J} B_j = \bigcap_{i,j} (A_i \cup B_j) \quad \text{with} \quad \bigcap_{i,j} = \bigcap_{(i,j) \in I \times J}$$

-

$$\bigcup_{i \in I} A_i \cap \bigcup_{j \in J} B_j = \bigcup_{i,j} (A_i \cap B_j) \quad \text{with} \quad \bigcup_{i,j} = \bigcup_{(i,j) \in I \times J}$$

Here, $I = \{1, 2, 3, \ldots, n\}$ and $J = \{1, 2, 3, \ldots, m\}$. Then:

$$I \times J = \{(i, j) \mid 1 \leq i \leq n, 1 \leq j \leq m, i, j \in \mathbb{N}\}$$

$$= \{(1,1), (1,2), \ldots, (1,m), (2,1), (2,2), \ldots, (2,m), \ldots, (n,1), (n,2), \ldots, (n,m)\}$$

3.6 Cardinality

The cardinality of a set is the number of elements in that set.

Let A and B be finite sets with $|A| = n$, $|B| = m$, and let X be the finite universal set. Then the following holds:

Cardinality of a Set

$$A = (A \cap B) \cup (A \setminus B), \quad |A| = |A \cap B| + |A \setminus B|$$

Cardinality of the complements

$$|A| = |X \setminus A| = |X| - |A|$$

$$A \setminus B = A \cap (X \setminus B), \quad |A \setminus B| = |A| - |A \cap B|$$

Cardinality of the Cartesian Product

$$|A \times B| = |A| \cdot |B| = n \cdot m$$

Inclusion-Exclusion Formula for Two Disjoint Sets

$$|A \cup B| = |A| + |B| = n + m$$

Inclusion-Exclusion Formula for Two Non-Disjoint Sets Let $|A \cap B| = k$, then:

$$|A \cup B| = |A| + |B| - |A \cap B| = n + m - k \quad \text{(since we do not count the intersection twice)}$$

Inclusion-Exclusion Formula for Three Non-Disjoint Sets

$$|A \cup B \cup C| = |(A \cup B) \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C|$$
$$= |A| + |B| - |A \cap B| + |C| - |(A \cap C) \cup (B \cap C)|$$
$$= |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

3.6.1 General Formula for the cardinality of the union of sets

$$\left| \bigcup_{i=1}^{n} M_i \right| = \sum_{I \subseteq \{1,\ldots,n\}, I \neq \emptyset} (-1)^{|I|-1} \left| \bigcap_{i \in I} M_i \right|$$

3.7 The Power Set

The Power Set of a set is the set of all subsets of a given set

$$P(x) := \{M : M \subset X\}$$

Its cardinality is 2^n with n being the number of elements in the original set X.

3.8 Family of Subsets

Let X be a non empty set. A subset $\mathscr{F}$ of the power set of X is called a set system of X

3.9 Partition

Let X be a non empty set. A subset $\mathscr{F}$ of the power set of X is called a partition if:
I $M \neq \emptyset \ \forall M \in \mathscr{F}$
II $\bigcap \mathscr{F} = X$
III $M_1 \bigcap M_2 \neq \emptyset \implies M_1 = M_2 \ \forall \ M_1, M_2 \in \mathscr{F}$

- Every equivalence relation corresponds to a partition:

  $$J : \{R : R \text{ is an equivalence relation on } X\} \to \{F : F \text{ is a partition of } X\}$$

  where $J(R) := X/R$ is a bijection.
- If F is a partition of X, then we can define an equivalence relation R_F by:

  $$R_F := \{(x, y) \in X \times X : \exists M \in F \text{ such that } x, y \in M\}$$

  Then R_F is an equivalence relation on X.

3.10 Family of Subsets Operations

Let $\mathscr{F}$ be a system of sets (a family of subsets) on the set X. We define:

$$\bigcup_{M \in F} M := \bigcup F := \{x \in X : \text{there exists } M \in F \text{ such that } x \in M\},$$

$$\bigcap_{M \in F} M := \bigcap F := \{x \in X : x \in M \text{ for all } M \in F\}.$$

4 Relations, Maps and Functions

A relation in mathematics is a connection or relationship between elements of two sets. It's formally defined as a subset of the Cartesian product of the sets.

For example, if we have sets A and B, a relation R from A to B consists of ordered pairs (a, b) where $a \in A$ and $b \in B$, such that a is related to b according to some rule or property.

Common types of relations include:

- Functions (special relations where each input has exactly one output)
- Equivalence relations (reflexive, symmetric, and transitive)
- Partial orders (reflexive, antisymmetric, and transitive)

Relations can be represented using diagrams, matrices, or sets of ordered pairs, and they're fundamental to many areas of mathematics including algebra, calculus, and discrete mathematics.

4.1 Types of relations

Let A be a set and X be a relation on A.

- **Reflexive:** $\forall a \in A : (a, a) \in X \quad$ (or written as $a \sim a$)
- **Irreflexive:** $\forall a \in A : (a, a) \notin X$
- **Symmetric:** $\forall a, b \in A : (a, b) \in X \Rightarrow (b, a) \in X \quad$ (or written as $(a \sim b) \Rightarrow (b \sim a)$)
- **Antisymmetric:** $\forall a, b \in A : (a, b) \in X$ and $(b, a) \in X \Rightarrow a = b$
- **Transitive:** $\forall a, b, c \in A : (a, b) \in X$ and $(b, c) \in X \Rightarrow (a, c) \in X \quad$ (or written as $(a \sim b)$ and $(b \sim c) \Rightarrow (a \sim c)$)
- **Total:** $\forall a, b \in A : a \neq b \Rightarrow (a, b) \in X$ or $(b, a) \in X$

4.2 Equivalence relation

An equivalence relation is a relation that is symmetric, transitive and reflexive.
Example:

$$R := \{(a, b) \in \mathbb{N} \times \mathbb{N} : a = b\}$$

4.3 The Graph

$$\text{graph}(f) := \{(x, f(x)) : x \in X\}$$

4.4 The identity

$$\text{id}(f) := \; idx := (x, x) : x \in X$$

4.5 Image and Domain

4.5.1 Image

The image (or range) of a relation R from set A to set B is the set of all elements in B that are related to at least one element in A. Formally, if $R \subseteq A \times B$ is a relation, then the image of R is defined as:

$$\text{Im}(R) = \{b \in B \mid \exists a \in A \text{ such that } (a, b) \in R\}$$

In other words, the image consists of all the ¡¡output¿¿ values that appear in the ordered pairs of the relation. For example, if $R = \{(1, 4), (2, 5), (3, 4), (2, 6)\}$, then $\text{Im}(R) = \{4, 5, 6\}$.

4.5.2 Domain

The domain of a relation R from set A to set B is the set of all elements in A that are related to at least one element in B. Formally, if $R \subseteq A \times B$ is a relation, then the domain of R is defined as:

$$\text{Dom}(R) = \{a \in A \mid \exists b \in B \text{ such that } (a, b) \in R\}$$

In other words, the domain consists of all the ¡¡input¿¿ values that appear in the ordered pairs of the relation. For example, if $R = \{(1, 4), (2, 5), (3, 4), (2, 6)\}$, then $\text{Dom}(R) = \{1, 2, 3\}$.

4.6 Equivalence Class

Let $\sim$ be an equivalence relation on a set A. For an element $x \in A$, the **equivalence class** of x, denoted by $[x]$, is the set of all elements in A that are equivalent to x. Formally, it is defined as:

$$[x] := \{y \in A \mid x \sim y\} \subseteq A$$

In other words, the equivalence class of x contains all elements y in A such that x is related to y under the equivalence relation $\sim$.

4.7 Quotient Space

Let $\sim$ be an equivalence relation on a set A. The **quotient space** of A by $\sim$, denoted by $A/\sim$ (or sometimes A/R), is the set of all distinct equivalence classes of elements in A. Formally, it is defined as:

$$A/\sim := \{[x] \mid x \in A\}$$

The quotient space $A/\sim$ is a partition of the original set A into disjoint equivalence classes. Each element of the quotient space is an equivalence class $[x]$, which itself is a subset of A.

4.8 Definition of a Map

A map (or function) from a set A to a set B, denoted as $f : A \to B$, is a relation that associates each element of the set A with exactly one element of the set B.

Formally, a function $f : A \to B$ is a subset of $A \times B$ such that for every $a \in A$, there exists exactly one $b \in B$ where $(a, b) \in f$. We typically write $f(a) = b$ to indicate that f maps the element a to the element b.

Example: Let $A = \{1, 2, 3\}$ and $B = \{x, y, z\}$. A possible function $f : A \to B$ could be defined as:

$$f(1) = x, \quad f(2) = y, \quad f(3) = z$$

This function can also be represented as the set of ordered pairs $\{(1, x), (2, y), (3, z)\}$.

4.9 Composition of Maps

The composition of two functions is the operation of applying one function to the result of another function. If we have functions $f : A \to B$ and $g : B \to C$, then the composition of g and f, denoted as $g \circ f$ (read as "g composed with f"), is a function from A to C defined by:

$$(g \circ f)(a) = g(f(a)) \quad \text{for all } a \in A$$

The composition applies f first, then applies g to the result. Note that the codomain of f must match the domain of g for the composition to be defined.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$ and $g : \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = x + 3$. Then:

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 3$$

$$(f \circ g)(x) = f(g(x)) = f(x + 3) = (x + 3)^2 = x^2 + 6x + 9$$

Note that $g \circ f \neq f \circ g$ in general, which shows that function composition is not commutative.

4.10 Types of Functions

4.10.1 Injective Functions

An injective function (also called a one-to-one function) is a function that maps distinct elements from the domain to distinct elements in the codomain.

Formally, a function $f : A \to B$ is injective if for all $a_1, a_2 \in A$:

$$a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$$

Equivalently, using the contrapositive:

$$f(a_1) = f(a_2) \Rightarrow a_1 = a_2$$

4.10.2 Surjective Functions

A surjective function (also called an onto function) is a function whose image equals its codomain, meaning that every element in the codomain has at least one preimage in the domain.

Formally, a function $f : A \to B$ is surjective if:

$$\forall b \in B, \exists a \in A \text{ such that } f(a) = b$$

4.10.3 Bijective Functions

A bijective function (also called a one-to-one correspondence) is a function that is both injective and surjective. In other words, every element in the codomain is mapped to by exactly one element in the domain.

Formally, a function $f : A \to B$ is bijective if it is both:

- Injective: $\forall a_1, a_2 \in A, a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$
- Surjective: $\forall b \in B, \exists a \in A$ such that $f(a) = b$

Bijective functions establish a perfect pairing between elements of the domain and codomain, where each element in the domain corresponds to exactly one element in the codomain, and vice versa. A bijection allows us to define an inverse function $f^{-1} : B \to A$.

4.11 Propositions on Images and Preimages under Set Operations

Let $f : X \to Y$ be a function.

4.11.1 Union and Cut Sets

i For subsets $A_1, A_2 \subseteq X$, we have:

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2) \quad \text{and} \quad f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$$

ii For subsets $B_1, B_2 \subseteq Y$, we have:

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2) \quad \text{and} \quad f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$$

Proof. We prove the second part of 1 and the first part of 2; the remaining statements are left as exercises.

1. Let $y \in f(A_1 \cap A_2)$. Then there exists $x \in A_1 \cap A_2$ such that $f(x) = y$. Since $x \in A_1$ and $x \in A_2$, it follows that $y \in f(A_1)$ and $y \in f(A_2)$, hence $y \in f(A_1) \cap f(A_2)$. Therefore, every element of $f(A_1 \cap A_2)$ is also an element of $f(A_1) \cap f(A_2)$, so:

$$f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$$

2. Let $x \in f^{-1}(B_1 \cup B_2)$. Then $f(x) \in B_1 \cup B_2$, which means $f(x) \in B_1$ or $f(x) \in B_2$. Thus, $x \in f^{-1}(B_1)$ or $x \in f^{-1}(B_2)$, which implies:

$$x \in f^{-1}(B_1) \cup f^{-1}(B_2)$$

Hence, both sets contain the same elements and are therefore equal:

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$$

4.11.2 Union and Cut of the whole Domain and Range

Let $f : X \to Y$ be a function.

i Let $\mathcal{F}$ be a collection of subsets of X. Then:

$$f\left(\bigcup_{A\in\mathcal{F}} A\right) = \bigcup_{A\in\mathcal{F}} f(A) \quad \text{and} \quad f\left(\bigcap_{A\in\mathcal{F}} A\right) \subseteq \bigcap_{A\in\mathcal{F}} f(A)$$

ii Let $\mathcal{G}$ be a collection of subsets of Y. Then:

$$f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right) = \bigcup_{B\in\mathcal{G}} f^{-1}(B) \quad \text{and} \quad f^{-1}\left(\bigcap_{B\in\mathcal{G}} B\right) = \bigcap_{B\in\mathcal{G}} f^{-1}(B)$$

Proof (partial). We show the first statement of part (ii); the rest follows analogously.
Let $x \in f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right)$. Then:

$$f(x) \in \bigcup_{B\in\mathcal{G}} B \quad \Leftrightarrow \quad \exists B \in \mathcal{G} \text{ such that } f(x) \in B \quad \Leftrightarrow \quad \exists B \in \mathcal{G} \text{ such that } x \in f^{-1}(B)$$

Hence:

$$x \in \bigcup_{B\in\mathcal{G}} f^{-1}(B)$$

It follows that:

$$f^{-1}\left(\bigcup_{B\in\mathcal{G}} B\right) = \bigcup_{B\in\mathcal{G}} f^{-1}(B)$$

4.12 Inverse of a Function

The inverse of a function $f : A \to B$ is a function $f^{-1} : B \to A$ that reverses the operation of f. That is, if f maps an element $a \in A$ to an element $b \in B$, then the inverse function f^{-1} maps b back to a. Formally, a function $f : A \to B$ has an inverse $f^{-1} : B \to A$ if and only if f is bijective (both injective and surjective). The inverse function satisfies the following properties:

$$f^{-1}(f(a)) = a \quad \text{for all } a \in A$$

$$f(f^{-1}(b)) = b \quad \text{for all } b \in B$$

In other words, composing a function with its inverse yields the identity function. That is:

$$f^{-1} \circ f = id_A \quad \text{and} \quad f \circ f^{-1} = id_B$$

where id_A and id_B are the identity functions on sets A and B, respectively.

4.12.1 Steps to Find the Inverse of a Function

To find the inverse of a function $f(x)$, follow these steps:

1. Replace $f(x)$ with y: $y = f(x)$
2. Interchange the variables x and y: $x = f(y)$
3. Solve for y in terms of x: $y = f^{-1}(x)$
4. Verify that the resulting function is indeed the inverse by checking that $f^{-1}(f(x)) = x$ and $f(f^{-1}(x)) = x$

4.12.2 Example: Finding the Inverse of $f(x) = 2x + 3$

Let's apply the steps above to find the inverse of $f(x) = 2x + 3$.

Step 1: Replace $f(x)$ with y.

$$y = 2x + 3$$

Step 2: Interchange the variables x and y.

$$x = 2y + 3$$

Step 3: Solve for y in terms of x.

$$x = 2y + 3 \tag{1}$$
$$x - 3 = 2y \tag{2}$$
$$\frac{x-3}{2} = y \tag{3}$$

So, the inverse function is:

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

Step 4: Verify that $f^{-1}(f(x)) = x$ and $f(f^{-1}(x)) = x$.
Let's verify $f^{-1}(f(x)) = x$:

$$f^{-1}(f(x)) = f^{-1}(2x+3) \tag{4}$$
$$= \frac{(2x+3)-3}{2} \tag{5}$$
$$= \frac{2x}{2} \tag{6}$$
$$= x \tag{7}$$

And let's verify $f(f^{-1}(x)) = x$:

$$f(f^{-1}(x)) = f\left(\frac{x-3}{2}\right) \tag{8}$$
$$= 2\left(\frac{x-3}{2}\right) + 3 \tag{9}$$
$$= (x-3) + 3 \tag{10}$$
$$= x \tag{11}$$

Since both compositions yield the identity function, $f^{-1}(x) = \frac{x-3}{2}$ is indeed the inverse of $f(x) = 2x + 3$.

4.12.3 Properties of the Inverse Function

Let $f : X \to Y$ be a function.

- Assume $x \sim y$ if $f(x) = f(y)$ so is $\sim$ an equivalence relation.
- Consider the quotient set $X_f := X/\sim$, where $\sim$ is the equivalence relation defined by $x \sim x' \iff f(x) = f(x')$. Let $q_f : X \to X_f$ be the canonical projection defined by $q_f(x) = [x]$, and let $\iota_f : f(X) \to Y$ be the inclusion map, $y \mapsto y$. Then the function

  $$\hat{f} : X_f \to f(X), \quad [x] \mapsto f(x)$$

  is a bijection, and the original map f can be written as the composition

  $$f = \iota_f \circ \hat{f} \circ q_f.$$

  Here, q_f is surjective, $\hat{f}$ is bijective, and ι_f is injective. This yields the following commutative diagram:

$$\begin{tikzcd}
X \arrow[rr, "f"] \arrow[d, "q_f"] & & Y \\
X_f \arrow[r, "\hat{f}"] & f(X) \arrow[ur, "\iota_f"] &
\end{tikzcd}$$

4.13 Transformations of a Function

Transformations modify the appearance of a function's graph without altering its basic shape. Here, we examine how different algebraic changes to a function $f(x)$ affect its graph:

- **Vertical Translation:** $f(x) + a$
  - Shifts the graph **upward** if $a > 0$, and **downward** if $a < 0$.
  - Each point on the graph moves vertically by a units.
- **Horizontal Translation:** $f(x + a)$
  - Shifts the graph **left** if $a > 0$, and **right** if $a < 0$.
  - This is opposite of what might be expected: adding to x shifts the graph in the negative direction.
- **Vertical Scaling (Stretch/Compression):** $af(x)$
  - If $|a| > 1$: the graph is **stretched** vertically (taller and narrower).
  - If $0 < |a| < 1$: the graph is **compressed** vertically (shorter and wider).
  - If $a < 0$: includes a reflection across the **x-axis**.
- **Horizontal Scaling (Stretch/Compression):** $f(ax)$
  - If $|a| > 1$: the graph is **compressed** horizontally (narrower).
  - If $0 < |a| < 1$: the graph is **stretched** horizontally (wider).
  - If $a < 0$: includes a reflection across the **y-axis**.
- **Reflection across the x-axis:** $-f(x)$
  - Flips the graph upside-down over the x-axis.
  - Each point (x, y) becomes $(x, -y)$.
- **Reflection across the y-axis:** $f(-x)$
  - Flips the graph left-to-right over the y-axis.
  - Each point (x, y) becomes $(-x, y)$.

5 Mathematical Proofs

In this section I will provide with some examples of different types of proofs.

5.1 Proof by Direct Argument

For any integer n, if n is even, then n^2 is even.

Proof. We will prove this theorem by direct argument.

Assume n is an even integer. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even, as it can be expressed as $2m$ where $m = 2k^2$ is an integer. Therefore, we conclude that if n is even, then n^2 is even. □

5.2 Proof by Contradiction

If n is an integer such that n^2 is even, then n is even.

Proof. We will prove this theorem by contradiction. Assume that n is an integer such that n^2 is even, but n is odd. Then we can write $n = 2k + 1$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

This shows that n^2 is odd, which contradicts our assumption that n^2 is even. Therefore, our assumption that n is odd must be false, and thus n must be even. □

5.3 Proof by Induction

For all $n \in \mathbb{N}$, the sum of the first n positive integers is given by:

$$S(n) = 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

Proof. We will prove this theorem by induction on n.

Base Case: For $n = 1$:

$$S(1) = 1 = \frac{1(1+1)}{2}$$

The base case holds.

Inductive Step: Assume that the statement holds for some $n = k$, i.e., assume that:

$$S(k) = 1 + 2 + 3 + \ldots + k = \frac{k(k+1)}{2}$$

We need to show that the statement holds for $n = k + 1$:

$$S(k+1) = S(k) + (k+1)$$

By the inductive hypothesis, we have:

$$\begin{aligned} S(k+1) &= \frac{k(k+1)}{2} + (k+1) \\ &= \frac{k(k+1)}{2} + \frac{2(k+1)}{2} \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

Thus, the statement holds for $n = k + 1$. By the principle of mathematical induction, the statement holds for all $n \in \mathbb{N}$. □

5.4 Proof by Exhaustion

The only integer solutions to the equation $x^2 + y^2 = 1$ are $(0, 1), (1, 0), (0, -1), (-1, 0)$.

Proof. We will prove this theorem by exhaustion. We will check all possible integer values of x and y such that $x^2 + y^2 = 1$.

The possible integer values for x and y are $-1, 0, 1$. We will check each case:

- If $x = 0$: - Then $y^2 = 1$ gives $y = 1$ or $y = -1$. - Solutions: $(0, 1), (0, -1)$. - If $x = 1$: - Then $y^2 = 0$ gives $y = 0$. - Solution: $(1, 0)$. - If $x = -1$: - Then $y^2 = 0$ gives $y = 0$. - Solution: $(-1, 0)$.

Thus, the only integer solutions to the equation are:

$$(0, 1), (1, 0), (0, -1), (-1, 0)$$

□

5.5 Proof by Cases

For any integer n, n^2 is even if and only if n is even.

Proof. We will prove this theorem by cases.

Case 1: Assume n is even. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even.

Case 2: Assume n is odd. Then we can write $n = 2k + 1$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

This shows that n^2 is odd.

Since both cases have been considered, we conclude that n^2 is even if and only if n is even. □

5.6 Proof by Construction

There exists an irrational number x such that x^2 is rational.

Proof. We will construct an irrational number x such that x^2 is rational.

Let $x = \sqrt{2}$. We know that $\sqrt{2}$ is irrational. Now, we compute x^2:

$$x^2 = (\sqrt{2})^2 = 2$$

Since 2 is a rational number, we have constructed an irrational number $x = \sqrt{2}$ such that $x^2 = 2$ is rational. Therefore, the theorem is proved. □

5.7 Proof by Counterexample

The statement "All prime numbers are odd" is false.

Proof. To prove this theorem, we will provide a counterexample.

The number 2 is a prime number, as its only divisors are 1 and 2. However, 2 is even, which contradicts the statement that all prime numbers are odd.

Therefore, the statement "All prime numbers are odd" is false. □

5.8 Proof by Contrapositive

If n is an integer such that n^2 is odd, then n is odd.

Proof. We will prove this theorem by contrapositive. The contrapositive of the statement is: If n is an integer such that n is even, then n^2 is even.

Assume n is even. Then we can write $n = 2k$ for some integer k.

Now, we compute n^2:

$$n^2 = (2k)^2 = 4k^2 = 2(2k^2)$$

This shows that n^2 is even.

Since the contrapositive statement is true, the original statement "If n^2 is odd, then n is odd" is also true. □

5.9 Proof by Reduction to Absurdity

The square root of 2 is irrational.

Proof. We will prove this theorem by reduction to absurdity. Assume that $\sqrt{2}$ is rational. Then we can write:

$$\sqrt{2} = \frac{p}{q}$$

where p and q are integers with no common factors (i.e., the fraction is in simplest form).

Squaring both sides gives:

$$2 = \frac{p^2}{q^2}$$

Rearranging gives:

$$p^2 = 2q^2$$

This implies that p^2 is even, and therefore p must be even (since the square of an odd number is odd). Let $p = 2k$ for some integer k. Substituting this back into the equation gives:

$$(2k)^2 = 2q^2$$

$$4k^2 = 2q^2$$

$$2k^2 = q^2$$

This implies that q^2 is even, and therefore q must also be even. Since both p and q are even, they have a common factor of 2, which contradicts our assumption that p and q have no common factors. Therefore, our assumption that $\sqrt{2}$ is rational must be false, and thus $\sqrt{2}$ is irrational. □

5.10 Proof by Analogy

The set of rational numbers is dense in the set of real numbers.

Proof. We will prove this theorem by analogy.

Consider the set of rational numbers $\mathbb{Q}$ and the set of real numbers $\mathbb{R}$. The density of $\mathbb{Q}$ in $\mathbb{R}$ means that between any two real numbers, there exists a rational number.

For example, between the real numbers 1 and 2, we can find the rational number $\frac{3}{2} = 1.5$. Similarly, between any two real numbers a and b (where $a < b$), we can find a rational number $r = \frac{a+b}{2}$.

This shows that the set of rational numbers is dense in the set of real numbers. Therefore, the theorem is proved by analogy. □

6 The Natural Numbers

In this section we will take a look at the natural numbers, which are the numbers we use for counting. The natural numbers are defined as follows: This not going not be a deep dive just a look at the axioms and the basic construction of the natural numbers. The natural numbers are defined as follows:

We will now define the set of natural numbers, $\mathbb{N}$, via the following 9 axioms. These axioms are known as the **Peano Axioms**. The first 4 axioms define equality on the set $\mathbb{N}$.

Axiom 1: For every $x \in \mathbb{N}$, we have $x = x$. (Reflexivity)

Axiom 2: For every $x, y \in \mathbb{N}$, if $x = y$ then $y = x$. (Symmetry)

Axiom 3: For every $x, y, z \in \mathbb{N}$, if $x = y$ and $y = z$ then $x = z$. (Transitivity)

Axiom 4: For all x, y, if $x \in \mathbb{N}$ and $x = y$, then $y \in \mathbb{N}$. (Closure of Equality)

The remaining 5 axioms define the structure of $\mathbb{N}$:

Axiom 5: $1 \in \mathbb{N}$

Axiom 6: If $x \in \mathbb{N}$, then the successor $S(x) \in \mathbb{N}$.

Axiom 7: There is no $x \in \mathbb{N}$ such that $S(x) = 1$.

Axiom 8: For all $x, y \in \mathbb{N}$, if $S(x) = S(y)$, then $x = y$.

Axiom 9: Let $P(x)$ be a statement about the natural number x. If:

- $P(1)$ is true, and
- for all $n \in \mathbb{N}$, if $P(n)$ is true, then $P(S(n))$ is also true,

then $P(x)$ is true for all $x \in \mathbb{N}$. (Mathematical Induction)

As shorthand, we denote:

$$S(1) = 2, \quad S(S(1)) = 3, \quad S(S(S(1))) = 4, \quad \text{and so on.}$$

6.1 Propositions and Proofs

6.1.1 Proposition 1: $n \neq m \implies S(n) \neq S(m)$

Proof: We will prove this by contradiction. Assume $n \neq m$ and $S(n) = S(m)$. By Axiom 8, we have $n = m$, which is a contradiction. Therefore, $S(n) \neq S(m)$.

6.1.2 Proposition 2: For any $n \in \mathbb{N}$, $n \neq S(n)$

Proof: $M = \{n \in mathbbN \| n \neq S(n)\}$ By $1 \neq S(n)$ for any $n \in \mathbb{N}$, this implies that 1 is part of the set M. This implies that $S(n) \neq S(S(n)) \implies S(n) \in M$ By Axiom 9 $M = \mathbb{N}$

6.1.3 Proposition 3: $n \neq 1 \ \exists m \in \mathbb{N} \mid n = S(m)$

Proof:

$$M = \{1\} \cup \{n \in \mathbb{N} | \text{Proposition 3 is true}\}$$

We know that 1 ins in the set M. An by prosposition 1 we know that

$$S(n) = S(S(m)) \implies S(n) \in M$$

And by Axiom 9 we know that $M = \mathbb{N}$

6.2 Definition of Addition in $\mathbb{N}$

For any pair n,m $\in \mathbb{N}$ there is a unique way to define

$$Add(n,m) = n + m$$

1. **Base Case:** $n + 1 = S(n)$
2. **Inductive Step:** $n + S(m) = S(n+m) \iff S(n+m)$

Uniqueness: Suppose: A & B satisfy our conditions. Fix n and then let $M = \{m \in \mathbb{N} | A(n,m) = B(n,m)\}$. Then

$$A(n,1) = S(n) = B(n,1) \implies 1 \in M$$

$$m \in M \implies A(n,m) = B(n,m) \implies A(n,S(m)) = S(A(n,m)) = S(B(n,m)) = B(n,S(m))$$
$$\implies A(n,S(m)) = B(n,S(m))$$

by Axiom 9 we know that $M = \mathbb{N}$ and A = B
Construction: For $n = 1$ Define $A(n,m) = S(m)$

1. $A(n,1) = S(n) = S(1)$
2. $A(n,S(m)) = S(A(n,m)) = S(S(m))$

Define: $A(S(n),S(m)) = S(A(n,m))$

1. $A(S(n),1) = S(A(n,1)) = S(S(n))$
2. $A(S(n),S(m)) = S(A(n,m)) = S(S(m))$

Commutativity of Addition: The prosposition says $n + m = m + n$ for any $n, m \in \mathbb{N}$. We will prove this by induction on m. Fix n and consider $M = \{n \in \mathbb{N} | A(n,m) = A(n,m)\}$
Now recall that $A(n,1) = S(n)$
For $n = 1$:

$$A(n,1) = S(1) = A(1,n) \implies 1 \in \mathbb{N}$$

also $A(n,k) = 1 + k \implies 1 + m = S(m) \implies 1 + m = m + 1 \implies 1 \in \mathbb{N}$
Suppose: $n \in \mathbb{N} \implies n + m = m + n$ or $A(n,m) = A(m,n)$
By construction $A(S(n),m) = S(A(n,m))$ and by definition $A(S(n),m) = S(A(n,m)) = A(m,S(n)) = S(n) + m = m + S(n) \implies S(n) \in M \implies$ by induction $M = \mathbb{N}$.

7 The Archimedean Principle

The Archimedean Principle is a fundamental property of the real numbers $\mathbb{R}$, and it essentially states that the natural numbers are unbounded in $\mathbb{R}$.

7.1 Statement

For any real number $x \in \mathbb{R}$, there exists a natural number $n \in \mathbb{N}$ such that $n > x$.

In other words, no matter how large a real number you choose, there is always a natural number that is larger. Similarly, for any positive real number $\epsilon > 0$, there exists a natural number $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$.

7.2 Proof

We prove the Archimedean Principle by contradiction.

Assume that there exists some real number $x \in \mathbb{R}$ such that $n \leq x$ for all $n \in \mathbb{N}$. That is, x is an upper bound for the set $\mathbb{N} \subset \mathbb{R}$.

Let $S = \sup(\mathbb{N})$, the least upper bound of $\mathbb{N}$. Then $S - 1 < \sup(\mathbb{N})$, so $S - 1$ is not an upper bound of $\mathbb{N}$. Hence, there exists $n_0 \in \mathbb{N}$ such that:

$$n_0 > S - 1 \Rightarrow n_0 + 1 > S$$

But $n_0 + 1 \in \mathbb{N}$, which contradicts the assumption that S is an upper bound of $\mathbb{N}$. Therefore, our assumption must be false, and the theorem is proven. $\mathfrak{QED}$

7.3 Equivalent Formulations

The Archimedean Principle is often stated in different but equivalent ways:

- For any $\epsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$.
- For any $a, b \in \mathbb{R}$ with $a > 0$, there exists $n \in \mathbb{N}$ such that $na > b$.

7.4 Applications

1. **Density of Rational Numbers:** The Archimedean Principle helps in proving that between any two real numbers, there exists a rational number.
2. **Limits and Infinitesimals:** It ensures that sequences like $\left\{\frac{1}{n}\right\}$ converge to 0, foundational in real analysis and calculus.
3. **Bounding Functions:** It is used in analysis to show that functions do not grow faster than natural numbers in certain contexts.
4. **Non-Existence of Infinitely Small Numbers:** The principle implies that real numbers do not contain infinitesimals (nonzero numbers smaller than all $\frac{1}{n}$), distinguishing $\mathbb{R}$ from non-standard number systems.

8 Fundamental Theorem of Arithmetic

- The Fundamental Theorem of Arithmetic says that every integer greater than 1 can be factored uniquely into a product of primes.
- Euclid's lemma says that if a prime divides a product of two numbers, it must divide at least one of the numbers.
- The least common multiple $[a, b]$ of nonzero integers a and b is the smallest positive integer divisible by both a and b.

Fundamental Theorem of Arithmetic: Every integer greater than 1 can be written in the form

$$p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

where $n_i \geq 0$ and the p_i are distinct primes. The factorization is unique, except possibly for the order of the factors.

Example.

$$4312 = 2 \cdot 2156 = 2 \cdot 2 \cdot 1078 = 2 \cdot 2 \cdot 2 \cdot 539 = 2 \cdot 2 \cdot 2 \cdot 7 \cdot 77 = 2 \cdot 2 \cdot 2 \cdot 7 \cdot 7 \cdot 11$$

That is,

$$4312 = 2^3 \cdot 7^2 \cdot 11$$

8.1 Lemmas

Lemma. If $m \mid pq$ and $\gcd(m, p) = 1$, then $m \mid q$.

Proof. Write $1 = \gcd(m, p) = am + bp$ for some $a, b \in \mathbb{Z}$. Then

$$q = amq + bpq$$

Since $m \mid amq$ and $m \mid bpq$ (because $m \mid pq$), we conclude $m \mid q$.

Lemma. If p is prime and $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_i$ for some i.

Proof. (Case $n = 2$): Suppose $p \mid a_1 a_2$, and $p \nmid a_1$. Then $\gcd(p, a_1) = 1$, and by the previous lemma, $p \mid a_2$.

For general $n > 2$: Assume the result is true for $n - 1$. Suppose $p \mid a_1 a_2 \cdots a_n$. Group as $(a_1 a_2 \cdots a_{n-1}) a_n$.

By the $n = 2$ case, either $p \mid a_n$ or $p \mid a_1 a_2 \cdots a_{n-1}$, and by induction, $p \mid a_i$ for some i.

8.2 Proof of the Fundamental Theorem

Existence: Use induction on $n > 1$. Base case: $n = 2$ is prime.

Inductive step: If n is prime, done. Otherwise $n = ab$, with $1 < a, b < n$. By induction, both a and b factor into primes, so n does too.

Uniqueness: Suppose:

$$p_1^{m_1} \cdots p_j^{m_j} = q_1^{n_1} \cdots q_k^{n_k}$$

with all p_i and q_i distinct primes.

Since p_1 divides the LHS, it divides the RHS. So $p_1 \mid q_i^{n_i}$ for some i, hence $p_1 = q_i$. Reorder so $p_1 = q_1$. Then:

If $m_1 > n_1$, divide both sides by $q_1^{n_1}$:

$$p_1^{m_1 - n_1} \cdots p_j^{m_j} = q_2^{n_2} \cdots q_k^{n_k}$$

But then p_1 divides LHS but not RHS, contradiction. So $m_1 = n_1$. Cancel and repeat.

Eventually, all p_i match with some q_i, and the exponents are equal. So the factorizations are the same up to order.

8.3 Least Common Multiple

The least common multiple of a and b, denoted $[a, b]$, is the smallest positive integer divisible by both.

Example:

$$[6, 4] = 12, \quad [33, 15] = 165$$

Fact:

$$[a, b] \cdot \gcd(a, b) = ab$$

Let:

$$a = p_1 \cdots p_l q_1 \cdots q_m, \quad b = q_1 \cdots q_m r_1 \cdots r_n$$

Then:

$$\begin{aligned} \gcd(a, b) &= q_1 \cdots q_m \\ [a, b] &= p_1 \cdots p_l q_1 \cdots q_m r_1 \cdots r_n \\ ab &= p_1 \cdots p_l q_1^2 \cdots q_m^2 r_1 \cdots r_n \end{aligned}$$

So:

$$[a, b] \cdot \gcd(a, b) = ab$$

Example:

$$\gcd(36, 90) = 18, \quad [36, 90] = 180, \quad 36 \cdot 90 = 32400 = 18 \cdot 180$$

9 Real Numbers

Let K be an ordered field.

- On the set
$$\mathrm{ch}(K) := \{x : \mathbb{N} \to K \mid x \text{ is a Cauchy sequence}\}$$
and on the set
$$c(K) := \{x : \mathbb{N} \to K \mid x \text{ is a convergent sequence}\},$$
we can define an addition and a multiplication using 1.2.55 and 1.2.57 as follows:
If $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ are Cauchy sequences (respectively, convergent sequences), then their sum is defined as
$$x + y := (x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}} := (x_n + y_n)_{n\in\mathbb{N}},$$
and their product is defined as
$$x \cdot y := (x_n)_{n\in\mathbb{N}} \cdot (y_n)_{n\in\mathbb{N}} := (x_n \cdot y_n)_{n\in\mathbb{N}}.$$

The sum and product satisfy all field axioms except for the existence of the multiplicative inverse. The zero element is $0_{\mathbb{N}} = (0, 0, \ldots)$, the unit element is $1_{\mathbb{N}} = (1, 1, \ldots)$, and the additive inverse of $x = (x_n)_{n\in\mathbb{N}}$ is $-x = (-x_n)_{n\in\mathbb{N}}$. We demonstrate the distributive law as an example:

Let $x = (x_n)_{n\in\mathbb{N}},\ y = (y_n)_{n\in\mathbb{N}},\ z = (z_n)_{n\in\mathbb{N}}$ be Cauchy sequences (convergent sequences). Then we have:

$$\begin{aligned} x(y+z) &= (x_n)_{n\in\mathbb{N}} \cdot ((y_n)_{n\in\mathbb{N}} + (z_n)_{n\in\mathbb{N}}) = (x_n)_{n\in\mathbb{N}} \cdot (y_n + z_n)_{n\in\mathbb{N}} = (x_n(y_n + z_n))_{n\in\mathbb{N}} \\ &= (x_n y_n + x_n z_n)_{n\in\mathbb{N}} = (x_n y_n)_{n\in\mathbb{N}} + (x_n z_n)_{n\in\mathbb{N}} = xy + xz. \end{aligned}$$

We now aim to construct the ordered field $\mathbb{R}$ of the real numbers; it will have the following properties:

α There exists an injective mapping $j : \mathbb{Q} \to \mathbb{R}$ which respects addition, multiplication, and order, such that the following holds: For all $z, w \in \mathbb{R}$ with $z < w$, there exists an $x \in \mathbb{Q}$ such that
$$z < j(x) < w.$$

β Every Cauchy sequence in $\mathbb{R}$ converges.

Via j, we identify $\mathbb{Q}$ with $j(\mathbb{Q})$ and consider $\mathbb{Q}$ as a subset of $\mathbb{R}$. In $\mathbb{R}$, the following will additionally hold:

γ For all $y > 0$ and $n \in \mathbb{N}$, the equation $x^n = y$ has a solution.

δ Every bounded above subset of $\mathbb{R}$ has a supremum.

We define the following relation on the set $\mathrm{ch}(\mathbb{Q})$ of all Cauchy sequences in $\mathbb{Q}$:

$$x \sim y \quad \text{if and only if} \quad x - y \text{ is a null sequence.}$$

That is, $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ if and only if
$$x_n - y_n \to 0 \quad (n \to \infty).$$

9.1 Definition

The set
$$\mathbb{R} := \{[x]_\sim : x \in \mathrm{ch}(\mathbb{Q})\}$$
is called the set of real numbers.

Analogous to the construction of the rational numbers, the real numbers consist of equivalence classes. Roughly speaking, an equivalence class consists of those Cauchy sequences in $\mathbb{Q}$ that exhibit the same "limit behavior."

Equipped with the addition
$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad [x], [y] \mapsto [x] + [y] := [x + y],$$
and the multiplication
$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad [x], [y] \mapsto [x] \cdot [y] := [xy],$$
$\mathbb{R}$ is a field. The zero element is $[0_{\mathbb{N}}]$, and the unit element is $[1_{\mathbb{N}}]$.

10 Complex Numbers

10.1 What are Complex Numbers?

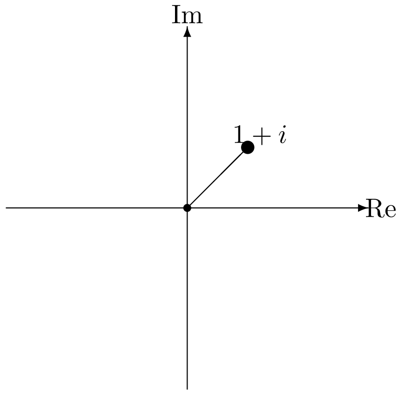
A complex number is a number of the form:

$$z = a + bi,$$

where $a, b \in \mathbb{R}$, and i is the imaginary unit defined by $i^2 = -1$. The set of all complex numbers is denoted by $\mathbb{C}$.

10.2 The Complex Plane

Complex numbers can be represented graphically in the **complex plane**, where the horizontal axis represents the real part and the vertical axis the imaginary part.



The point $1 + i$ is located at (1,1), showing 1 unit on the real axis and 1 unit on the imaginary axis.

10.3 Conjugate of a Complex Number

The **conjugate** of a complex number $z = a + bi$ is denoted $\overline{z}$ and is defined as:

$$\overline{z} = a - bi$$

Geometrically, it reflects the point z across the real axis in the complex plane. Conjugates are useful in division and in finding the modulus, since:

$$z \cdot \overline{z} = a^2 + b^2 = |z|^2$$

10.4 Operations in Cartesian Coordinates

Let $z_1 = a + bi$ and $z_2 = c + di$ be two complex numbers.

- **Addition:**
$$z_1 + z_2 = (a + c) + (b + d)i$$
- **Multiplication:**
$$z_1 \cdot z_2 = (ac - bd) + (ad + bc)i$$
- **Quotient:**
$$\frac{z_1}{z_2} = \frac{(a + bi)}{(c + di)} \frac{(c - di)}{(c - di)} = \frac{(a + bi)(c - di)}{c^2 + d^2} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2}$$

10.5 Polar Coordinates

A complex number can also be expressed in polar form as:

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

where:

$$r = |z| = \sqrt{a^2 + b^2} \quad \text{(modulus)}$$
$$\theta = \arg(z) = \tan^{-1}\left(\frac{b}{a}\right) \quad \text{(argument)}$$

Important: The value of θ depends on the quadrant where the complex number lies:

- Quadrant I: $a > 0, b > 0$ — use $\tan^{-1}(b/a)$
- Quadrant II: $a < 0, b > 0$ — add π to $\tan^{-1}(b/a)$
- Quadrant III: $a < 0, b < 0$ — add π to $\tan^{-1}(b/a)$
- Quadrant IV: $a > 0, b < 0$ — use $\tan^{-1}(b/a)$

10.6 Multiplication and Division in Polar Coordinates

Given:

$$z_1 = r_1 e^{i\theta_1}, \quad z_2 = r_2 e^{i\theta_2},$$

- **Multiplication:**
$$z_1 \cdot z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$
- **Division:**
$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}$$

This polar form is especially useful in simplifying powers and roots of complex numbers using De Moivre's Theorem.

10.7 Exponentiation and Roots (De Moivre's Theorem)

Let $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$ be a complex number in polar form.

10.7.1 Exponentiation

To raise z to the power $n \in \mathbb{N}$, we use De Moivre's Theorem:

$$z^n = r^n(\cos(n\theta) + i\sin(n\theta)) = r^n e^{in\theta}$$

10.7.2 Roots of Complex Numbers

To find the nth roots of a complex number $z = re^{i\theta}$, we use the formula:

$$z^{1/n} = r^{1/n}\left(\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right)\right), \quad k = 0, 1, \ldots, n-1$$

This yields n distinct roots, each separated by an angle of $\frac{2\pi}{n}$ in the complex plane.

10.8 Example: Solve $z^4 = 1 + \sqrt{3}i$

Step 1: Convert RHS to polar form.
Let $w = 1 + \sqrt{3}i$. Real part: $a = 1$, Imaginary part: $b = \sqrt{3}$

$$r = |w| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{1+3} = 2$$
$$\theta = \arg(w) = \tan^{-1}\left(\frac{\sqrt{3}}{1}\right) = \frac{\pi}{3}$$

So,
$$w = 2\left(\cos\left(\frac{\pi}{3}\right) + i\sin\left(\frac{\pi}{3}\right)\right)$$

Step 2: Solve $z^4 = w \Rightarrow z = w^{1/4}$
Using the root formula:

$$z_k = 2^{1/4}\left(\cos\left(\frac{\pi + 2k\pi}{12}\right) + i\sin\left(\frac{\pi + 2k\pi}{12}\right)\right), \quad k = 0, 1, 2, 3$$

So the four roots are:

$$z_0 = 2^{1/4}\left(\cos\left(\frac{\pi}{12}\right) + i\sin\left(\frac{\pi}{12}\right)\right)$$
$$z_1 = 2^{1/4}\left(\cos\left(\frac{5\pi}{12}\right) + i\sin\left(\frac{5\pi}{12}\right)\right)$$
$$z_2 = 2^{1/4}\left(\cos\left(\frac{9\pi}{12}\right) + i\sin\left(\frac{9\pi}{12}\right)\right) = 2^{1/4}\left(\cos\left(\frac{3\pi}{4}\right) + i\sin\left(\frac{3\pi}{4}\right)\right)$$
$$z_3 = 2^{1/4}\left(\cos\left(\frac{13\pi}{12}\right) + i\sin\left(\frac{13\pi}{12}\right)\right)$$

These represent the four complex 4th roots of $1+\sqrt{3}i$, equally spaced around the circle of radius $2^{1/4}$ in the complex plane.

10.9 Solving Equations with Complex Numbers

Solving equations in $\mathbb{C}$ can involve various forms. Here are the most common cases:

1. **Linear Equations:** Solve for z in $az + b = 0$, where $a, b \in \mathbb{C}$, $a \neq 0$:
$$z = -\frac{b}{a}$$

2. **Equations Involving the Conjugate:** Solve for z in equations like $z + \overline{z} = 4$. Let $z = x + iy$, then $\overline{z} = x - iy$. So:
$$z + \overline{z} = 2x \quad \Rightarrow \quad x = 2 \quad \Rightarrow \quad z = 2 + iy$$
The imaginary part remains free unless further constraints are given.

3. **Modulus Equations:** Solve $|z| = r$: Let $z = x + iy$, then:
$$\sqrt{x^2 + y^2} = r \Rightarrow x^2 + y^2 = r^2$$
This is a circle of radius r centered at the origin in the complex plane.

4. **Equations Involving $z \cdot \overline{z}$:** Recall $z \cdot \overline{z} = |z|^2$. For example, solve:
$$z \cdot \overline{z} = 9 \Rightarrow |z| = 3$$
Again, a circle in the complex plane of radius 3.

5. **Quadratic Equations:** Complex roots occur naturally. For example:
$$z^2 + 1 = 0 \Rightarrow z^2 = -1 \Rightarrow z = \pm i$$

6. **General Polynomial Equations:** Use De Moivre's Theorem or polar form. Example:
$$z^n = w \Rightarrow z_k = \sqrt[n]{|w|} \cdot e^{i\left(\frac{\arg(w)+2k\pi}{n}\right)}, \quad k = 0, 1, \ldots, n-1$$

10.9.1 Example 1:

Solve:

$$\left(\frac{2+3i}{1+i}+\frac{4+5i}{2-2i}\right)\hat{z}=\frac{i+2}{i}$$
$$\left(\frac{-3-i}{4}\right)\hat{z}=\frac{i+2}{i}$$
$$\hat{z}=\frac{i+2}{i}:\frac{-3-i}{4}$$

10.9.2 Example 2:

Solve:

$$z-3i+(2-i)\hat{z}+2=0$$

In this case we let $z=x+iy$ and $\hat{z}=x-iy$.

$$z=3i-(2-i)\hat{z}-2$$

$$x+iy=3i-(2-i)(x-iy)-2$$
$$x+iy=3i-[2x-2yi-xi+yi]-2$$
$$x+yi=3i-2+2x+2yi+xi+y$$
$$x+yi=(y-2-2x)+i(3+2y+x)$$
$$x=y-2-2x\ y=3+2y+x$$
$$x=\frac{y-2}{3}\quad y=3+2y+\frac{y-2}{3}=-7$$
$$x=\frac{-7-2}{3}=-3$$
$$z=-3-7i\ \ \hat{z}=-3+7i$$

10.10 The Complex Logarithm

The logarithm of a complex number is multi-valued due to the periodic nature of the complex exponential.
Let $z=re^{i\theta}$ with $r>0$, $\theta\in\mathbb{R}$. Then:

$$\log z=\ln r+i(\theta+2\pi k),\quad k\in\mathbb{Z}$$

Here:

- $\ln r$ is the natural (real) logarithm of the modulus.
- θ is the principal argument $\arg(z)\in(-\pi,\pi]$.
- The term $2\pi k$ accounts for the infinitely many branches of the logarithm in $\mathbb{C}$.

Principal Value: The principal value of the complex logarithm is often written:

$$\operatorname{Log} z=\ln|z|+i\operatorname{Arg}(z),\quad \text{where } \operatorname{Arg}(z)\in(-\pi,\pi]$$

Example: Let $z=-1$. Then:

$$|z|=1,\quad \arg(z)=\pi,\quad \Rightarrow \log(-1)=i(\pi+2\pi k),\quad k\in\mathbb{Z}$$
$$\Rightarrow \operatorname{Log}(-1)=i\pi$$

The multi-valued nature of $\log z$ is crucial in advanced complex analysis, especially in defining analytic continuations and branch cuts.

10.11 Complex Exponents

$a^x \approx 1\left(a+\alpha\frac{x}{N}\right)^N \to e^z := \lim_{N\to\inf}\left(1+\frac{z}{N}\right)^N$ The process above is called linearization of the exponential function by ¡¡zooming¿¿ $\alpha \approx \frac{dy}{d}$
For $e^{ci} = \cos\theta + i\sin\theta$ every exponentiation of a complex number is a rotation in the complex plane.

$$e^{ic} = \lim_{N\to\inf}\left(1+\frac{ic}{N}\right)^N$$

Now imagine that in a sector of a circunference you put trinagles one above the other with base of lenght one and a height of $\frac{C}{N}$ and an angle of δ

$$\tan\delta \approx \delta \text{ for } \delta \ll 1$$

$$1+\frac{ci}{N} = 1\angle\frac{c}{N}, N \gg 1$$

$$e^{ic} = \lim_{N\to\inf}\left(1+\frac{c}{N}\right)^N \to e^{ci} = 1\angle c = \cos c + i\sin c$$

10.12 Euler's Formula Proof

We know that

$$e^{i\pi} = -1 \text{ and } e^{i\theta} = |r|(\cos\theta + i\sin\theta)$$

$$e^z = \lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n \implies e^{i\pi} = \lim_{n\to\infty}\left(1+\frac{i\pi}{n}\right)^n = -1$$

$$\implies \lim_{n\to\infty}|r_n| = 1 \quad \lim_{n\to\infty}\theta = 0+$$

Now we can demonstrate the formula.

$$|r_n| = \left(1+\left|\frac{z}{n}\right)^n\right| \implies \left(\sqrt{1+\frac{\pi}{n}}\right)$$

$$\theta = \sum_{k=1}^{n} n\arctan\frac{\pi}{2} = n\arctan\frac{\pi}{n}$$

$$\lim_{n\to\infty}\left(\sqrt{1+\frac{\pi}{n}}\right)^n = \lim_{n\to\infty}\left(1+\frac{\pi}{2n}\right)^{\frac{n}{2}} = \lim_{n\to\infty} e^{ln\left(1+\frac{\pi}{2}\right)\frac{n}{2}} = e^0 = 1$$

$$\lim_{n\to\infty}\theta = \lim_{n\to\infty} n\arctan\frac{\pi}{n} = \lim_{n\to\infty} n^{-1}\arctan\frac{\pi}{n} = 0$$

Thus for $e^{i\pi} = 1$ for $x = \pi\forall x \in \lim r_n(x) = 1$ and $\lim\theta(x) = x$
$\mathfrak{QED}$

11 Topology

In this section, we introduce essential vocabulary used in topology. Each term is accompanied by a brief explanation and its formal mathematical definition.

11.1 Introduction to topological nomenclature

- **Open Set**
  A subset $U \subseteq X$ of a topological space is called *open* if for every point $x \in U$, there exists an $\varepsilon > 0$ such that the open ball $B_\varepsilon(x) \subseteq U$.
  Intuitively, an open set contains none of its boundary points and every point has some ¡¡wiggle¿¿ room around it.

- **Closed Set**
  A subset $A \subseteq X$ is called *closed* if its complement $X \setminus A$ is open. Equivalently, A contains all its limit points.
  That is, A is closed if it includes its boundary.

- **Interior Point**
  A point $x \in A$ is an *interior point* of $A \subseteq X$ if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq A$.
  The set of all interior points of A is called the *interior* of A, denoted $\mathrm{int}(A)$.

- **Boundary Point**
  A point $x \in X$ is a *boundary point* of a set $A \subseteq X$ if every open ball around x contains both points in A and in $X \setminus A$.
  The set of all boundary points is called the *boundary* of A, denoted ∂A.

- **Accumulation Point / Limit Point**
  A point $x \in X$ is an *accumulation point* of a set $A \subseteq X$ if every open ball $B_\varepsilon(x)$ contains a point of $A \setminus \{x\}$.
  In other words, points of A cluster arbitrarily close to x, even if $x \notin A$.

- **Isolated Point**
  A point $x \in A$ is an *isolated point* if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \cap A = \{x\}$.
  That is, x stands alone in A without other points of A nearby.

- **Compact Set**
  A set $K \subseteq X$ is *compact* if every open cover of K has a finite subcover.
  In $\mathbb{R}^n$, this is equivalent to K being closed and bounded (by the Heine–Borel theorem).

- **Dense Set**
  A subset $D \subseteq X$ is *dense* in X if every point $x \in X$ is either in D or is a limit point of D.
  Equivalently, the closure of D is X, i.e., $\overline{D} = X$.

- **Open Ball** $(B_\varepsilon(x))$
  For a metric space (X, d), the *open ball* centered at $x \in X$ with radius $\varepsilon > 0$ is defined as:

  $$B_\varepsilon(x) := \{y \in X \mid d(x, y) < \varepsilon\}$$

  It represents the set of all points within distance ε from x, excluding the boundary.

12 Fractions, Roots, and Exponents

This is a small chapter to remember the properties of fractions, roots, and exponents. It is not a complex chapter, but it is useful to have it in the compendium.

12.1 Fractions

- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$
- $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$
- $\frac{a}{b} \div \frac{c}{d} = \frac{ad}{bc}$
- $\frac{a}{b} = \frac{c}{d} \iff ad = bc$
- $\frac{ac}{bc} = \frac{a}{b}$

12.2 Roots and Exponents

- $\sqrt{a} \cdot \sqrt{b} = \sqrt{ab}$
- $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$
- $\sqrt[nm]{a} = \sqrt[n]{\sqrt[m]{a}}$
- $a\sqrt[n]{b} = \sqrt[n]{a^n b}$
- $a^m \cdot a^n = a^{m+n}$
- $a^m \div a^n = a^{m-n}$
- $(a^m)^n = a^{mn}$
- $a^{-n} = \frac{1}{a^n}$
- $a^{\frac{m}{n}} = \sqrt[n]{a^m}$
- $a^{\frac{1}{2}} = \sqrt{a}$
- $a^{\frac{1}{3}} = \sqrt[3]{a}$
- $a^{\frac{1}{n}} = \sqrt[n]{a}$
- $a^0 = 1$
- $a^1 = a$

13 Logarithms

The logarithm of a number x with base b is the exponent to which the base must be raised to produce that number. It is denoted as:

$$\log_b(x) = y \iff b^y = x$$

where $b > 0$, $b \neq 1$, and $x > 0$.

$$\log_2(8) = 3 \quad \text{because} \quad 2^3 = 8$$

13.1 Properties of Logarithms

- $\log_b(xy) = \log_b(x) + \log_b(y)$
- $\log_b\left(\frac{x}{y}\right) = \log_b(x) - \log_b(y)$
- $\log_b(x^k) = k \cdot \log_b(x)$
- $\log_b(b) = 1$
- $\log_b(1) = 0$
- $\log_{b^k}(x^w) = \frac{1}{k} \cdot \log_b(x)$
- $\log_b\left(\frac{1}{x}\right) = -\log_b(x)$
- $\log_b(b^x) = x$
- $\log_b(x) = \frac{\log_k(x)}{\log_k(b)}$ for any positive $k \neq 1$
- $e^{\ln(x)} = x$
- If $0 < a < 1$ then $\ln(a)$ is a negative number.

13.2 Fundamental Identity of Logarithms

$$a^{\log_a(x)} = x$$

13.3 Change of Base Formula

$$\log_b(x) = \frac{\log_k(x)}{\log_k(b)}$$

where k is any positive number different from 1.

13.4 The Chain Rule

$$\log_y(a)\log_a(b) = \log_y(b)$$

13.5 The derivative of the Natural Logarithm

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

Let $f(x) = \ln(x)$, then

$$f\prime(x) = \lim_{h\to 0} \frac{\ln(x+h) - \ln(x)}{h}$$

$$= f\prime(x) = \lim_{h\to 0} \frac{\ln\left(\frac{x+h}{x}\right)}{h}$$

$$= f\prime(x) = \lim_{h\to 0} \frac{\ln\left(1 + \frac{h}{x}\right)}{h}$$

$$= f\prime(x) = \lim_{h\to 0} \ln\left(1 + \frac{h}{x}\right)^{\frac{1}{h}}$$

Now let $n = \frac{h}{x}$

$$f\prime(x) = \lim_{n \to 0} \ln{(1+n)^{\frac{x}{h}\frac{1}{x}}}$$

$$f\prime(x) = \lim_{n \to 0} \frac{1}{x} \ln{(1+n)^{\frac{1}{n}}}$$

$$f\prime(x) = \frac{1}{x} \ln\left(\lim_{n \to 0} (1+n)^{\frac{1}{n}}\right)$$

$$f\prime(x) = \frac{1}{x} \ln(e)$$

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

Therefore, the derivative of the natural logarithm is:

$$\frac{d}{dx} \ln(x) = \frac{1}{x}$$

$\mathfrak{QED}$

14 Sum and Product Notation

In mathematics, the sum and product notations are compact ways to represent repeated addition and multiplication, respectively. These notations are essential for working with sequences, series, and algebraic expressions.

14.1 Sum Notation $\sum$

Definition

The summation symbol $\sum$ represents the addition of a sequence of terms:

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

where i is the index of summation, m is the lower bound, and n is the upper bound.

Properties

- **Linearity:**

$$\sum_{i=m}^{n}(a_i + b_i) = \sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i$$

$$\sum_{i=m}^{n} c \cdot a_i = c \cdot \sum_{i=m}^{n} a_i$$

- **Splitting:**

$$\sum_{i=m}^{n} a_i = \sum_{i=m}^{k} a_i + \sum_{i=k+1}^{n} a_i \quad (m \leq k < n)$$

Change of Index

Let $j = i + k$, then:

$$\sum_{i=m}^{n} a_i = \sum_{j=m+k}^{n+k} a_{j-k}$$

Example:

$$\sum_{i=1}^{4} a_i = \sum_{j=2}^{5} a_{j-1}$$

Power Sums and Their Formulas

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

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

$$\sum_{i=1}^{n} i^3 = \left[\frac{n(n+1)}{2}\right]^2$$

$$\sum_{i=1}^{n} i^k = \text{(Higher-order polynomial in } n\text{)}$$

Derivation of $\sum_{i=1}^{n} i^2$ We use the method of finite differences or induction. Assume a quadratic form:

$$\sum_{i=1}^{n} i^2 = An^3 + Bn^2 + Cn$$

Plug in small values of n (e.g., 1, 2, 3), solve the system of equations to find:

$$A = \frac{1}{3}, \quad B = \frac{1}{2}, \quad C = \frac{1}{6} \Rightarrow \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

Telescoping Sum

A telescoping sum is a sum where intermediate terms cancel out, leaving only the first and last terms.
Example:

$$\sum_{i=1}^{n} \left(\frac{1}{i} - \frac{1}{i+1} \right) = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n} - \frac{1}{n+1} = 1 - \frac{1}{n+1}$$

Geometric Series

Finite Geometric Series: For a geometric sequence $a, ar, ar^2, \ldots, ar^{n-1}$:

$$\sum_{i=0}^{n-1} ar^i = a \cdot \frac{1-r^n}{1-r}, \quad r \neq 1$$

Infinite Geometric Series: If $|r| < 1$, then:

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

14.2 Product Notation $\prod$

Definition

The product notation $\prod$ represents repeated multiplication:

$$\prod_{i=m}^{n} a_i = a_m \cdot a_{m+1} \cdot \cdots \cdot a_n$$

Properties

– **Multiplicativity:**

$$\prod_{i=m}^{n} (a_i \cdot b_i) = \left(\prod_{i=m}^{n} a_i \right) \cdot \left(\prod_{i=m}^{n} b_i \right)$$

– **Power Rule:**

$$\prod_{i=m}^{n} a^k = a^{k(n-m+1)}$$

Change of Index

Let $j = i + k$, then:

$$\prod_{i=m}^{n} a_i = \prod_{j=m+k}^{n+k} a_{j-k}$$

Example:

$$\prod_{i=1}^{3} a_i = \prod_{j=2}^{4} a_{j-1}$$

Telescoping Product

A telescoping product occurs when consecutive terms simplify or cancel.

Example:

$$\prod_{i=1}^{n} \frac{i}{i+1} = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{n}{n+1} = \frac{1}{n+1}$$

15 Means and Proofs

In math there are a lot of means. In this section I will show some of them with the correspoinding proofs.

- **Arithmetic Mean:** The arithmetic mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$A = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

- **Geometric Mean:** The geometric mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

- **Harmonic Mean:** The harmonic mean of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

- **Quadratic Mean:** The quadratic mean (or root mean square) of n numbers $x_1, x_2, \ldots, x_n$ is given by:

$$Q = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}}$$

15.1 Proof of the Arithmetic Mean-Geometric Mean Inequality

Let $a_1, a_2, \ldots, a_n > 0$. We will prove by induction that:

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \cdots a_n}$$

with equality if and only if $a_1 = a_2 = \cdots = a_n$.

Base Case: $n = 2$

We want to prove:

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2}$$

Let $a_1, a_2 > 0$. Then by the identity

$$\left(\frac{a_1 - a_2}{2}\right)^2 \geq 0,$$

we get

$$\frac{a_1^2 - 2a_1 a_2 + a_2^2}{4} \geq 0 \Rightarrow a_1^2 + a_2^2 \geq 2a_1 a_2.$$

So,

$$(a_1 + a_2)^2 \geq 4a_1 a_2 \Rightarrow \left(\frac{a_1 + a_2}{2}\right)^2 \geq a_1 a_2,$$

and taking square roots gives the desired result:

$$\frac{a_1 + a_2}{2} \geq \sqrt{a_1 a_2}.$$

For $n \geq 2$

$$A_{n+1} := (\sum_{i=1}^{n+1} a_i)/(n+1) = \frac{a_1 + a_2 + \cdots + a_n + a_{n+1}}{n+1}$$

$$G_{n+1} := \sqrt[n+1]{a_1 a_2 \cdots a_n a_{n+1}} = \sqrt[n+1]{(a_1 a_2 \cdots a_n) a_{n+1}}$$

$$A_{n-1}^{n+1} = (\frac{a_1 + a_2 + \cdots + a_{n+1}}{n+1})^{n-1} = A_{n+1}^{n-1}$$

$$G_{n+1}^{n+1} := (\sqrt[n+1]{a_1 a_2 \cdots a_n a_{n+1}})^{n+1} = (a_1 a_2 \cdots a_n) a_{n+1} = \sqrt[n]{(a_1 a_2 \cdots a_n)^n} a_{n+1}^{n+1} = G_n^n a_{n+1}^{n+1}$$

Then

$$G_{n+1}^{n+1}A_{n+1}^{n-1} = G_n^n a_{n+1}^{n+1}A_{n+1}^{n-1} \leq A_n^n A_{n+1}^{n-1}$$

This comes from

$$G_n \leq A_n$$

$$G_n^n \leq A_n^n$$

$$G_n^n a_{n+1} \leq A_n^n a_{n+1}$$

$$G_n^n a_{n+1}^{n+1}A_{n+1}^{n-1} \leq A_n^n A_{n+1}^{n-1} = \left(A_n^n \left(a_{n+1}A_{n+1}\right)^{n-1}\right)^{\frac{n}{n}}$$

$$\leq A_n^n \left(\frac{a_{n+1} + A_{n+1} + \ldots A_{n+1}}{n}\right)^n$$

$$A_n^n \left(\frac{a_{n+1} + (n-1)A_{n+1}}{n}\right)^n$$

$$\left(A_n \frac{a_{n+1} + (n-1)A_{n+1}}{n}\right)^n$$

Note that

$$\left(A_n \frac{a_{n+1} + (n-1)A_{n+1}}{n}\right) \to \left(\sqrt{A_n \frac{a_{n+1} + (n-1)A_{n+1}}{n}}\right)^2$$

$$\leq \left(\frac{A_n + \frac{a_{n+1}+(n-1)A_{n+1}}{n}}{2}\right)^{2n}$$

Now with power of n we have

$$\left(A_n + \frac{a_{n+1} + (n-1)A_{n+1}}{n}\right)^n \leq \left(\frac{A_n + \frac{a_{n+1}+(n-1)A_{n+1}}{n}}{2}\right)^{2n}$$

$$= \left(\frac{A_n}{2} + \frac{a_{n+1} + (n-1)A_{n+1}}{2n}\right)^{2n}$$

$$= \left(\frac{A_n n}{2n} + \frac{a_{n+1} + (n-1)A_{n+1}}{2n}\right)^{2n}$$

$$= \left(\frac{A_n n + a_{n+1} + (n-1)A_{n+1}}{2n}\right)^{2n}$$

$$= \left(\frac{A_{n+1}(n+1) + (n-1)A_{n+1}}{2n}\right)^{2n}$$

$$= \left(\frac{2nA_{n+1}}{2n}\right)^{2n} = \left(A_{n+1}\right)^{2n}$$

Now we have prooven that $G_{n+1}^{n+1}A_{n+1}^{n-1} \leq \left(A_{n+1}\right)^{2n}$ by dividing both sides by A_{n+1}^{n-1} we get

$$G_{n+1}^{n+1} \leq A_{n+1}^{n+1}$$

QED

15.2 Proof of the Harmonic Mean Geometric Mean Inequality

Let $a_1, a_2, \ldots, a_n > 0$. We will prove the inequality.
We know that $G_n \leq A_n$

$$G_n \leq A_n$$

$$\sqrt[n]{x_1 \cdot x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} \leq \sqrt[n]{x_1 \cdot x_2 \cdots x_n}$$

This concludes the proof of the harmonic mean-geometric mean inequality.
$\mathfrak{QED}$

16 Solving Polynomial Equations

In this section, we discuss the solution formulas for polynomial equations of degrees 2 and 3: the PQ formula, the ABC formula, and the Cubic formula. We also derive each of them step by step.

16.1 The PQ Formula

The PQ formula solves quadratic equations of the form:

$$x^2 + px + q = 0$$

Derivation

To derive the PQ formula, we complete the square:

$$\begin{aligned}
x^2 + px + q &= 0 \\
x^2 + px &= -q \\
x^2 + px + \left(\frac{p}{2}\right)^2 &= -q + \left(\frac{p}{2}\right)^2 \\
\left(x + \frac{p}{2}\right)^2 &= \left(\frac{p}{2}\right)^2 - q \\
x + \frac{p}{2} &= \pm\sqrt{\left(\frac{p}{2}\right)^2 - q} \\
x &= -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}
\end{aligned}$$

PQ formula:

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

16.2 The ABC Formula (Quadratic Formula)

The general quadratic equation is:

$$ax^2 + bx + c = 0 \quad \text{with } a \neq 0$$

Derivation

We normalize the equation by dividing through by a and complete the square:

$$\begin{aligned}
ax^2 + bx + c &= 0 \\
x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\
x^2 + \frac{b}{a}x &= -\frac{c}{a} \\
x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 &= -\frac{c}{a} + \left(\frac{b}{2a}\right)^2 \\
\left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \\
x + \frac{b}{2a} &= \pm\frac{\sqrt{b^2 - 4ac}}{2a} \\
x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\end{aligned}$$

ABC formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

16.3 The Cubic Formula

To solve a general cubic equation:

$$ax^3 + bx^2 + cx + d = 0$$

we first reduce it to a depressed cubic using a substitution.

Step 1: Depress the cubic

Let $x = t - \frac{b}{3a}$, then the equation becomes:

$$t^3 + pt + q = 0$$

with:

$$p = \frac{3ac - b^2}{3a^2}, \quad q = \frac{2b^3 - 9abc + 27a^2 d}{27a^3}$$

Step 2: Solve the depressed cubic using Cardano's method

Assume a solution of the form:

$$t = u + v$$

Then substitute and simplify:

$$(u + v)^3 + p(u + v) + q = 0$$

Expanding and setting:

$$u^3 + v^3 + (3uv + p)(u + v) + q = 0$$

To eliminate the $(u + v)$ term, set:

$$3uv + p = 0 \quad \Rightarrow \quad uv = -\frac{p}{3}$$

Now:

$$u^3 + v^3 = -q$$

Let:

$$u^3 = A, \quad v^3 = B \quad \Rightarrow \quad A + B = -q, \quad AB = -\frac{p^3}{27}$$

These are the roots of the quadratic:

$$z^2 + qz - \frac{p^3}{27} = 0$$

Solve for A and B, then take cube roots to get u and v. The final solution is:

$$x = u + v$$

Cardano's Formula (for depressed cubic)

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}$$

This formula gives one real root. The other roots (if real) can be found using trigonometric or complex methods depending on the discriminant.

17 The Fundamental Theorem of Algebra

(Fundamental Theorem of Algebra). Given any positive integer $n \geq 1$ and any choice of complex numbers $a_0, a_1, \ldots, a_n$, such that $a_n \neq 0$, the polynomial equation

$$a_n z^n + \cdots + a_1 z + a_0 = 0 \tag{1}$$

has at least one solution $z \in \mathbb{C}$.

Gist of the Proof

For readers familiar with Newton's method for solving equations, one starts with a reasonably close approximation to a root, then adjusts the approximation by moving closer in an appropriate direction. We will employ the same strategy here, showing that if one assumes that the argument where the polynomial function achieves its minimum absolute value is not a root, then there is a nearby argument where the polynomial function has an even smaller absolute value, contradicting the assumption that the argument of the minimum absolute value is not a root.

Definitions and Axioms

In the following, $p(z)$ will denote the n-th degree polynomial

$$p(z) = p_0 + p_1 z + p_2 z^2 + \cdots + p_n z^n,$$

where the coefficients p_i are any complex numbers, with neither p_0 nor p_n equal to zero (otherwise the polynomial is equivalent to one of lesser degree).

We will utilize a fundamental completeness property of real and complex numbers, namely that a continuous function on a closed set achieves its minimum at some point in the domain. This can be taken as an axiom, or can be easily proved by applying other well-known completeness axioms, such as the Cauchy sequence axiom or the nested interval axiom.

17.1 Theorem 1

Every polynomial with real or complex coefficients has at least one complex root.

Proof

Suppose that $p(z)$ has no roots in the complex plane. First note that for large z, say $|z| > 2 \max_i |p_i/p_n|$, the z^n term of $p(z)$ is greater in absolute value than the sum of all the other terms. Thus, given some $B > 0$, for any sufficiently large s, we have $|p(z)| > B$ for all z with $|z| \geq s$. We will take $B = 2|p(0)| = 2|p_0|$.

Since $|p(z)|$ is continuous on the interior and boundary of the circle with radius s, it follows by the completeness axiom that $|p(z)|$ achieves its minimum value at some point t in this circle. But since $|p(0)| < \frac{1}{2}|p(z)|$ for all z on the circumference of the circle, it follows that $|p(z)|$ achieves its minimum at some point t in the interior.

Now rewrite the polynomial $p(z)$ by translating the argument z by t, thus producing a new polynomial

$$q(z) = p(z + t) = q_0 + q_1 z + q_2 z^2 + \cdots + q_n z^n,$$

and similarly translate the circle. Presumably the polynomial $q(z)$, defined on some circle centered at the origin, has a minimum absolute value $M > 0$ at $z = 0$. Note that $M = |q(0)| = |q_0|$.

Our proof strategy is to construct some point x, close to the origin, such that $|q(x)| < |q(0)|$, thus contradicting the assumption that $|q(z)|$ has a minimum nonzero value at $z = 0$.

Construction of x such that $|q(x)| < |q(0)|$

Let the first nonzero coefficient of $q(z)$ following q_0 be q_m, so that

$$q(z) = q_0 + q_m z^m + q_{m+1} z^{m+1} + \cdots + q_n z^n.$$

We choose

$$x = r\left(-\frac{q_0}{q_m}\right)^{1/m},$$

where r is a small positive real value, and $\left(-\frac{q_0}{q_m}\right)^{1/m}$ denotes any m-th root of $\left(-\frac{q_0}{q_m}\right)$.

Comment

Unlike the real numbers, in the complex number system the m-th roots of any complex number are guaranteed to exist. If $z = z_1 + iz_2$, then the m-th roots of z are given by

$$\left\{R^{1/m}\cos\left(\frac{\theta + 2k\pi}{m}\right) + iR^{1/m}\sin\left(\frac{\theta + 2k\pi}{m}\right) \middle| k = 0, 1, \ldots, m-1\right\},$$

where $R = \sqrt{z_1^2 + z_2^2}$ and $\theta = \arctan(z_2/z_1)$.

Proof that $|q(x)| < |q(0)|$

With the definition of x, we can write

$$q(x) = q_0 - q_0 r^m + q_{m+1} r^{m+1}\left(-\frac{q_0}{q_m}\right)^{(m+1)/m} + \cdots + q_n r^n\left(-\frac{q_0}{q_m}\right)^{n/m} = q_0 - q_0 r^m + E,$$

where the extra terms E can be bounded as follows. Assume $q_0 \leq q_m$, and define $s = r\left|\frac{q_0}{q_m}\right|^{1/m}$. Then

$$|E| \leq r^{m+1}\max_i |q_i| \left|\frac{q_0}{q_m}\right|^{(m+1)/m}(1 + s + s^2 + \cdots + s^{n-m-1}) \leq \frac{r^{m+1}\max_i |q_i|}{1-s}\left|\frac{q_0}{q_m}\right|^{(m+1)/m}.$$

Thus $|E|$ can be made arbitrarily small compared to $|q_0 r^m| = |q_0| r^m$ by choosing r small enough. For example, select r so that $|E| < \frac{|q_0| r^m}{2}$. Then:

$$|q(x)| = |q_0 - q_0 r^m + E| < |q_0 - \frac{q_0 r^m}{2}| = |q_0|\left(1 - \frac{r^m}{2}\right) < |q_0| = |q(0)|,$$

which contradicts the assumption that $|q(z)|$ has a minimum nonzero value at $z = 0$.

17.2 Theorem 2

Every polynomial of degree n with real or complex coefficients has exactly n complex roots, when counting multiplicities.

Proof

If α is a root of the polynomial $p(z)$ of degree n, then by dividing $p(z)$ by $(z - \alpha)$, we get:

$$p(z) = (z - \alpha)q(z) + r,$$

where $q(z)$ has degree $n - 1$ and r is a constant. But since $p(\alpha) = r = 0$, we conclude:

$$p(z) = (z - \alpha)q(z).$$

Continuing by induction, we conclude that the original polynomial $p(z)$ has exactly n complex roots, counted with multiplicities.

18 The Binomial Coefficient

In this short section we will cover the definition and Properties of the binomial coefficient without going to deep into the combinatoric meaning or Pascal's Triangle.

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

or as Product

$$\prod_{j=1}^{n} \frac{n-k+j}{j}$$

with n and k as natural numbers (for the moment).

18.1 Properties

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

18.2 The binomial Theoreme

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

19 Proportionality and the Rule of Three

19.1 Proportionality

Two quantities are said to be **proportional** if their ratio remains constant.

Direct Proportionality

Two quantities a and b are in *direct proportion* if:

$$\frac{a}{b} = k \quad \Rightarrow \quad a = k \cdot b$$

where k is the constant of proportionality.

Example: If 2 pencils cost \$1, then 4 pencils cost \$2. The ratio is constant: $\frac{2}{1} = \frac{4}{2}$.

Inverse Proportionality

Two quantities a and b are in *inverse proportion* if their product is constant:

$$a \cdot b = k$$

Example: If 4 workers finish a job in 6 hours, then 2 workers would need 12 hours:

$$4 \cdot 6 = 2 \cdot 12 = 24$$

19.2 The Rule of Three (Simple)

The **Rule of Three** is a method to find a fourth value when three values are known and a proportional relationship is assumed.

Direct Rule of Three (Simple)

Given: $a : b = c : x$, solve for x:

$$x = \frac{b \cdot c}{a}$$

Example: If 3 apples cost \$6, how much do 5 apples cost?

$$x = \frac{6 \cdot 5}{3} = 10$$

Inverse Rule of Three (Simple)

If the relationship is inverse:

$$a : b = x : c \quad \Rightarrow \quad x = \frac{a \cdot c}{b}$$

Example: If 5 people finish a task in 8 hours, how long will 10 people need?

$$x = \frac{5 \cdot 8}{10} = 4$$

19.3 The Rule of Three (Compound)

The **Compound Rule of Three** (or *composed rule of three*) involves more than two variables.

Example: If 4 machines produce 120 items in 5 hours, how many items will 6 machines produce in 8 hours?

Step 1: Set up proportionally:

$$\text{Items} \propto \text{Machines} \quad \text{(direct)} \text{Items} \propto \text{Time} \quad \text{(direct)}$$

Step 2: Adjust the quantity:

$$\text{Initial: 4 machines, 5 hrs } \rightarrow 120 \text{ items}$$
$$\text{New: 6 machines, 8 hrs } \rightarrow x \text{ items}$$

Step 3: Use proportionality:

$$x = 120 \cdot \frac{6}{4} \cdot \frac{8}{5} = 120 \cdot 1.5 \cdot 1.6 = 288$$

Answer: 288 items.

Summary Table

Type	**Formula**
Direct Proportion	$x = \frac{b \cdot c}{a}$
Inverse Proportion	$x = \frac{a \cdot c}{b}$
Compound Rule of Three	Multiply by all direct and divide by inverse ratios

20 Factorization Techniques

Factorization is the process of writing a mathematical expression as a product of its factors. This is a fundamental technique in algebra used to simplify expressions, solve equations, and analyze functions.

20.1 Common Factor

Factor out the greatest common divisor (GCD) of all terms.

Example:

$$6x^2 + 9x = 3x(2x + 3)$$

20.2 Difference of Squares

A difference of squares follows the identity:

$$a^2 - b^2 = (a - b)(a + b)$$

Example:

$$x^2 - 16 = (x - 4)(x + 4)$$

20.3 Sum/Difference of Cubes

Sum of Cubes:

$$a^3 + b^3 = (a + b)(a^2 - ab + b^2)$$

Difference of Cubes:

$$a^3 - b^3 = (a - b)(a^2 + ab + b^2)$$

Example:

$$x^3 + 8 = (x + 2)(x^2 - 2x + 4)$$

20.4 Trinomial: General Form $ax^2 + bx + c$

Find two numbers m and n such that:

$$m \cdot n = a \cdot c \quad \text{and} \quad m + n = b$$

Then rewrite and factor by grouping.

Example:

$$6x^2 + 11x + 3 = 6x^2 + 9x + 2x + 3 = 3x(2x + 3) + 1(2x + 3) = (3x + 1)(2x + 3)$$

20.5 Trinomial: Special Case $x^2 + bx + c$

This is the case where $a = 1$. Find two numbers whose product is c and sum is b.

Example:

$$x^2 + 5x + 6 = (x + 2)(x + 3)$$

20.6 Perfect Square Trinomials

These follow the identities:

$$(a + b)^2 = a^2 + 2ab + b^2$$
$$(a - b)^2 = a^2 - 2ab + b^2$$

Example:

$$x^2 + 6x + 9 = (x + 3)^2$$

20.7 Substitution

Substitute a more complex expression with a single variable, factor, then back-substitute.

Example:

$$x^4 + 2x^2 + 1 \Rightarrow \text{Let } y = x^2 \Rightarrow y^2 + 2y + 1 = (y + 1)^2 \Rightarrow (x^2 + 1)^2$$

20.8 Rationalization of Radicals

Rationalizing removes radicals from the denominator.

Example (Single Radical):

$$\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

Example (Binomial):

$$\frac{1}{\sqrt{3}+1} = \frac{\sqrt{3}-1}{(\sqrt{3}+1)(\sqrt{3}-1)} = \frac{\sqrt{3}-1}{2}$$

20.9 Ruffini's Rule (Horner's Method)

Used to divide a polynomial by a binomial of the form $(x - r)$.

Steps:

- Write coefficients of the polynomial.
- Bring down the first coefficient.
- Multiply it by r, add to next coefficient.
- Repeat until the remainder.

Example: Divide $P(x) = x^3 - 6x^2 + 11x - 6$ by $x - 1$:

$$\begin{array}{r|rrrr} 1 & 1 & -6 & 11 & -6 \\ & & 1 & -5 & 6 \\ \hline & 1 & -5 & 6 & 0 \end{array} \Rightarrow Q(x) = x^2 - 5x + 6$$

20.10 Long Division of Polynomials

Use the same algorithm as numerical long division.

Example:

$$\begin{array}{r l} & X^2 + 2X + 2 \\ \cline{2-2} X - 1 \big) & X^3 + X^2 + 0X - 1 \\ & -X^3 + X^2 \\ \cline{2-2} & 2X^2 + 0X \\ & -2X^2 + 2X \\ \cline{2-2} & 2X - 1 \\ & -2X + 2 \\ \cline{2-2} & 1 \end{array}$$

21 Partial Fractions

21.1 The Simplest Case

In the most common partial fraction decomposition, we split:

$$\frac{N(x)}{(x-a_1)(x-a_2)\cdots(x-a_d)}$$

into a sum of the form:

$$\frac{A_1}{x-a_1}+\cdots+\frac{A_d}{x-a_d}$$

We now show that this decomposition can always be achieved, under the assumption that the a_i are all different and $N(x)$ is a polynomial of degree at most $d-1$.

21.1.1 Lemma 1

Let $N(x)$ and $D(x)$ be polynomials of degree n and d, respectively, with $n \leq d$. Suppose that a is not a root of $D(x)$. Then there exists a polynomial $P(x)$ of degree $< d$ and a number A such that:

$$\frac{N(x)}{D(x)(x-a)} = \frac{P(x)}{D(x)} + \frac{A}{x-a}$$

Proof: Let $z = x - a$. Define:

$$\tilde{N}(z) = N(z+a), \quad \tilde{D}(z) = D(z+a)$$

Then:

$$\frac{\tilde{N}(z)}{\tilde{D}(z)z} = \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A}{z} \Rightarrow \frac{\tilde{P}(z)z + A\tilde{D}(z)}{\tilde{D}(z)z}$$

We equate:

$$\tilde{P}(z)z + A\tilde{D}(z) = \tilde{N}(z)$$

Choosing $A = \frac{\tilde{N}(0)}{\tilde{D}(0)}$, the constant terms match. The remainder has no constant term and is divisible by z, so:

$$\tilde{P}(z)z = \tilde{N}(z) - A\tilde{D}(z)$$

Thus, $\tilde{P}(z)$ is a polynomial of degree $< d$.

21.1.2 Recursive Decomposition

Now, consider:

$$\frac{N(x)}{(x-a_1)(x-a_2)\cdots(x-a_d)}$$

Apply Lemma 1 recursively:

$$\frac{N(x)}{(x-a_1)\cdots(x-a_d)} = \frac{A_1}{x-a_1} + \frac{P(x)}{(x-a_2)\cdots(x-a_d)}$$

Then:

$$\frac{P(x)}{(x-a_2)\cdots(x-a_d)} = \frac{A_2}{x-a_2} + \frac{Q(x)}{(x-a_3)\cdots(x-a_d)}$$

Continue until:

$$\frac{N(x)}{(x-a_1)\cdots(x-a_d)} = \frac{A_1}{x-a_1} + \cdots + \frac{A_d}{x-a_d}$$

21.2 Lemma 2

Let $N(x)$ and $D(x)$ be polynomials of degree n and d respectively, with $n < d+m$. Suppose that a is NOT a zero of $D(x)$. Then there is a polynomial $P(x)$ of degree $p < d$ and numbers $A_1, \ldots, A_m$ such that

$$\frac{N(x)}{D(x)(x-a)^m} = \frac{P(x)}{D(x)} + \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_m}{(x-a)^m}$$

Proof: To save writing, let $z = x - a$. Then $\tilde{N}(z) = N(z+a)$ and $\tilde{D}(z) = D(z+a)$ are polynomials of degree n and d respectively, $\tilde{D}(0) = D(a) \neq 0$ and we have to find a polynomial $\tilde{P}(z)$ of degree $p < d$ and numbers $A_1, \ldots, A_m$ such that

$$\begin{aligned}\frac{\tilde{N}(z)}{\tilde{D}(z)z^m} &= \frac{\tilde{P}(z)}{\tilde{D}(z)} + \frac{A_1}{z} + \frac{A_2}{z^2} + \cdots + \frac{A_m}{z^m} \\ &= \frac{\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_m\tilde{D}(z)}{\tilde{D}(z)z^m}\end{aligned}$$

or equivalently, such that

$$\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_{m-1}z\tilde{D}(z) + A_m\tilde{D}(z) = \tilde{N}(z)$$

Now look at the polynomial on the left hand side. Every single term on the left hand side, except for the very last one, $A_m\tilde{D}(z)$, has at least one power of z. So the constant term on the left hand side is exactly the constant term in $A_m\tilde{D}(z)$, which is $A_m\tilde{D}(0)$. The constant term on the right hand side is $\tilde{N}(0)$. So the constant terms on the left and right hand sides are the same if we choose $A_m = \frac{\tilde{N}(0)}{\tilde{D}(0)}$. Recall that $\tilde{D}(0) \neq 0$. Now move $A_m\tilde{D}(z)$ to the right hand side.

$$\tilde{P}(z)z^m + A_1 z^{m-1}\tilde{D}(z) + A_2 z^{m-2}\tilde{D}(z) + \cdots + A_{m-1}z\tilde{D}(z) = \tilde{N}(z) - A_m\tilde{D}(z)$$

The constant terms in $\tilde{N}(z)$ and $A_m\tilde{D}(z)$ are the same, so the right hand side contains no constant term and the right hand side is of the form $\tilde{N}_1(z)z$ with $\tilde{N}_1$ a polynomial of degree at most $d+m-2$. (Recall that $\tilde{N}$ is of degree at most $d+m-1$ and $\tilde{D}$ is of degree at most d.) Divide the whole equation by z.

$$\tilde{P}(z)z^{m-1} + A_1 z^{m-2}\tilde{D}(z) + A_2 z^{m-3}\tilde{D}(z) + \cdots + A_{m-1}\tilde{D}(z) = \tilde{N}_1(z)$$

Now, we can repeat the previous argument. The constant term on the left hand side, which is exactly $A_{m-1}\tilde{D}(0)$ matchs the constant term on the right hand side, which is $\tilde{N}_1(0)$ if we choose $A_{m-1} = \frac{\tilde{N}_1(0)}{\tilde{D}(0)}$. With this choice of A_{m-1}

$$\tilde{P}(z)z^{m-1} + A_1 z^{m-2}\tilde{D}(z) + A_2 z^{m-3}\tilde{D}(z) + \cdots + A_{m-2}z\tilde{D}(z) = \tilde{N}_1(z) - A_{m-1}\tilde{D}(z) = \tilde{N}_2(z)z$$

with $\tilde{N}_2$ a polynomial of degree at most $d+m-3$. Divide by z and continue. After m steps like this, we end up with

$$\tilde{P}(z)z = \tilde{N}_{m-1}(z) - A_1\tilde{D}(z)$$

after having chosen $A_1 = \frac{\tilde{N}_{m-1}(0)}{\tilde{D}(0)}$. There is no constant term on the right side so that $\tilde{N}_{m-1}(z) - A_1\tilde{D}(z)$ is of the form $\tilde{N}_m(z)z$ with $\tilde{N}_m$ a polynomial of degree $d-1$. Choosing $\tilde{P}(z) = \tilde{N}_m(z)$ completes the proof.

Now back to

$$\frac{N(x)}{(x-a_1)^{n_1} \times \cdots \times (x-a_d)^{n_d}}$$

Apply Lemma 2, with $D(x) = (x-a_2)^{n_2} \times \cdots \times (x-a_d)^{n_d}$, $m = n_1$ and $a = a_1$. It says

$$\frac{N(x)}{(x-a_1)^{n_1} \times \cdots \times (x-a_d)^{n_d}} = \frac{P(x)}{(x-a_2)^{n_2} \times \cdots \times (x-a_d)^{n_d}} + \frac{A_{1,1}}{x-a_1} + \frac{A_{1,2}}{(x-a_1)^2} + \cdots + \frac{A_{1,n_1}}{(x-a_1)^{n_1}}$$

Apply Lemma 2 a second time, with $D(x) = (x-a_3)^{n_3} \times \cdots \times (x-a_d)^{n_d}$, $N(x) = P(x)$, $m = n_2$ and $a = a_2$. And so on. Eventually, we end up with

$$\frac{N(x)}{(x-a_1)^{n_1} \times \cdots \times (x-a_d)^{n_d}} = \left[\frac{A_{1,1}}{x-a_1} + \cdots + \frac{A_{1,n_1}}{(x-a_1)^{n_1}}\right] + \cdots + \left[\frac{A_{d,1}}{x-a_d} + \cdots + \frac{A_{d,n_d}}{(x-a_d)^{n_d}}\right]$$

22 Solving Inequalities

Inequalities in mathematics express the relative order of two values. Unlike equations that assert equality, inequalities use symbols such as $<$, $>$, $\leq$, and $\geq$ to indicate that one value is less than, greater than, less than or equal to, or greater than or equal to another. Solving an inequality involves finding the set of all values that satisfy the given relationship.

22.1 Basic Principles

Solving inequalities often involves similar algebraic manipulations as solving equations, with one crucial difference: when multiplying or dividing both sides of an inequality by a negative number, the direction of the inequality sign must be reversed. For example, if $a < b$ and $c < 0$, then $ac > bc$.

22.2 Inequalities Involving Absolute Value

Inequalities involving absolute values require careful consideration due to the piecewise definition of the absolute value function:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

To solve inequalities with absolute values, we typically break down the problem into cases based on the values that make the expressions inside the absolute value equal to zero. These critical points divide the number line into intervals, and we analyze the inequality in each interval.

22.2.1 Example: $|x + 1| \leq 10 + |2x - 6|$

To solve the inequality $|x + 1| \leq 10 + |2x - 6|$, we first identify the critical points where the expressions inside the absolute values are zero:

- $x + 1 = 0 \implies x = -1$
- $2x - 6 = 0 \implies x = 3$

These critical points divide the number line into three intervals: $(-\infty, -1)$, $[-1, 3)$, and $[3, \infty)$. We will analyze the inequality in each interval.

Case 1: $x < -1$ In this interval, $x+1 < 0$, so $|x+1| = -(x+1) = -x-1$. Also, $2x-6 < 2(-1)-6 = -8 < 0$, so $|2x - 6| = -(2x - 6) = -2x + 6$. Substituting these into the inequality, we get:

$$\begin{aligned} -x - 1 &\leq 10 + (-2x + 6) \\ -x - 1 &\leq 16 - 2x \\ 2x - x &\leq 16 + 1 \\ x &\leq 17 \end{aligned}$$

Combining this with the condition $x < -1$, the solution in this interval is $(-\infty, -1)$.

Case 2: $-1 \leq x < 3$ In this interval, $x + 1 \geq 0$, so $|x + 1| = x + 1$. Also, $2x - 6 < 2(3) - 6 = 0$, so $|2x - 6| = -(2x - 6) = -2x + 6$. Substituting these into the inequality, we get:

$$\begin{aligned} x + 1 &\leq 10 + (-2x + 6) \\ x + 1 &\leq 16 - 2x \\ x + 2x &\leq 16 - 1 \\ 3x &\leq 15 \\ x &\leq 5 \end{aligned}$$

Combining this with the condition $-1 \leq x < 3$, the solution in this interval is $[-1, 3)$.

Case 3: $x \geq 3$ In this interval, $x + 1 > 0$, so $|x + 1| = x + 1$. Also, $2x - 6 \geq 2(3) - 6 = 0$, so $|2x - 6| = 2x - 6$. Substituting these into the inequality, we get:

$$\begin{aligned} x + 1 &\leq 10 + (2x - 6) \\ x + 1 &\leq 4 + 2x \\ 1 - 4 &\leq 2x - x \\ -3 &\leq x \end{aligned}$$

Combining this with the condition $x \geq 3$, the solution in this interval is $[3, \infty)$.

Overall Solution To find the complete solution to the inequality, we take the union of the solutions from each case:

$$(-\infty, -1) \cup [-1, 3) \cup [3, \infty) = (-\infty, \infty)$$

Therefore, the inequality $|x + 1| \leq 10 + |2x - 6|$ is true for all real numbers x.

23 Geometry

23.1 Formulas for Area and Perimeter of Geometric Figures

- **Square**
  - Area: $A = a^2$
  - Perimeter: $P = 4a$
- **Rectangle**
  - Area: $A = l \cdot w$
  - Perimeter: $P = 2(l + w)$
- **Circle**
  - Area: $A = \pi r^2$
  - Circumference: $C = 2\pi r$
- **Triangle (General)**
  - Area: $A = \frac{1}{2} b \cdot h$
  - Perimeter: $P = a + b + c$
- **Equilateral Triangle**
  - Area: $A = \frac{\sqrt{3}}{4} a^2$
  - Perimeter: $P = 3a$
- **Isosceles Triangle**
  - Area: $A = \frac{b}{4} \sqrt{4a^2 - b^2}$
- **Trapezoid**
  - Area: $A = \frac{1}{2}(a + b)h$
  - Perimeter: $P = a + b + c + d$
- **Parallelogram**
  - Area: $A = b \cdot h$
  - Perimeter: $P = 2(a + b)$

23.2 Formulas for Area and Volume of 3D Geometric Figures

- **Cube**
  - Surface Area: $A = 6a^2$
  - Volume: $V = a^3$
- **Cylinder**
  - Surface Area: $A = 2\pi r(h + r)$
  - Volume: $V = \pi r^2 h$
- **Cone**
  - Surface Area: $A = \pi r(r + l)$ with l = slant height
  - Volume: $V = \frac{1}{3} \pi r^2 h$
- **Sphere**
  - Surface Area: $A = 4\pi r^2$

- Volume: $V = \frac{4}{3}\pi r^3$

- **Square Pyramid**
  - Surface Area: $A = a^2 + 2a \cdot l$
  - Volume: $V = \frac{1}{3}a^2 h$
- **Triangular Pyramid (Tetrahedron)**
  - Volume: $V = \frac{1}{3}A_b \cdot h$
- **Prism (General)**
  - Surface Area: $A = 2A_b + P_b \cdot h$
  - Volume: $V = A_b \cdot h$

23.3 Thales' Theorem

Thales' Theorem states: If A, B, and C are points on a circle where the line AC is the diameter, then the angle $\angle ABC$ is a right angle.

$$\angle ABC = 90^\circ \quad \text{if } AC \text{ is a diameter of the circle}$$

23.4 Conversion Between Radians and Degrees

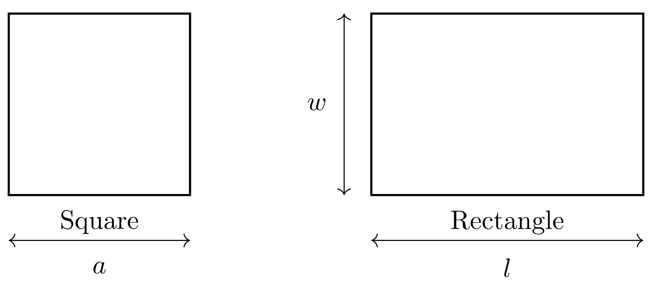
To convert between radians and degrees, use the following formulas:

$$\text{Degrees} = \text{Radians} \times \left(\frac{180^\circ}{\pi}\right)$$
$$\text{Radians} = \text{Degrees} \times \left(\frac{\pi}{180^\circ}\right)$$

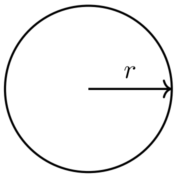
Geometric Figures with Formulas (2D)

Square and Rectangle

$A = a^2 \quad P = 4a$ $\qquad$ $A = l \cdot w \quad P = 2(l + w)$

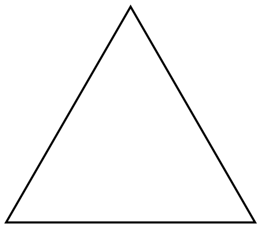


Circle and Triangle Types



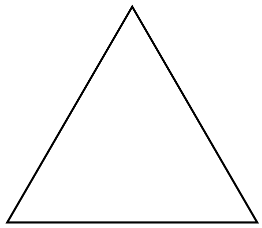
Circle

$A = \pi r^2 \quad C = 2\pi r$



Equilateral

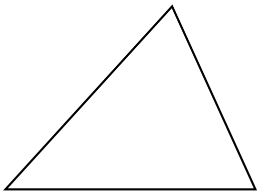
$A = \frac{\sqrt{3}}{4}a^2$



Isosceles

$A = \frac{b}{4}\sqrt{4a^2 - b^2}$

Scalene Triangle, Trapezoid, Parallelogram

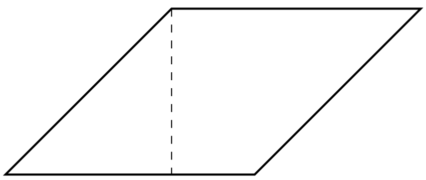


Scalene Triangle

$A = \frac{1}{2}b \cdot h$

Trapezoid

$A = \frac{1}{2}(a+b)h$

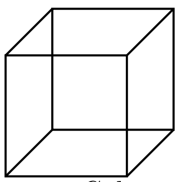


Parallelogram

$A = b \cdot h \quad P = 2(a+b)$

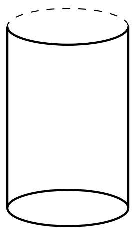
23.5 3D Geometric Figures with Formulas

23.5.1 Cube and Cylinder



Cube

$A = 6a^2 \quad V = a^3$

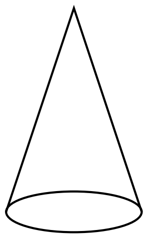


Cylinder

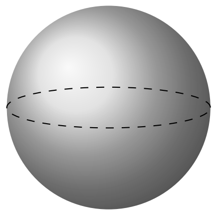
$A = 2\pi r(h+r)$

$V = \pi r^2 h$

23.5.2 Cone and Sphere

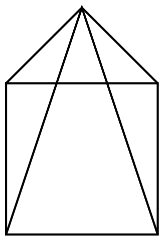


Cone
$A = \pi r(r + l)$
$V = \frac{1}{3}\pi r^2 h$

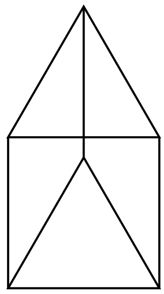


Sphere
$A = 4\pi r^2 \quad V = \frac{4}{3}\pi r^3$

23.5.3 Pyramid and Prism



Square Pyramid
$A = a^2 + 2al$
$V = \frac{1}{3}a^2 h$



Triangular Prism
$A = 2A_b + P_b h$
$V = A_b \cdot h$

23.6 Intercept Theorems

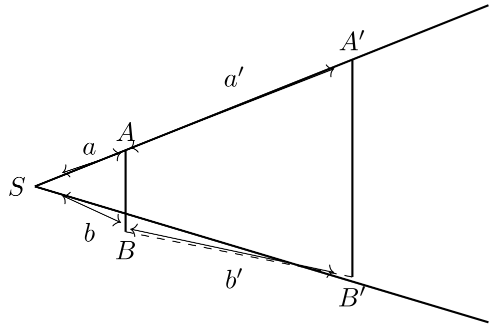
The intercept theorems describe relationships between segment lengths when two rays from a point intersect two parallel lines. They are based on similar triangles and allow us to calculate unknown lengths using proportions.

23.6.1 First Intercept Theorem

If two rays start from a common point and are intersected by two parallel lines, then:

$$\frac{a}{a'} = \frac{b}{b'}$$

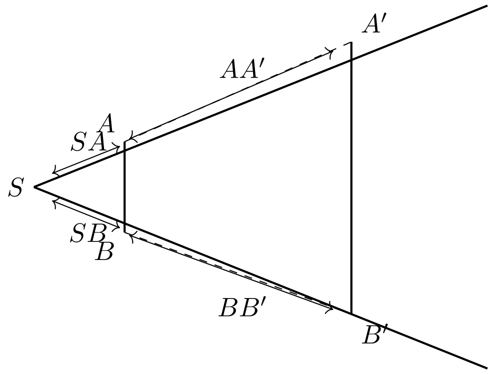
where a and b are segments on one ray, and a', b' are the corresponding segments on the other ray.



23.6.2 Second Intercept Theorem (General Form)

If a ray intersects two parallel lines, the segments from the origin point to the lines are in the same ratio as the segments along the parallels:

$$\frac{SA}{SA'} = \frac{SB}{SB'} \quad \text{and} \quad \frac{AB}{A'B'} = \frac{SA}{SA'}$$



24 Trigonometry

In this section, we will cover a variety of trigonometric functions and their properties.

24.1 Trigonometric Functions

- $\sin(x) = \frac{1}{\csc(x)}$
- $\cos(x) = \frac{1}{\sec(x)}$
- $\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{1}{\cot(x)}$
- $\csc(x) = \frac{1}{\sin(x)}$
- $\sec(x) = \frac{1}{\cos(x)}$
- $\cot(x) = \frac{\cos(x)}{\sin(x)} = \frac{1}{\tan(x)}$

24.2 SOH-CAH-TOA

- $\sin(x) = \frac{\text{opposite}}{\text{hypotenuse}}$
- $\cos(x) = \frac{\text{adjacent}}{\text{hypotenuse}}$
- $\tan(x) = \frac{\text{opposite}}{\text{adjacent}}$

24.3 Pythagorean Identities

- $\sin^2(x) + \cos^2(x) = 1$
- $1 + \tan^2(x) = \sec^2(x)$
- $1 + \cot^2(x) = \csc^2(x)$

24.4 Sum and Difference Formulas

- $\sin(a + b) = \sin(a)\cos(b) + \cos(a)\sin(b)$
- $\sin(a - b) = \sin(a)\cos(b) - \cos(a)\sin(b)$
- $\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b)$
- $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$
- $\tan(a + b) = \frac{\tan(a)+\tan(b)}{1-\tan(a)\tan(b)}$
- $\tan(a - b) = \frac{\tan(a)-\tan(b)}{1+\tan(a)\tan(b)}$

24.5 Double Angle Formulas

- $\sin(2x) = 2\sin(x)\cos(x)$
- $\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$
- $\tan(2x) = \frac{2\tan(x)}{1-\tan^2(x)}$
- $\csc(2x) = \frac{2\csc(x)\sec(x)}{1-\tan^2(x)}$
- $\sec(2x) = \frac{1+\tan^2(x)}{2\tan(x)}$
- $\cot(2x) = \frac{1-\tan^2(x)}{2\tan(x)}$

24.6 Half Angle Formulas

- $\sin\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1-\cos(x)}{2}}$
- $\cos\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1+\cos(x)}{2}}$
- $\tan\left(\frac{x}{2}\right) = \frac{\sin(x)}{1+\cos(x)} = \frac{1-\cos(x)}{\sin(x)} = \frac{\tan(x)}{1+\tan^2(x)}$
- $\csc\left(\frac{x}{2}\right) = \frac{1}{\sin\left(\frac{x}{2}\right)} = \pm\sqrt{\frac{2}{1-\cos(x)}}$
- $\sec\left(\frac{x}{2}\right) = \frac{1}{\cos\left(\frac{x}{2}\right)} = \pm\sqrt{\frac{2}{1+\cos(x)}}$
- $\cot\left(\frac{x}{2}\right) = \frac{1}{\tan\left(\frac{x}{2}\right)} = \frac{1+\cos(x)}{\sin(x)} = \frac{1-\tan^2(x)}{2\tan(x)}$

24.7 Product to Sum Formulas

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

24.8 Power Reducing Formulas

- $\sin^2(x) = \frac{1-\cos(2x)}{2}$
- $\cos^2(x) = \frac{1+\cos(2x)}{2}$
- $\tan^2(x) = \frac{1-\cos(2x)}{1+\cos(2x)}$
- $\csc^2(x) = \frac{1}{\sin^2(x)} = \frac{2}{1-\cos(2x)}$
- $\sec^2(x) = \frac{1}{\cos^2(x)} = \frac{2}{1+\cos(2x)}$
- $\cot^2(x) = \frac{1}{\tan^2(x)} = \frac{1+\cos(2x)}{1-\cos(2x)}$

24.9 Even and Odd Functions

- $\sin(-x) = -\sin(x)$
- $\cos(-x) = \cos(x)$
- $\tan(-x) = -\tan(x)$
- $\csc(-x) = -\csc(x)$
- $\sec(-x) = \sec(x)$
- $\cot(-x) = -\cot(x)$

24.10 Graph Identity

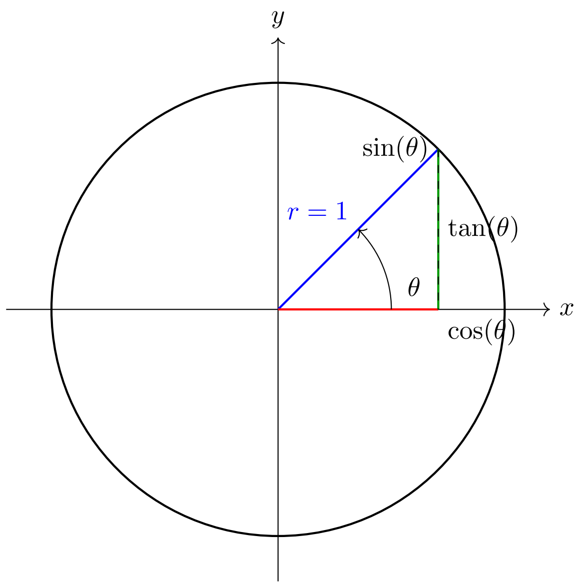
$\cos(x) = \sin(x + \frac{\pi}{2})$

24.11 Trigonometric Values

θ (radians)	$\sin(\theta)$	$\cos(\theta)$	$\tan(\theta)$
0	0	1	0
$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1
$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$
$\frac{\pi}{2}$	1	0	undefined
$\frac{2\pi}{3}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$-\sqrt{3}$
$\frac{3\pi}{4}$	$\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	-1
$\frac{5\pi}{6}$	$\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{3}}$
π	0	-1	0
$\frac{7\pi}{6}$	$-\frac{1}{2}$	$-\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{3}}$
$\frac{5\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$	1
$\frac{4\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$-\frac{1}{2}$	$\sqrt{3}$
$\frac{3\pi}{2}$	-1	0	undefined
$\frac{5\pi}{3}$	$-\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$-\sqrt{3}$
$\frac{7\pi}{4}$	$-\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	-1
$\frac{11\pi}{6}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{3}}$
2π	0	1	0

Unit Circle and Trigonometric Functions

The unit circle is a circle with a radius of 1 centered at the origin. Trigonometric functions like $\sin(\theta)$, $\cos(\theta)$, and $\tan(\theta)$ can be defined based on the coordinates of a point on the circle corresponding to angle θ.



Definitions on the Unit Circle:

– $\cos(\theta)$ is the **x-coordinate** of the point on the circle.

– $\sin(\theta)$ is the **y-coordinate**.

- $\tan(\theta) = \dfrac{\sin(\theta)}{\cos(\theta)}$ is the ratio of the opposite side to the adjacent side in the triangle.

25 Linear Systems of Equations

In this section, we will discuss the solution of linear systems of equations. A linear system of equations is a set of equations that can be expressed in the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where a_{ij} are the coefficients of the variables x_j, and b_i are the constants on the right-hand side of the equations. The goal is to find the values of $x_1, x_2, \ldots, x_n$ that satisfy all equations simultaneously.

25.1 Matrix Representation

A linear system can be represented in matrix form as:

$$A\mathbf{x} = \mathbf{b}$$

where A is the coefficient matrix, $\mathbf{x}$ is the vector of variables, and $\mathbf{b}$ is the vector of constants. The coefficient matrix A is an $m \times n$ matrix, where m is the number of equations and n is the number of variables. The vector $\mathbf{x}$ is an $n \times 1$ column vector, and the vector $\mathbf{b}$ is an $m \times 1$ column vector. The system can be solved using various methods, including:

- Gaussian elimination
- LU decomposition
- Matrix inversion (if A is square and invertible)
- Iterative methods (e.g., Jacobi, Gauss-Seidel)
- Special methods for sparse matrices
- Special methods for structured matrices (e.g., banded, Toeplitz)
- Special methods for large-scale problems (e.g., conjugate gradient, GMRES)

25.2 Gaussian Elimination

Gaussian elimination is a method for solving linear systems by transforming the system into an upper triangular form. The steps involved in Gaussian elimination are:

1. Forward elimination: Transform the system into an upper triangular form by eliminating the variables from the equations.
2. Back substitution: Solve for the variables starting from the last equation and substituting back into the previous equations.

The forward elimination process involves performing row operations on the augmented matrix $[A|\mathbf{b}]$ to create zeros below the diagonal. The row operations include:

- Swapping two rows
- Multiplying a row by a non-zero scalar
- Adding or subtracting a multiple of one row from another row

Once the matrix is in upper triangular form, back substitution is used to find the values of the variables. The last equation gives the value of the last variable, which can then be substituted into the previous equations to find the other variables.

25.3 Gauss-Jordan Elimination

Gauss-Jordan elimination is an extension of Gaussian elimination that transforms the matrix into reduced row echelon form (RREF). In RREF, each leading entry in a row is 1, and all entries above and below the leading entry are zeros. The steps involved in Gauss-Jordan elimination are:

1. Forward elimination: Transform the system into an upper triangular form.
2. Back substitution: Transform the upper triangular matrix into RREF by eliminating the entries above the leading 1s.
3. Solve for the variables directly from the RREF matrix.

The Gauss-Jordan elimination method is particularly useful for finding the inverse of a matrix, as it can be applied to the augmented matrix $[A|I]$, where I is the identity matrix. If the left side of the augmented matrix becomes I, then the right side will be the inverse of A. **Example:** Solve the following system of equations using Gaussian elimination:

$$\begin{aligned} 2x + 3y + z &= 1 \\ 4x + y - z &= 2 \\ -2x + y + 3z &= 3 \end{aligned}$$

Solution: The augmented matrix for the system is:

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

Performing row operations to eliminate the variables, we can transform the matrix into upper triangular form:

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

Now, we can perform back substitution to find the values of x, y, and z:

$$\begin{aligned} z &= 1 \\ -5y - 3z &= 0 \implies y = -\frac{3}{5} \\ 2x + 3y + z &= 1 \implies x = \frac{1}{5} \end{aligned}$$

Thus, the solution to the system is:

$$\begin{aligned} x &= \frac{1}{5} \\ y &= -\frac{3}{5} \\ z &= 1 \end{aligned}$$

25.4 Homogeneous Linear Equations

A linear equation is said to be **homogeneous** if its constant term is zero. That is, it can be written in the form:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

Such equations always have at least the trivial solution $x_1 = x_2 = \cdots = x_n = 0$.

25.5 Particular Solution

A particular solution to a linear system of equations is a specific solution that satisfies the system. It can be found using various methods, including substitution, elimination, or matrix methods. A particular solution is not unique; there may be multiple particular solutions depending on the system. A particular solution can be found by substituting specific values for the variables and solving for the remaining variables. For example, in the system

$$\begin{aligned} 2x + 3y &= 5 \\ 4x - y &= 1 \end{aligned}$$

we can substitute $x = 1$ into the first equation to find y:

$$\begin{aligned} 2(1) + 3y &= 5 \\ 3y &= 3 \\ y &= 1 \end{aligned}$$

Thus, $(x, y) = (1, 1)$ is a particular solution to the system. However, this is not the only solution; other values of x may yield different values of y.

25.6 General = Particular + Homogeneous

The general solution of a linear system of equations is the complete set of solutions that satisfy the system. It can be expressed as the sum of a particular solution and the general solution of the associated homogeneous system. The general solution can be written as:

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$$

where $\mathbf{x}_p$ is a particular solution to the non-homogeneous system, and $\mathbf{x}_h$ is the general solution to the homogeneous system. The homogeneous system is obtained by setting the right-hand side of the equations to zero:

$$A\mathbf{x} = \mathbf{0}$$

The theoreme says:

Any linear system's solution set has the form:

$$\left\{ \vec{p} + c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \;\middle|\; c_1, \ldots, c_k \in \mathbb{R} \right\}$$

where $\vec{p}$ is a particular solution to the system, and the vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ form a basis of the solution space to the corresponding homogeneous system. The number k equals the number of **free variables** the system has after applying Gaussian elimination.

25.7 Linear Combination Lemma

Any linear combination of linear combinations is a linear combination.

25.8 Example: Gaussian Elimination with 3 Equations and 4 Unknowns

Consider the following system of linear equations:

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 4 \\ 2x_1 + 5x_2 + x_3 + 3x_4 &= 10 \\ x_1 + 3x_2 + 2x_3 + 2x_4 &= 7 \end{aligned}$$

Step 1: Augmented Matrix

$$\begin{bmatrix} 1 & 2 & 1 & 1 & 4 \\ 2 & 5 & 1 & 3 & 10 \\ 1 & 3 & 2 & 2 & 7 \end{bmatrix}$$

Step 2: Eliminate below pivot in column 1

- Row 2 = Row 2 - 2 × Row 1
- Row 3 = Row 3 - Row 1

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

Step 3: Eliminate below pivot in column 2

- Row 3 = Row 3 - Row 2

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

Step 4: Back Substitution

From Row 3:

$$2x_3 = 1 \Rightarrow x_3 = \frac{1}{2}$$

From Row 2:

$$x_2 - x_3 + x_4 = 2 \Rightarrow x_2 = 2 + x_3 - x_4 = 2 + \frac{1}{2} - x_4 = \frac{5}{2} - x_4$$

From Row 1:

$$x_1 + 2x_2 + x_3 + x_4 = 4 \Rightarrow x_1 = 4 - 2x_2 - x_3 - x_4$$

Substitute:

$$x_1 = 4 - 2\left(\frac{5}{2} - x_4\right) - \frac{1}{2} - x_4 = 4 - 5 + 2x_4 - \frac{1}{2} - x_4 = -1 - \frac{1}{2} + x_4 = -\frac{3}{2} + x_4$$

General Solution

Let $x_4 = t$ (free variable), then:

$$\begin{aligned} x_1 &= -\frac{3}{2} + t \\ x_2 &= \frac{5}{2} - t \\ x_3 &= \frac{1}{2} \\ x_4 &= t \end{aligned} \qquad \text{with } t \in \mathbb{R}$$

Solution Set:

$$\left\{ \begin{pmatrix} -\frac{3}{2} \\ \frac{5}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; t \in \mathbb{R} \right\}$$

25.9 The Determinant, the cross product and the solutions of linear Systems of Equations

A linear system of three equation has the following properties:

- There is a unique solution if the determinant of the coefficient matrix is non-zero.

$$\langle (axb), c \rangle = \det(a, b, c) \neq 0$$

- There are infinitely many solutions if the determinant of the coefficient matrix is zero.

$$\langle (axb), c \rangle = \det(a, b, c) = 0$$

- There is no solution if the determinant of the coefficient matrix is zero and the system is inconsistent.

$$\langle (axb), c \rangle = 0$$

26 Analytical Geometry

In this section, we will cover the topics for the geometry of $\mathbb{R}^2$ and $\mathbb{R}^3$. An maybe also in higher dimensions.

26.1 Vectors and Points

In analytical geometry, points and vectors are the basic elements.

A point in $\mathbb{R}^3$ is represented as $\vec{P} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Or as $(x_1, x_2, \cdots, x_n)$.

A vector is an object with direction and magnitude (in this case), also represented as $\vec{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$.

Or as $(x_1, x_2, \cdots, x_n)^T$.

26.2 Vector Addition and Scalar Multiplication

Given two vectors $\vec{a}$ and $\vec{b}$:

$$\vec{a} \pm \vec{b} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} \pm \begin{pmatrix} b_x \\ b_y \end{pmatrix} = \begin{pmatrix} a_x \pm b_x \\ a_y \pm b_y \end{pmatrix}$$

For a scalar λ and vector $\vec{a}$:

$$\lambda\vec{a} = \lambda \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \begin{pmatrix} \lambda a_x \\ \lambda a_y \end{pmatrix}$$

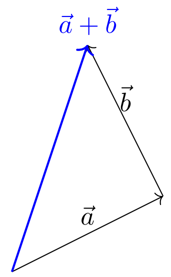


Figure 1: Vector addition: $\vec{a} + \vec{b}$

26.3 Equation of a Line

A line is defined by a point $\vec{P}$ and a direction vector $\vec{v}$:

$$\vec{r}(t) = \vec{P} + t\vec{v}, \quad t \in \mathbb{R}$$

26.4 Equation of a Plane

A plane is defined by a point $\vec{P}$ and a normal vector $\vec{n}$:

$$P : \vec{x} = \vec{s} + \lambda_1\vec{v_1} + \lambda_2\vec{v_2}$$

26.5 Scalar (Dot) Product

The scalar product of vectors $\vec{a}$ and $\vec{b}$ is a function $\langle x, y\rangle := V \times V \to \mathbb{K}$ with the following properties:

$\forall \vec{a}, \vec{b} \in V$:

(i) $\langle \vec{a}, \vec{b}\rangle = \langle \vec{b}, \vec{a}\rangle$

(ii) $\langle \vec{a}, \vec{b} + \vec{c}\rangle = \langle \vec{a}, \vec{b}\rangle + \langle \vec{a}, \vec{c}\rangle$

(iii) $\langle \vec{a}, \lambda\vec{b}\rangle = \lambda\langle \vec{a}, \vec{b}\rangle$

(iv) $\langle \vec{a}, \vec{b} \rangle = 0 \Leftrightarrow \vec{a} = \vec{0} \vee \vec{b} = \vec{0}$

Formula:

$$\langle \vec{a}, \vec{b} \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

26.6 Length (Norm) of a Vector

The norm of a vector $\vec{a}$ is described by different norms:
It also has the following properties:

(i) $\|\vec{a}\| \geq 0$ and $\|\vec{a}\| = 0 \Leftrightarrow \vec{a} = \vec{0}$

(ii) $\|\lambda \vec{a}\| = |\lambda| \|\vec{a}\|$

(iii) $\|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\|$ (Triangle inequality)

(iv) $\|\vec{a} + \vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\langle \vec{a}, \vec{b} \rangle$ (Pythagorean theorem)

This are the most common norms, although we will be using primarly the euclidean norm:

1. **Euclidean norm**:
$$\|\vec{a}\| = \sqrt{\langle \vec{a}, \vec{a} \rangle} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

2. **Manhattan norm**:
$$\|\vec{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n|$$

3. **Maximum norm**:
$$\|\vec{a}\|_\infty = \max(|a_1|, |a_2|, \cdots, |a_n|)$$

26.7 Angle Relations

The angle θ between two vectors is:

$$\cos\theta = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}, \quad \theta = \arccos\left(\frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}\right)$$

Line-Line Angle

Use direction vectors $\vec{v}_1$ and $\vec{v}_2$:

$$\theta = \arccos\left(\frac{\langle \vec{v}_1, \vec{v}_2 \rangle}{\|\vec{v}_1\| \|\vec{v}_2\|}\right)$$

Line-Plane Angle

Let $\vec{v}$ be the line's direction and $\vec{n}$ the plane's normal:

$$\theta = \arcsin\left(\frac{\langle \vec{v}, \vec{n} \rangle}{\|\vec{v}\| \|\vec{n}\|}\right)$$

Plane-Plane Angle

Angle between planes is angle between their normals:

$$\theta = \arccos\left(\frac{\langle \vec{n}_1, \vec{n}_2 \rangle}{\|\vec{n}_1\| \|\vec{n}_2\|}\right)$$

26.8 Line Relations

Two lines can be:

- **Identical**: same direction vector and point
- **Parallel**: direction vectors are proportional
- **Intersecting**: one solution for t_1, t_2 such that $\vec{r}_1(t_1) = \vec{r}_2(t_2)$
- **Skew**: not parallel, do not intersect

To find the relation:

1. Check if direction vectors are scalar multiples $\Rightarrow$ parallel
2. Solve $\vec{P}_1 + t\vec{v}_1 = \vec{P}_2 + s\vec{v}_2$ for t and s:
   - Solution exists $\Rightarrow$ intersect
   - No solution $\Rightarrow$ skew
3. If same point and direction vector $\Rightarrow$ identical

26.9 Normalization of a vector

To normalize a vector $\vec{a}$, we divide it by its length:

$$\hat{\vec{a}} = \frac{\vec{a}}{||\vec{a}||} = \frac{\begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}}{\sqrt{a_x^2 + a_y^2 + a_z^2}} = \begin{pmatrix} \frac{a_x}{||\vec{a}||} \\ \frac{a_y}{||\vec{a}||} \\ \frac{a_z}{||\vec{a}||} \end{pmatrix}$$

26.10 Orthogonal Vectors and the Orthogonal Projection

Two vectors $\vec{a}$ and $\vec{b}$ are orthogonal if:

$$\langle \vec{a}, \vec{b} \rangle = 0$$

The orthogonal projection of vector $\vec{a}$ onto vector $\vec{b}$ is given by:

$$\mathrm{p}_{\vec{b}}(\vec{a}) = \frac{\langle \vec{a}, \vec{b} \rangle}{||\vec{b}||^2} \cdot \vec{b}$$

Proof of the projection formula

$$p||q \implies \alpha q = p \ \forall \alpha \in \mathbb{K} \text{ and } \forall \ p, q \in V$$

$$\langle a - p, q \rangle = 0 \implies \langle a - \alpha q, q \rangle = 0 \implies \langle a, q \rangle - \alpha \langle q, q \rangle = 0$$

$$\alpha = \frac{\langle a, q \rangle}{\langle q, q \rangle}$$

Therefore the orthogonal projection of vector $\vec{a}$ on $\vec{b}$ is given by multipliying $\vec{b}$ by the scalar α.
$\mathfrak{QED}$

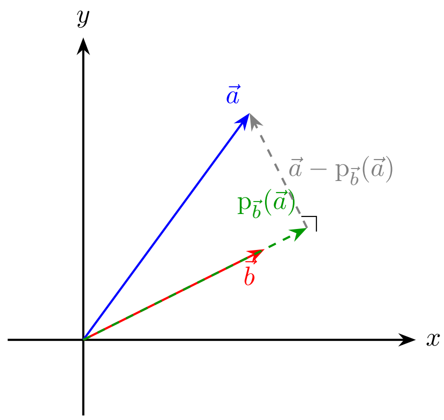


Figure 2: Orthogonal projection of $\vec{a}$ onto $\vec{b}$

26.11 The Cross Product

The cross product of two vectors $\vec{a}$ and $\vec{b}$ in $\mathbb{R}^3$ is defined as:

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

The cross product has the following properties:

(i) $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$

(ii) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$

(iii) $(\lambda_1 \vec{a}) \times (\lambda_2 \vec{b}) = \lambda_1 \lambda_2 (\vec{a} \times \vec{b})$

(iv) $\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \|\vec{b}\| \sin(\theta)$

(v) $\langle \vec{a}, (\vec{b} \times \vec{c}) \rangle = 0$ (scalar triple product)

(vi) $\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a} = \lambda \vec{b}$ for some $\lambda \in \mathbb{R}$ (parallel vectors)

(vii) $\vec{a} \times \vec{b} = \vec{0} \Leftrightarrow \vec{a} = \vec{0} \vee \vec{b} = \vec{0}$ (zero vector)

The cross product is not defined in $\mathbb{R}^2$. The cross product is not commutative, but it is associative:

$$(\vec{a} \times \vec{b}) \times \vec{c} = \vec{a} \times (\vec{b} \times \vec{c})$$

The cross product is distributive over vector addition:

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$$

The cross product is anti-commutative:

$$\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$$

The cross product is not associative:

$$(\vec{a} \times \vec{b}) \times \vec{c} \neq \vec{a} \times (\vec{b} \times \vec{c})$$

The cross product is not distributive over scalar multiplication:

$$\lambda(\vec{a} \times \vec{b}) \neq (\lambda \vec{a}) \times \vec{b}$$

The lenght of the cross product in $\mathbb{R}^3$ is the area of the parallelogram spanned by the two vectors:

26.12 Orthogonal vectors in $\mathbb{R}^2$ $\mathbb{R}^3$

26.12.1 Orthogonal vectors in $\mathbb{R}^2$

(i) Interchange the components

(ii) Change the sign of one component

26.12.2 Orthogonal vectors in $\mathbb{R}^3$

(i) Interchange two components

(ii) the one that was not changed, set to zero

(iii) Change the sign of first component

Example:

$$\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

26.13 Hessian Normal Form

The Hessian normal form is a way of expressing the equation of a plane in three-dimensional space using a normalized normal vector. It is particularly useful in computational geometry and physics, where signed distances from points to planes are important.

Geometric Interpretation

The Hessian normal form represents a plane by specifying:

- a unit normal vector $\vec{n} = (a, b, c)$ to the plane,
- and the shortest distance d from the origin to the plane.

This form is derived by normalizing the general plane equation. A plane in 3D can be written as:

$$ax + by + cz + d = 0,$$

where (a, b, c) is a normal vector to the plane and d is the dot product of the normal vector with a poin p . If we divide all terms by $\sqrt{a^2 + b^2 + c^2}$, we normalize the normal vector:

$$\frac{a}{\sqrt{a^2+b^2+c^2}}x + \frac{b}{\sqrt{a^2+b^2+c^2}}y + \frac{c}{\sqrt{a^2+b^2+c^2}}z + \frac{d}{\sqrt{a^2+b^2+c^2}} = 0.$$

Let:

$$\vec{n} = \left(\frac{a}{\sqrt{a^2+b^2+c^2}}, \frac{b}{\sqrt{a^2+b^2+c^2}}, \frac{c}{\sqrt{a^2+b^2+c^2}}\right), \quad d' = \frac{d}{\sqrt{a^2+b^2+c^2}},$$

then the equation becomes:

$$\vec{n} \cdot \vec{r} + d' = 0,$$

which is the Hessian normal form. Here, $\vec{r} = (x, y, z)$ is any point on the plane, and $\vec{n}$ is the unit normal.

Signed Distance to the Plane

This form allows easy calculation of the signed distance from any point $\vec{p}$ to the plane:

$$\text{distance} = \vec{n} \cdot \vec{p} + d',$$

which is positive if $\vec{p}$ lies on the same side of the plane as the normal vector.

Derivation Illustration

To visualize the derivation, imagine a plane with unit normal vector $\vec{n}$, and a point P in space. The shortest distance from P to the plane is the projection of the vector $\vec{p} - \vec{q}$ onto $\vec{n}$, where $\vec{q}$ is any point on the plane. This leads to:

$$\text{distance} = (\vec{p} - \vec{q}) \cdot \vec{n}.$$

This gives the signed distance formula and thus motivates the Hessian form.

26.14 Converting from the parametric form to the Hessian normal form

Steps:

(i) Find the normal vector $\vec{n}$ of the plane

(ii) Normalize the normal vector

(iii) Find the distance d from the origin to the plane

(iv) Write the Hessian normal form

26.15 Converting from the Hessian normal form to the parametric form

Steps:

(i) Find a point on the plane

(ii) Find two direction vectors in the plane

(iii) Write the parametric form

26.16 Properties of lines and planes

- Two planes are parallel if their normal vectors are scalar multiples of each other.
- A line and a plane are parallel if the direction vector of the line is orthogonal to the normal vector of the plane.
- A line intersects a plane if there exists a point on the line that satisfies the equation of the plane.
- Two planes intersect in a line if their normal vectors are not parallel.
- Three planes can intersect in a point, a line, or not at all.
- If we have a line G and a poin on the line, for every vector $\vec{n}$ that is orthogonal to the direction vector of the line: $x \in G \iff \langle x, n \rangle$
- If p and q are two points in the line G with a normal vector then $\langle p, n \rangle = \langle q, n \rangle$
- Let E be a plane with the origin p and the direction vector $\vec{v}$ and $\vec{w}$, then there exist a nomarl vector and $x \in E \iff \langle x, n \rangle = \langle p, n \rangle$

26.17 Convert Normal Vector in Two Direction Vectors

Steps:

(i) Given the normal vector $\vec{n} = (a, b, c)$, interchange a and b and multiply b by -1

(ii) Set the other component to 0. This gives you the first direction vector $\vec{v} = (-b, a, 0)$

(iii) Take the original normal vector $\vec{n}$ and interchange a and c and multiply c by -1

(iv) Set the other component to 0. This gives you the second direction vector $\vec{w} = (-c, 0, a)$

26.18 Intersection between Line and Plane

To find the intersection between a line and a plane, we can use the following steps:

(i) Write the parametric form of the line: $\vec{r}(t) = \vec{P} + t\vec{v}$, where $\vec{P}$ is a point on the line and $\vec{v}$ is the direction vector.

(ii) Write the equation of the plane in Hessian normal form: $\langle \vec{n}, \vec{x} - \vec{P} \rangle = 0$, where $\vec{n}$ is the normal vector and $\vec{P}$ is a point on the plane.

(iii) Substitute the parametric form of the line into the equation of the plane.

(iv) Solve for t to find the intersection point.

(v) Substitute t back into the parametric form of the line to find the intersection point.

Example: Given the line:

$$g(t) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and the plane:

$$E(u, m) = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + u \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + m \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

we want to find the intersection point.

Step 1: Determine the normal vector of the plane using the cross product of the two direction vectors:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\vec{n} = \vec{v}_1 \times \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Step 2: Use the normal vector and a point on the plane to write the plane equation:

$$\langle \vec{n}, \vec{x} - \vec{Q} \rangle = 0, \quad \vec{Q} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Step 3: Plug the line into the plane equation:

$$\vec{x}(t) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \Rightarrow \vec{x}(t) - \vec{Q} = \begin{pmatrix} t \\ t-1 \\ t-1 \end{pmatrix}$$

Now compute the dot product:

$$\langle \vec{n}, \vec{x}(t) - \vec{Q} \rangle = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} t \\ t-1 \\ t-1 \end{pmatrix} = -t + (t-1) + (t-1) = t - 2$$

Step 4: Solve for t:

$$t - 2 = 0 \Rightarrow t = 2$$

Step 5: Substitute $t = 2$ into the line:

$$\vec{x}(2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$$

Result: The line intersects the plane at the point

$$\boxed{\begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}}$$

Proof:

$$E : \langle x, n \rangle = \langle p, n \rangle$$
$$G : x = p + t \cdot v$$
$$\langle x, n \rangle = c$$
$$\langle p + t \cdot v, n \rangle = c$$
$$\langle p, n \rangle + t \cdot \langle v, n \rangle = c$$
$$t = \frac{c - \langle p, n \rangle}{\langle v, n \rangle}$$

$\mathfrak{QED}$

26.19 Distances between points, lines and planes

Distance between two points

The distance between two points $\vec{P_1}$ and $\vec{P_2}$ in $\mathbb{R}^n$ is given by:

$$d(\vec{P_1}, \vec{P_2}) = \|\vec{P_1} - \vec{P_2}\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + \cdots + (z_1 - z_2)^2}$$

where $\vec{P_1} = (x_1, y_1, \cdots, z_1)$ and $\vec{P_2} = (x_2, y_2, \cdots, z_2)$.

Distance between a point and a hyperplane

The distance between a point $\vec{P}$ and a hyperplane defined by the equation $\langle \vec{n}, \vec{x} - \vec{P_0} \rangle = 0$ is given by:

$$d(\vec{P}, \text{hyperplane}) = \frac{|\langle \vec{n}, \vec{P} - \vec{P_0} \rangle|}{\|\vec{n}\|}$$

where $\vec{P_0}$ is a point on the hyperplane and $\vec{n}$ is the normal vector of the hyperplane.

Distance between two lines

The distance between two lines in $\mathbb{R}^3$ can be calculated using the formula:

$$d = \frac{|\langle \vec{v_1} \times \vec{v_2}, \vec{P_2} - \vec{P_1} \rangle|}{\|\vec{v_1} \times \vec{v_2}\|}$$

where $\vec{P_1}$ and $\vec{P_2}$ are points on the two lines, and $\vec{v_1}$ and $\vec{v_2}$ are the direction vectors of the lines.

Distance between a point and a line

The distance between a point $\vec{P}$ and a line defined by the parametric equation $\vec{r}(t) = \vec{P_0} + t\vec{v}$ is given by:

$$d(\vec{P}, \text{line}) = \frac{\|\vec{v} \times (\vec{P} - \vec{P_0})\|}{\|\vec{v}\|}$$

where $\vec{P_0}$ is a point on the line and $\vec{v}$ is the direction vector of the line.

Distance between two planes

The distance between two parallel planes defined by the equations $\langle \vec{n}, \vec{x} - \vec{P_1} \rangle = 0$ and $\langle \vec{n}, \vec{x} - \vec{P_2} \rangle = 0$ is given by:

$$d = \frac{|\langle \vec{n}, \vec{P_2} - \vec{P_1} \rangle|}{\|\vec{n}\|}$$

where $\vec{P_1}$ and $\vec{P_2}$ are points on the two planes, and $\vec{n}$ is the normal vector of the planes.

Distance between a point and a plane

The distance between a point $\vec{P}$ and a plane defined by the equation $\langle \vec{n}, \vec{x} - \vec{P_0} \rangle = 0$ is given by:

$$d(\vec{P}, \text{plane}) = \frac{|\langle \vec{n}, \vec{P} - \vec{P_0} \rangle|}{\|\vec{n}\|}$$

where $\vec{P_0}$ is a point on the plane and $\vec{n}$ is the normal vector of the plane.

26.19.1 Example: Distance Between Two Skew Lines

To find the shortest distance between two skew lines, we use the formula:

$$\text{distance} = \frac{|\langle (\vec{P_2} - \vec{P_1}), (\vec{v}_1 \times \vec{v}_2) \rangle|}{\|\vec{v}_1 \times \vec{v}_2\|}$$

Where:

- $\vec{P_1}$ and $\vec{P_2}$ are points on each line,
- $\vec{v}_1$ and $\vec{v}_2$ are the direction vectors,
- $\vec{v}_1 \times \vec{v}_2$ is the cross product of the direction vectors.

Given:

$$g_1 : \vec{r}_1(a) = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + a \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad g_2 : \vec{r}_2(b) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + b \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Step 1: Set

$$\vec{P_1} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \quad \vec{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\vec{P_2} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Step 2: Compute the vector between base points:

$$\vec{P_2} - \vec{P_1} = \begin{pmatrix} 1-2 \\ 2-2 \\ 3-2 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Step 3: Compute the cross product:

$$\vec{v}_1 \times \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} (1)(1) - (1)(2) \\ (1)(3) - (0)(1) \\ (0)(2) - (1)(3) \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix}$$

Step 4: Compute scalar triple product:

$$\langle (\vec{P_2} - \vec{P_1}), (\vec{v}_1 \times \vec{v}_2) \rangle = \langle \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -3 \end{pmatrix} \rangle = (-1)(-1) + (0)(3) + (1)(-3) = 1 + 0 - 3 = -2 \Rightarrow |\dots| = 2$$

Step 5: Magnitude of the cross product:

$$\|\vec{v}_1 \vec{v}_2\| = \sqrt{(-1)^2 + 3^2 + (-3)^2} = \sqrt{1 + 9 + 9} = \sqrt{19}$$

Final Answer:

$$\text{distance} = \frac{2}{\sqrt{19}} \approx 0.458$$

$$\boxed{\text{Shortest distance between the lines is } \frac{2}{\sqrt{19}}}$$

26.20 Foot of the Perpendicular and Mirror Point

26.20.1 Foot of the Perpendicular

The **foot of the perpendicular** from a point $\vec{P}$ to a line (or plane) is the point on the line (or plane) where the perpendicular from $\vec{P}$ meets it.

Line case: Given a line in parametric form:

$$g : \vec{r}(t) = \vec{A} + t\vec{v}$$

and a point $\vec{P}$ not on the line, the foot of the perpendicular $\vec{F}$ satisfies:

$$(\vec{P} - \vec{F}) \perp \vec{v} \quad \Rightarrow \quad (\vec{P} - (\vec{A} + t\vec{v})) \cdot \vec{v} = 0$$

Solve this scalar product for t, then compute:

$$\vec{F} = \vec{A} + t\vec{v}$$

Plane case: Given a plane in normal form:

$$\langle \vec{n}, \vec{x} - \vec{Q} \rangle = 0$$

then the foot of the perpendicular from point $\vec{P}$ to the plane is:

$$\vec{F} = \vec{P} - ((\vec{P} - \vec{Q}) \cdot \vec{n}) \cdot \vec{n}$$

26.20.2 Mirror Point

The **mirror point** (or reflected point) of $\vec{P}$ across a line or plane is the point $\vec{P'}$ such that the midpoint between $\vec{P}$ and $\vec{P'}$ is the foot of the perpendicular.
Formula:

$$\vec{P'} = 2\vec{F} - \vec{P}$$

Where $\vec{F}$ is the foot of the perpendicular from $\vec{P}$ to the line or plane.

27 Algebraic Structures

27.1 Introduction

Algebraic structures are mathematical systems consisting of a set equipped with one or more operations that satisfy certain axioms. They provide a unified language to study various objects in mathematics, from numbers and matrices to functions and vector spaces. Understanding these structures is fundamental in abstract algebra and has applications in computer science, cryptography, coding theory, and physics.

27.2 Operations: Internal and External

An **internal composition law** is a binary operation that takes two elements from a set and returns another element in the same set. Formally, for a set S and operation $\circ$, we have:

$$\circ : S \times S \to S$$

An **external composition law** involves a second set acting on the structure, such as scalar multiplication in vector spaces:

$$\cdot : K \times V \to V$$

where K is a field and V is a vector space.

27.3 Properties of Operations

Let $*$ be a binary operation on a set S. The most important properties include:

- **Associativity:** $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$
- **Commutativity:** $a * b = b * a$ for all $a, b \in S$
- **Identity Element:** There exists $e \in S$ such that $a * e = e * a = a$ for all $a \in S$
- **Inverse Element:** For every $a \in S$, there exists $a^{-1} \in S$ such that $a * a^{-1} = a^{-1} * a = e$
- **Distributivity:** $a \circ (b \bullet c) = (a \circ b) \bullet (a \circ c)$ and/or $(b \bullet c) \circ a = (b \circ a) \bullet (c \circ a)$

27.4 Homomorphisms and Isomorphisms

Let $(G, \oplus)$ and $(H, \oplus')$ be two algebraic structures.

- A **homomorphism** is a function $\varphi : G \to H$ such that:

$$\varphi(a \oplus b) = \varphi(a) \oplus' \varphi(b), \quad \forall a, b \in G$$

- An **isomorphism** is a bijective homomorphism. If such a map exists, we say the structures are **isomorphic**, written as $G \cong H$.

27.5 Common Algebraic Structures

The following table lists common algebraic structures along with their notation and defining properties. Let $\oplus$ denote the additive operation and $\odot$ the multiplicative one:

Name	Notation	Properties
Semigroup	$(S, \oplus)$	Associative
Monoid	$(M, \oplus)$	Associative, Identity element
Group	$(G, \oplus)$	Associative, Identity, Inverses
Abelian Group	$(A, \oplus)$	Group + Commutativity
Ring	$(R, \oplus, \odot)$	$(R, \oplus)$ is an abelian group, $(R, \odot)$ is a semigroup, Distributivity: $a \odot (b \oplus c) = a \odot b \oplus a \odot c$
Commutative Ring	$(R, \oplus, \odot)$	Ring + $(R, \odot)$ is commutative
Field	$(K, \oplus, \odot)$	Commutative Ring + $(K \setminus \{0\}, \odot)$ is an abelian group
Vector Space	$(V, \oplus, \cdot)$	$(V, \oplus)$ is an abelian group, $\cdot : K \times V \to V$ (scalar mult.), Distributivity, associativity, identities

28 Vector Spaces

28.1 Definition

A **vector space** is a set V with two operations, vector addition and scalar multiplication, such that:

- The set V is closed under vector addition.
- The set V is closed under scalar multiplication.
- Vector addition is commutative.
- Vector addition is associative.
- There exists a zero vector $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$.
- For every vector $\vec{v} \in V$, there exists a vector $-\vec{v} \in V$ such that $\vec{v} + (-\vec{v}) = \vec{0}$.
- Scalar multiplication is distributive with respect to vector addition: $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ for all $a \in F$ and $\vec{u}, \vec{v} \in V$.
- Scalar multiplication is distributive with respect to field addition: $(a+b)\vec{v} = a\vec{v} + b\vec{v}$ for all $a, b \in F$ and $\vec{v} \in V$.
- Scalar multiplication is associative: $a(b\vec{v}) = (ab)\vec{v}$ for all $a, b \in F$ and $\vec{v} \in V$.
- The multiplicative identity acts as a scalar: $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.

The set F is a field, and the elements of V are called **vectors**.

28.2 Examples

1. The set of all n-tuples of real numbers $\mathbb{R}^n$ is a vector space over the field of real numbers $\mathbb{R}$.
2. The set of all polynomials of degree less than or equal to n is a vector space over the field of real numbers $\mathbb{R}$.
3. The set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space over the field of real numbers $\mathbb{R}$.
4. The set of all $m \times n$ matrices with real entries is a vector space over the field of real numbers $\mathbb{R}$.

28.3 Subspaces

A subset W of a vector space V is a **subspace** of V if:

- The zero vector $\vec{0} \in W$.
- For all $\vec{u}, \vec{v} \in W$, $\vec{u} + \vec{v} \in W$.
- For all $a \in F$ and $\vec{v} \in W$, $a\vec{v} \in W$.

If W is a subspace of V, we write $W \subseteq V$.

Note: The intersection of two subspaces is also a subspace.

28.4 Linear Combinations

A **linear combination** of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in a vector space V is an expression of the form:

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_n\vec{v}_n$$

where $a_1, a_2, \ldots, a_n$ are scalars from the field F. The set of all linear combinations of a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is called the **span** of those vectors, denoted by $\text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$. The span of a set of vectors is a subspace of the vector space V. **Note:** The span of a set of vectors is the smallest subspace containing those vectors.

28.5 Properties of the subspaces

- The intersection of two subspaces is a subspace.
- The union of two subspaces is not necessarily a subspace.
- The sum of two subspaces U and W is defined as:

$$U + W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

  The sum of two subspaces is a subspace.

Note: The sum of two subspaces is the smallest subspace containing both subspaces.

- The direct sum of two subspaces U and W is defined as:

$$U \oplus W = \{\vec{u} + \vec{w} : \vec{u} \in U, \vec{w} \in W\}$$

  The direct sum of two subspaces is a subspace.

- The direct sum of two subspaces is the smallest subspace containing both subspaces, such that $U \cap W = \{\vec{0}\}$.
- The direct sum of two subspaces is denoted by $U \oplus W$.

28.6 Linear Independence

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ in a vector space V is said to be **linearly independent** if the only solution to the equation:

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \ldots + \lambda_n \vec{v}_n = 0$$

or

$$\sum_{i=1}^{n} \lambda_i \vec{v}_i = 0$$

is $a_1 = a_2 = \ldots = a_n = 0$. If there exists a non-trivial solution to this equation, then the set of vectors is said to be **linearly dependent**. A set of vectors is linearly independent if and only if the only linear combination of those vectors that equals the zero vector is the trivial combination where all coefficients are zero.

28.6.1 Properties of the linear independence

- A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent if and only if the only linear combination of those vectors that equals the zero vector is the trivial combination where all coefficients are zero.
- If a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent, then any subset of that set is also linearly independent.
- If a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly dependent, then at least one vector in that set can be expressed as a linear combination of the others.

28.7 Base

- A **base** of a vector space V is a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ that is linearly independent and spans the vector space V.
- The number of vectors in a base of a vector space is called the **dimension** of the vector space.
- The dimension of a vector space V is denoted by $\dim(V)$.
- If V has a finite base, then it is said to be **finite-dimensional**.
- If V does not have a finite base, then it is said to be **infinite-dimensional**.

28.8 Dimension

The **dimension** of a vector space V is the number of vectors in a base of V.

- The dimension of a vector space is denoted by $\dim(V)$.
- The dimension of a vector space can be finite or infinite.
- If the dimension of a vector space is finite, then it is said to be **finite-dimensional**.
- If the dimension of a vector space is infinite, then it is said to be **infinite-dimensional**.

28.8.1 How to find the base of a set vector

To find the base of a set of vectors, we can use the following steps:

1. Write the vectors as columns of a matrix.
2. Row reduce the matrix to echelon form.
3. The non-zero rows of the echelon form matrix correspond to the base of the vector space spanned by the original set of vectors.

The number of non-zero rows in the echelon form matrix is equal to the dimension of the vector space spanned by the original set of vectors.
Note: The base of a vector space is not unique. Different bases can span the same vector space.

28.9 Basis Extension Theorem

Let V be a vector space over a field K, and let

$$v_1, \dots, v_r, \quad w_1, \dots, w_s \in V.$$

Suppose that $(v_1, \dots, v_r)$ is a linearly independent tuple and that

$$\operatorname{span}(v_1, \dots, v_r, w_1, \dots, w_s) = V.$$

Then it is possible to extend $(v_1, \dots, v_r)$ to a basis of V by possibly adding suitable vectors from the set $\{w_1, \dots, w_s\}$.

Proof

If $\operatorname{span}(v_1, \dots, v_r) = V$, the statement is obvious. So assume

$$\operatorname{span}(v_1, \dots, v_r) \neq V.$$

Then there exists at least one w_i such that $w_i \notin \operatorname{span}(v_1, \dots, v_r)$; otherwise, if all $w_i \in \operatorname{span}(v_1, \dots, v_r)$, then

$$\operatorname{span}(v_1, \dots, v_r, w_1, \dots, w_s) = \operatorname{span}(v_1, \dots, v_r) = V,$$

which contradicts our assumption that $\operatorname{span}(v_1, \dots, v_r) \neq V$.

The tuple $(w_i, v_1, \dots, v_r)$ is linearly independent, because from

$$\sum_{j=1}^{r} \lambda_j v_j + \lambda w_i = 0$$

it follows that $\lambda = 0$ (since $w_i \notin \operatorname{span}(v_1, \dots, v_r)$), and then also $\lambda_j = 0$ for all j because the v_j are linearly independent.

Possibly, $(w_i, v_1, \dots, v_r)$ is still not a basis of V. Then we repeat the previous step and keep adding further w_i until the tuple extends $(v_1, \dots, v_r)$ to a basis of V. This process terminates after finitely many steps, since

$$\operatorname{span}(v_1, \dots, v_r, w_1, \dots, w_s) = V.$$

$\mathfrak{QED}$

Note: Every finitely generated vector space V has a basis.

28.10 Exchange Lemma

Let $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ be bases of a vector space V. Then, for every v_i, there exists a w_j such that if we replace v_i by w_j in the tuple $(v_1, \ldots, v_n)$, it still forms a basis of V.

Proof

Let $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$ be two bases of V. Suppose we remove v_i from the first basis. The truncated tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ satisfies

$$\text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \neq V,$$

because if $\text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) = V$, then v_i would lie in the span of the remaining vectors and could be written as a linear combination of them. This would contradict the assumption that $(v_1, \ldots, v_n)$ is linearly independent and a basis of V.

By the Basis Extension Theorem, we can extend the truncated tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ to a basis of V by adding vectors from $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_1, \ldots, w_m)$. Therefore, by the Basis Extension Theorem, there exists a w_j such that

$$w_j \notin \text{span}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n),$$

and the tuple $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_j)$ is linearly independent.

If this tuple does not form a basis, we can again apply the Basis Extension Theorem and add one of the vectors $v_1, \ldots, v_n$ to complete the basis. Clearly, the only possibility is to add v_i, but this would imply that the tuple $(v_1, \ldots, v_n, w_j)$ is not a basis, as w_j would then be linearly dependent on the other vectors. Therefore, $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n, w_j)$ must form a basis of V.
$\mathfrak{QED}$

28.11 Dimension of a sum of subspaces

Let U and W be two subspaces of a vector space V. Then the dimension of the sum of the two subspaces is given by:

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

28.12 Linear Independence of polynomials

Let P_n be the vector space of polynomials of degree at most n. The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is a basis for P_n. The dimension of P_n is $n + 1$.

- The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.
- The set of polynomials $\{1, x, x^2, \ldots, x^n\}$ spans the vector space P_n.
- The dimension of P_n is $n + 1$.

To prove that the set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent, we can use the following steps:

1. Assume that there exists a linear combination of the polynomials that equals zero:

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

   where $a_0, a_1, \ldots, a_n$ are scalars.

2. Since the left-hand side is a polynomial of degree at most n, it can only be equal to zero if all coefficients are zero.

3. Therefore, we have $a_0 = a_1 = \ldots = a_n = 0$, which proves that the set of polynomials $\{1, x, x^2, \ldots, x^n\}$ is linearly independent.

So you only have to prove that the set of coefficients vector is linearly independent.

28.13 Interpolation Polynomial

Given the $n+1$ points (x_k, y_k), with $0 \leq k \leq n$ and all x_k distinct, there exists exactly one polynomial $p_n \in P_n$ such that $y_k = p_n(x_k)$ for all $0 \leq k \leq n$. This polynomial is called the interpolation polynomial.

Proof

The uniqueness follows immediately from Remark 3.100. We prove the existence by induction on n. For $n = 0$, choose $p_0(x) = y_0$.

Now assume the statement is true for $n-1$. Let the polynomial p_{n-1} interpolate the points $(x_0, y_0), \ldots, (x_{n-1}, y_{n-1})$. Define

$$p_n(x) = p_{n-1}(x) + q(x),$$

where

$$q(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_{n-1})}{(x_n - x_0)(x_n - x_1)\cdots(x_n - x_{n-1})}(y_n - p_{n-1}(x_n)).$$

We have $q \in P_n$, and by Corollary 3.98, it follows that $p_n \in P_n$. Furthermore, $q(x_k) = 0$ for $k \leq n-1$ because a linear factor in the numerator always vanishes at x_k. Therefore, $p_n(x_k) = y_k$ for $k \leq n-1$. Additionally, we have

$$q(x_n) = y_n - p_{n-1}(x_n),$$

so that $p_n(x_n) = y_n$.
$\mathfrak{QED}$

28.13.1 Example of the interpolation polynomial

Consider the three points $(-2, 1)$, $(-1, -1)$, and $(1, 1)$. By Theorem 3.101, these points uniquely define an interpolating parabola p_2. This parabola can be determined using the definition of p_n from the proof of Theorem 3.101. For hand calculations and a small number of points to interpolate, the following approach is also useful. The general form of the polynomial is

$$p_2(x) = ax^2 + bx + c.$$

Substituting the three points into this form gives the system of equations:

$$\begin{aligned} 1 &= a + b + c \quad \text{(from the point (1, 1))} \\ -1 &= a - b + c \quad \text{(from the point (-1, -1))} \\ 1 &= 4a - 2b + c \quad \text{(from the point (-2, 1))} \end{aligned}$$

This leads to the system of equations:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 4 & -2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Solving this system gives $a = 1$, $b = 1$, and $c = -1$, so the interpolation polynomial is

$$p_2(x) = x^2 + x - 1.$$

29 Dot Product, Euclidean and Unitary Space

29.1 Scalar Product

Let V be a vector space over a field K. A mapping $\langle \cdot, \cdot \rangle : V \times V \to K$ is called a scalar product (or inner product) if the following conditions are satisfied:

SP1: Symmetry
For all $a, b \in V$:

$$\langle a, b \rangle = \begin{cases} \langle b, a \rangle & \text{if } K = \mathbb{R}, \\ \overline{\langle b, a \rangle} & \text{if } K = \mathbb{C}. \end{cases}$$

SP2: Linearity in the First Argument
For all $a, b, c \in V$:

$$\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$$

and

$$\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle.$$

SP3: Homogeneity in the First Argument
For all $\alpha \in K$, we have:

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle = \begin{cases} \langle a, \alpha b \rangle & \text{if } K = \mathbb{R}, \\ \langle a, \alpha b \rangle & \text{if } K = \mathbb{C}. \end{cases}$$

SP4: Positive Definiteness
For all $a \in V \setminus \{0\}$:

$$\langle a, a \rangle > 0,$$

and

$$\langle 0, 0 \rangle = 0.$$

29.2 Standard Scalar Product for Complex Number

Let $a = (a_i)_{i=1}^n$ and $b = (b_i)_{i=1}^n$ be vectors in $\mathbb{C}^n$. The standard scalar product is defined by

$$\langle a, b \rangle := \sum_{i=1}^{n} a_i b_i.$$

29.3 Scalar Product on $C[a, b]$

Let $f, g \in C[a, b]$. The scalar product on $C[a, b]$ is defined by

$$\langle f, g \rangle := \int_a^b f(x) \cdot g(x) \, dx.$$

29.4 Euclidean and Unitary Vector Spaces

A real vector space equipped with a scalar product is called an *Euclidean vector space*, while a complex vector space with a scalar product is called a *unitary vector space*.

29.5 Norms in Vector Spaces

Let V be a K-vector space and $a, b \in V$. A function $\| \cdot \| : V \to \mathbb{R}$ is called a norm if and only if the following conditions hold:

- N0: $\|a\| \in \mathbb{R}$,
- N1: $\|a\| \geq 0$,

- N2: $\|a\| = 0 \iff a = 0$,
- N3: $\forall \lambda \in K,\ \|\lambda a\| = |\lambda|\|a\|$,
- N4: (Triangle Inequality) $\|a + b\| \leq \|a\| + \|b\|$.

29.5.1 Induced Norm by a Scalar Product

As in the special case $V = \mathbb{R}^n$, a scalar product induces a norm. In a unitary (or Euclidean) space, the scalar product induces a (standard) norm defined by

$$\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}.$$

29.6 Cauchy-Schwarz Inequality in Unitary Vector Spaces

In all unitary vector spaces V, the Cauchy-Schwarz inequality holds:

$$|\langle a, b \rangle| \leq \|a\|\|b\| \quad \forall a, b \in V.$$

Proof of the Triangle Inequality

Both sides of the triangle inequality are real and, in particular, non-negative. Therefore, it is sufficient to prove that the squares of both sides satisfy the desired inequality, i.e., we need to show:

$$\langle a + b, a + b \rangle \leq (\|a\| + \|b\|)^2.$$

First, we expand the left-hand side:

$$\langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle.$$

Since $\langle b, a \rangle = \langle a, b \rangle$, we have:

$$\langle a, b \rangle + \langle b, a \rangle = 2\,\mathrm{Re}\langle a, b \rangle.$$

Now, we know that the absolute value of a complex number is always greater than or equal to its real part, so:

$$2\,\mathrm{Re}\langle a, b \rangle \leq 2|\langle a, b \rangle|.$$

Using the Cauchy-Schwarz inequality (3.6), we can further bound this by:

$$2\,\mathrm{Re}\langle a, b \rangle \leq 2\|a\|\|b\|.$$

Thus, we have:

$$\langle a + b, a + b \rangle \leq \langle a, a \rangle + 2\|a\|\|b\| + \langle b, b \rangle.$$

Using the definition of the norm, $\|a\|^2 = \langle a, a \rangle$ and $\|b\|^2 = \langle b, b \rangle$, we obtain:

$$\langle a + b, a + b \rangle \leq \|a\|^2 + 2\|a\|\|b\| + \|b\|^2.$$

This is exactly the expansion of $(\|a\| + \|b\|)^2$, which completes the proof.

30 Orthogonality

30.1 Orthogonality and Projection

Let $a, b \in V$. The vectors a and b are orthogonal to each other if

$$\langle a, b \rangle = 0.$$

This is written as $a \perp b$.

With the same proof as in Theorem 2.25, the Pythagorean Theorem holds

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2$$

for $a \perp b$ in all unitary vector spaces.

For the orthogonal projection $p_b(a)$ of a vector a onto b, with $b \neq 0$, the formula in every unitary vector space is:

$$p_b(a) = \frac{\langle a, b \rangle}{\langle b, b \rangle} b.$$

30.2 Orthogonal Projection and Orthogonal Complement

Let U be a finitely generated subspace of V and $a \in V$. A vector $p_U(a) \in U$ is called the orthogonal projection of a onto U if

$$a - p_U(a) \perp u \quad \forall u \in U \quad (3.9).$$

The question arises about well-definedness, i.e., whether such a vector $p_U(a)$ always exists and if it is unique. The following concept helps in the discussion of uniqueness.

For $M \subseteq V$, the **orthogonal complement** of M is defined as

$$M^\perp = \{v \in V \mid v \perp u \forall u \in M\}.$$

- $M^\perp$ is a subspace of V.
- Let U be a subspace of V. Then, we have $U \cap U^\perp = \{0\}$.

Proof: 1.) We need to check the closure property. For $u \in M$, $x, y \in M^\perp$, and $\lambda \in \mathbb{R}$, we have:

$$\langle x + y, u \rangle = \langle x, u \rangle + \langle y, u \rangle = 0$$
$$\langle \lambda x, u \rangle = \lambda \langle x, u \rangle = 0.$$

2.) Let $a \in U \cap U^\perp$. Then, we have $\langle a, u \rangle = 0$ for all $u \in U$, because $a \in U^\perp$, and in particular, $\langle a, a \rangle = 0$ since $a \in U$, which implies $a = 0$.

30.3 Orthogonality of the basis to the Complement

Let U be as before, and let $(u_1, \ldots, u_m)$ be a basis of U. For $v \in V$, we have:

$$v \in U^\perp \quad \text{if and only if} \quad \langle v, u_i \rangle = 0 \quad \forall 1 \leq i \leq m.$$

30.4 How to find the orthogonal complement

Let U be a finitely generated subspace of V with basis $(u_1, \ldots, u_m)$. To find the orthogonal complement $U^\perp$, we can solve the system of equations:

$$\langle v, u_i \rangle = 0 \quad \forall 1 \leq i \leq m.$$

This system can be expressed in matrix form as $A \cdot v = 0$, where A is the matrix whose rows are the vectors u_i and v is the vector we want to find in $U^\perp$.

30.5 Orthogonal and Orthonormal Systems

Let $B = (v_1, \dots, v_m)$ be an m-tuple of vectors in $V \setminus \{0\}$.

- B is called an orthogonal system in V if all the vectors v_i are pairwise orthogonal.
- An orthogonal system is called an orthonormal system if, in addition, $\|v_i\| = 1$ for all $i = 1, \dots, m$.
- An orthogonal system that forms a basis of V is called an orthogonal basis of V.
- An orthonormal system that forms a basis of V is called an orthonormal basis of V.

Notation:

OG-System OG-Basis ON-System ON-Basis

Using the Kronecker delta symbol $\delta_{i,j}$:

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

we have $\langle v_i, v_j \rangle = \delta_{i,j}$ for any orthonormal system.

30.6 Writing a vector in terms of the Orthogonal Basis

Example: The vectors

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

are orthogonal and normalized, but do not form a basis. Thus, they constitute an **orthonormal system**.
Example: The vectors

$$a_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

form a basis, are orthogonal and normalized. Therefore, they form an **orthonormal basis**.

30.7 Linear Independence of Orthogonal Systems

An orthogonal system $(v_1, \dots, v_n)$ is linearly independent.
Proof: Let $\sum_{i=1}^n \lambda_i v_i = 0$. Then, for any v_j, it follows that

$$\left\langle \sum_{i=1}^n \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j \rangle = 0.$$

Due to the orthogonality of the system, all terms vanish except $\lambda_j \langle v_j, v_j \rangle$, so

$$\lambda_j \langle v_j, v_j \rangle = 0.$$

Since $v_j \neq 0$ in any orthogonal system, we have $\langle v_j, v_j \rangle = \|v_j\|^2 > 0$, implying $\lambda_j = 0$ for all $1 \leq j \leq n$. □

30.8 Representation with Respect to an Orthogonal Basis

Let $B = (v_1, \dots, v_n)$ be an orthogonal basis of a vector space V. Then for every $v \in V$, we have:

$$v = \sum_{k=1}^n \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k,$$

that is, the coordinates of v with respect to the basis B are given by

$$\left(\frac{\langle v, v_k \rangle}{\|v_k\|^2} \right)^T_{1 \leq k \leq n}.$$

30.9 Orthogonal Projection and Direct Sum Decomposition

Let $B = (v_1, \ldots, v_m)$ be an orthogonal system in V, and let $U = L(B)$ be the subspace of V spanned by B.

- For every $v \in V$, the orthogonal projection of v onto U is given by:

$$p_U(v) = \sum_{i=1}^{m} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

- Every vector $v \in V$ can be uniquely written as the sum $v = p_U(v) + w$, with $w \in U^\perp$. In this case, we have:

$$w = v - p_U(v).$$

- $V = U \oplus U^\perp$, which means that every vector in V can be uniquely decomposed into a sum of a vector from U and a vector from $U^\perp$.
- If $\dim(V) = n$, then:

$$\dim(U) + \dim(U^\perp) = n \quad \text{for every subspace}\, U.$$

30.10 The Gram-Schmidt Process

The Gram-Schmidt process is a method for orthonormalizing a set of vectors in an inner product space. Given a finite set of linearly independent vectors $(v_1, \ldots, v_n)$, the process generates an orthonormal basis $(u_1, \ldots, u_n)$ as follows:

1. Set $u_1 = \frac{v_1}{\|v_1\|}$.
2. For $k = 2, \ldots, n$:
   (a) Set $w_k = v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j$.
   (b) Set $u_k = \frac{w_k}{\|w_k\|}$.
3. The resulting set $(u_1, \ldots, u_n)$ is an orthonormal basis of the subspace spanned by $(v_1, \ldots, v_n)$.

Note: The Gram-Schmidt process can be applied to any finite set of linearly independent vectors in an inner product space, and it is particularly useful for constructing orthonormal bases in Euclidean spaces.
Example: Let $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$, and $v_3 = (1, 1, 1)$. The Gram-Schmidt process yields:

$$u_1 = \frac{v_1}{\|v_1\|} = (1, 0, 0), \quad u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\|v_2 - \langle v_2, u_1 \rangle u_1\|} = (0, 1, 0)\,,$$

$$u_3 = \frac{v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2}{\|v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2\|} = (0, 0, 1)\,.$$

Thus, the orthonormal basis is $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.

30.11 Existence of Orthonormal Bases and Orthogonal Projections

- Every finitely generated unitary vector space has an orthonormal basis.
- I is, in general, false for vector spaces that are not finitely generated.
- We now provide the existence proof of the orthogonal projection.
- Let V be a unitary vector space and U a finitely generated subspace. Then for every $v \in V$, the orthogonal projection $p_U(v)$ of v onto U exists.
- Let V be a finitely generated unitary vector space and U any subspace. Then we have $V = U \oplus U^\perp$, and

$$\dim(V) = \dim(U) + \dim(U^\perp).$$

- Every hyperplane in $\mathbb{R}^n$ admits a normal form; the normal vector is unique up to scalar multiplication.
- Let V be as above and let $v_1, \ldots, v_m \in V$. If it is possible to construct orthonormal vectors $w_1, \ldots, w_m$ from them using the Gram-Schmidt process, then $(v_1, \ldots, v_m)$ are linearly independent.

30.12 Best Approximation

Let V be a unitary vector space and U a finitely generated subspace of V. A vector $v* \in U$ is called Best Approximation in U on v, if:

$$\|v * -v\| = \inf_{x\in U} \|x - v\|$$

also

$$\|v * -P_u(v)\| = \inf_{u\in U} \|u - v\|$$

31 Linear Maps

A linear map $f : V \to W$ is a function that satisfies the following properties:

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$ for all $v_1, v_2 \in V$.
2. $f(\lambda v) = \lambda f(v)$ for all $v \in V$ and $\lambda \in \mathbb{R}$.
3. $f(0) = 0$.

This kind of function is also called **Homomorphism**.
If $V = W$ is it called **Endomorphism**.

31.1 Types of Linear Maps

1. **Injective or Monomorphism** (One-to-One): A linear map $f : V \to W$ is injective if $f(v_1) = f(v_2)$ implies $v_1 = v_2$.
2. **Surjective or Epimorphism** (Onto): A linear map $f : V \to W$ is surjective if for every $w \in W$, there exists a $v \in V$ such that $f(v) = w$.
3. **Bijective or Isomorphism**: A linear map $f : V \to W$ is bijective if it is both injective and surjective.

Note: Two vector spcaces are called **isomorphic** if there exists a bijective linear map between them. In this case, we can say that the two vector spaces are **isomorphic** and we write $V \cong W$.

31.2 Properties of Linear Maps

1. The composition of two linear maps is a linear map.
2. The inverse of a bijective linear map is also a linear map.
3. The zero map $f : V \to W$ defined by $f(v) = 0$ for all $v \in V$ is a linear map.
4. The identity map $\mathrm{id}_V : V \to V$ defined by $\mathrm{id}_V(v) = v$ for all $v \in V$ is a linear map.
5. The sum of two linear maps $f : V \to W$ and $g : V \to W$ is a linear map defined by $(f + g)(v) = f(v) + g(v)$.
6. The scalar multiplication of a linear map $f : V \to W$ by a scalar c is a linear map defined by $(cf)(v) = c(f(v))$.
7. The composition of linear maps is associative, i.e., $(f \circ g) \circ h = f \circ (g \circ h)$.
8. The composition of linear maps is distributive over addition, i.e., $f \circ (g + h) = f \circ g + f \circ h$.
9. The composition of linear maps is compatible with scalar multiplication, i.e., $(cf) \circ g = c(f \circ g)$.
10. Linear Maps compose a vector space over the field of scalars.

$$(Hom(V, W), +, \cdot)$$

31.3 The Kernel of a Linear Map

The kernel of a linear map $f : V \to W$ is the set of all vectors in V that are mapped to the zero vector in W:

$$\ker(f) = \{v \in V \mid f(v) = 0\}.$$

31.4 The Image of a Linear Map

The image of a linear map $f : V \to W$ is the set of all vectors in W that can be expressed as $f(v)$ for some $v \in V$:

$$\mathrm{Im}(f) = \{w \in W \mid w = f(v) \text{ for some } v \in V\}.$$

31.5 The Rank of a Linear Map

The rank of a linear map $f : V \to W$ is the ension of its image:

$$\text{rank}(f) = (\text{Im}(f)).$$

31.6 The Nullity of a Linear Map

The nullity of a linear map $f : V \to W$ is the ension of its kernel:

$$\text{nullity}(f) = (\ker(f)).$$

31.7 The Rank-Nullity Theorem

The rank-nullity theorem states that for a linear map $f : V \to W$:

$$(\ker(f)) + \dim(\text{Im}(f)) = \dim(V).$$

31.8 Proof of the injectivity of the Kernel

Suppose we have a linear map f and two vectors v_1 and v_2 which are not equal. We are going to assume that they both map to the $\vec{0}$ therefore:

$$f(v_1) = f(v_2) = \vec{0}$$

Because of the linearity of the map we can write:

$$f(v_1 - v_2) = f(v_1) - f(v_2) = \vec{0} - \vec{0} = \vec{0}$$

And that contradicts the assumption that v_1 and v_2 are not equal. Therefore, the kernel of a linear map is injective.

31.9 Dimension Formula for Linear Mappings

Let $f : V \to W$ be linear and $\dim(V) = n$. Then

$$\dim(\ker(f)) + \text{rank}(f) = n.$$

Remember that $\ker(f)$ forms a subspace of V and therefore $\dim(\ker(f)) := r \leq n$. We extend an arbitrary basis $(v_1, \ldots, v_r)$ of $\ker(f)$ to a basis $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ of V. Setting $w_{r+i} = f(v_{r+i})$ for $i = 1, \ldots, n-r$, we have $\forall v \in V$:

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n) \\ &= \lambda_1 \underbrace{f(v_1)}_{=0} + \cdots + \lambda_r \underbrace{f(v_r)}_{=0} + \lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n) \\ &= \lambda_{r+1} f(v_{r+1}) + \cdots + \lambda_n f(v_n) \\ &= \lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n \end{aligned}$$

Thus, $\text{Im}(f) = \text{span}(w_{r+1}, \ldots, w_n)$. We now show that $w_{r+1}, \ldots, w_n$ are linearly independent. Let

$$\lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n = 0.$$

From

$$0 = \lambda_{r+1} w_{r+1} + \cdots + \lambda_n w_n = f(\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n)$$

it follows that

$$\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n \in \ker(f).$$

Therefore,

$$\lambda_{r+1} v_{r+1} + \cdots + \lambda_n v_n = \lambda_1 v_1 + \cdots + \lambda_r v_r$$

for some $\lambda_1, \ldots, \lambda_r$. Since $v_1, \ldots, v_n$ are linearly independent, we have $\lambda_1 = \cdots = \lambda_n = 0$. Thus, the vectors $w_{r+1}, \ldots, w_n$ are linearly independent.

It follows that $\dim(\text{Im}(f)) = \text{rank}(f) = n - r$ and therefore

$$\dim(\ker(f)) + \dim(\text{Im}(f)) = \dim(\ker(f)) + \text{rank}(f) = r + n - r = n.$$

$\mathfrak{QED}$

31.10 Identifying the type of linear map

- If $\dim(\ker(f)) = 0$ then the map is injective.
- If $\dim(\ker(f)) = \dim(V)$ then the map is surjective.
- If $\dim(\ker(f)) = \dim(V)$ and $\dim(W) = 0$ then the map is bijective.

31.11 Image of the basis

Let $(V) = \dim(W)$, For the basis $(v_1, \cdots, v_n)$ of V we have the image $(w_1, \cdots, w_n)$ of W. Then:

- If $\dim(\ker(f)) = 0$ then the map is injective and $(w_1, \cdots, w_n)$ is a basis of W.
- If $\dim(\ker(f)) = \dim(V)$ then the map is surjective and $(w_1, \cdots, w_n)$ is a spanning set of W.
- If $\dim(\ker(f)) = \dim(V)$ and $\dim(W) = 0$ then the map is bijective and $(w_1, \cdots, w_n)$ is a basis of W.

31.12 Linear Maps and Matrices

Let V and W be finite-ensional vector spaces over the same field K. If $\dim V = n$ and $\dim W = m$, then a linear map $f : V \to W$ can be represented by an $m \times n$ matrix. The action of the linear map on a vector can be expressed as:

$$f(x) = A \cdot x,$$

where A is the matrix representation of f and v is the vector represented in a column format. The columns of the matrix A are the images of the basis vectors of V under the linear map f.

32 Matrices

In this section, we explore the fundamental concepts of matrices, their operations, and important properties that form the foundation of linear algebra.

32.1 Definition of a Matrix

A matrix is a rectangular array of numbers, symbols, or expressions arranged in rows and columns. Formally, an $m \times n$ matrix A consists of mn elements a_{ij} where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

The set of all $m \times n$ matrices with real entries is denoted by $\mathbb{R}^{m \times n}$. Special cases include:

- Square matrix: A matrix with the same number of rows and columns ($m = n$)
- Column vector: An $m \times 1$ matrix
- Row vector: A $1 \times n$ matrix
- Identity matrix I_n: An $n \times n$ matrix with ones on the main diagonal and zeros elsewhere
- Zero matrix: A matrix where all entries are zero

32.2 Matrix Addition and Subtraction

Matrix addition and subtraction are defined for matrices of the same dimensions.

32.2.1 Addition

For matrices $A, B \in \mathbb{R}^{m \times n}$, their sum $C = A + B$ is defined as:

$$c_{ij} = a_{ij} + b_{ij} \quad \text{for all } i = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, n$$

32.2.2 Subtraction

Similarly, the difference $C = A - B$ is defined as:

$$c_{ij} = a_{ij} - b_{ij} \quad \text{for all } i = 1, 2, \ldots, m \text{ and } j = 1, 2, \ldots, n$$

Matrix addition satisfies the following properties:

$$\begin{gathered} A + B = B + A \quad \text{(Commutativity)} \\ (A + B) + C = A + (B + C) \quad \text{(Associativity)} \\ A + O = A \quad \text{(Identity element)} \\ A + (-A) = O \quad \text{(Inverse element)} \end{gathered}$$

where O is the zero matrix.

32.3 Matrix Multiplication

Matrix multiplication is defined between matrices where the number of columns in the first matrix equals the number of rows in the second matrix.

For $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{p \times n}$, their product $C = AB \in \mathbb{R}^{m \times n}$ is defined as:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ip} b_{pj}$$

Matrix multiplication satisfies the following properties:

$$
\begin{aligned}
A(BC) &= (AB)C \quad \text{(Associativity)} \\
A(B+C) &= AB + AC \quad \text{(Left distributivity)} \\
(A+B)C &= AC + BC \quad \text{(Right distributivity)} \\
AI_n = A \quad &\text{and} \quad I_m A = A \quad \text{(Identity)}
\end{aligned}
$$

Note that matrix multiplication is generally not commutative, i.e., $AB \neq BA$ in most cases.

32.4 The Transpose of a Matrix

The transpose of a matrix $A \in \mathbb{R}^{m\times n}$, denoted $A^T \in \mathbb{R}^{n\times m}$, is obtained by interchanging rows and columns:

$$(A^T)_{ij} = a_{ji} \quad \text{for all } i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, m$$

Properties of the transpose include:

$$
\begin{aligned}
(A^T)^T &= A \\
(A+B)^T &= A^T + B^T \\
(AB)^T &= B^T A^T \\
(\alpha A)^T &= \alpha A^T \quad \text{for any scalar } \alpha
\end{aligned}
$$

Special matrices related to the transpose include:

- Symmetric matrix: $A = A^T$
- Skew-symmetric matrix: $A = -A^T$

32.5 The Equivalence of Matrices

Two matrices A and B are said to be equivalent if one can be transformed into the other through a finite sequence of elementary row operations. We write $A \sim B$ to denote this equivalence.

The elementary row operations are:

- Interchanging two rows: $R_i \leftrightarrow R_j$
- Multiplying a row by a non-zero scalar: $R_i \mapsto \alpha R_i$ where $\alpha \neq 0$
- Adding a multiple of one row to another: $R_i \mapsto R_i + \alpha R_j$ where $i \neq j$

Matrix equivalence is an equivalence relation, satisfying reflexivity, symmetry, and transitivity. Equivalent matrices represent the same linear system in different bases.

32.5.1 Row Echelon Form (REF)

A matrix is in row echelon form if:

- All rows consisting entirely of zeros are at the bottom of the matrix.
- The leading entry (first non-zero element) of each non-zero row is to the right of the leading entry of the row above it.
- All entries in a column below a leading entry are zeros.

32.5.2 Reduced Row Echelon Form (RREF)

A matrix is in reduced row echelon form if:

- It is in row echelon form.
- Each leading entry is 1.
- Each leading entry is the only non-zero entry in its column.

The Gauß-Jordan elimination algorithm proceeds as follows:

- Start with the leftmost non-zero column.
- Find the pivot (non-zero element) in this column. If necessary, swap rows to move a non-zero element to the pivot position.
- Divide the pivot row by the pivot value to make the pivot equal to 1.
- Eliminate all other entries in the pivot column by subtracting appropriate multiples of the pivot row.
- Cover the pivot row and column, and repeat steps 1-4 on the submatrix until all rows are processed.
- For RREF, eliminate all entries above each pivot as well.

32.6 The Inverse of a Matrix and Its Properties

For a square matrix $A \in \mathbb{R}^{n\times n}$, the inverse matrix A^{-1} (if it exists) satisfies:

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

32.6.1 Properties of the Inverse

$$\begin{aligned}
(A^{-1})^{-1} &= A \\
(AB)^{-1} &= B^{-1}A^{-1} \\
(A^T)^{-1} &= (A^{-1})^T \\
\det(A^{-1}) &= \frac{1}{\det(A)}
\end{aligned}$$

32.6.2 Finding the Inverse

There are several methods to find the inverse of a matrix:

Gauß-Jordan Method Form the augmented matrix $[A|I_n]$ and apply Gauß-Jordan elimination to transform it into $[I_n|A^{-1}]$:

1. Create the augmented matrix $[A|I_n]$
2. Apply row operations to transform the left side into I_n
3. The right side will be A^{-1}

Adjoint Method For an $n \times n$ matrix A:

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$$

where $\text{adj}(A)$ is the adjoint (or adjugate) of A, defined as the transpose of the cofactor matrix.

A matrix is invertible if and only if its determinant is non-zero. Such matrices are called non-singular or regular matrices.

32.7 The Rank of a Matrix and How to Find It

The rank of a matrix A, denoted $\text{rank}(A)$ or $\text{rg}(A)$, is the dimension of the column space (or equivalently, the row space) of A.

Equivalent definitions of rank include:

- The maximum number of linearly independent columns of A
- The maximum number of linearly independent rows of A
- The order of the largest non-zero minor of A
- The number of non-zero rows in any row echelon form of A

32.7.1 Finding the Rank

To find the rank of a matrix:

1. Transform the matrix into row echelon form using Gauß-Jordan elimination
2. Count the number of non-zero rows in the resulting matrix

Properties of rank include:

$$\begin{aligned}
\operatorname{rank}(A) &\leq \min(m, n) \text{ for } A \in \mathbb{R}^{m \times n} \\
\operatorname{rank}(A^T) &= \operatorname{rank}(A) \\
\operatorname{rank}(AB) &\leq \min(\operatorname{rank}(A), \operatorname{rank}(B)) \\
\operatorname{rank}(A + B) &\leq \operatorname{rank}(A) + \operatorname{rank}(B)
\end{aligned}$$

For a square matrix $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

- A is invertible
- $\operatorname{rank}(A) = n$
- $\det(A) \neq 0$
- The columns of A are linearly independent
- The rows of A are linearly independent
- $Ax = 0$ has only the trivial solution $x = 0$

32.8 The Definitions of Column Space, Row Space, and Null Space

These fundamental spaces associated with a matrix $A \in \mathbb{R}^{m \times n}$ provide important insights into its structure.

32.8.1 Column Space

The column space of A, denoted $\operatorname{Col}(A)$, is the span of the columns of A:

$$\operatorname{Col}(A) = \{\vec{y} \in \mathbb{R}^m : \vec{y} = A\vec{x} \text{ for some } \vec{x} \in \mathbb{R}^n\}$$

This is also called the range or image of the linear transformation represented by A. The dimension of the column space equals the rank of A.

32.8.2 Row Space

The row space of A, denoted $\operatorname{Row}(A)$, is the span of the rows of A:

$$\operatorname{Row}(A) = \operatorname{Col}(A^T)$$

The dimension of the row space also equals the rank of A.

32.8.3 Null Space

The null space (or kernel) of A, denoted $\operatorname{Null}(A)$ or $\operatorname{Ker}(A)$, is the set of all vectors that A maps to zero:

$$\operatorname{Null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$$

The dimension of the null space is called the nullity of A, denoted $\operatorname{nullity}(A)$.

32.8.4 Left Null Space

The left null space of A is the null space of A^T:

$$\text{Null}(A^T) = \{\vec{y} \in \mathbb{R}^m : A^T\vec{y} = \vec{0}\} = \{\vec{y} \in \mathbb{R}^m : \vec{y}^T A = \vec{0}^T\}$$

The Rank-Nullity Theorem connects these spaces:

$$\text{rank}(A) + \text{nullity}(A) = n$$

To find a basis for these spaces:

- Column space: Take the linearly independent columns of A
- Row space: Take the non-zero rows from any row echelon form of A
- Null space: Solve the homogeneous system $A\vec{x} = \vec{0}$ and express the general solution in terms of free variables

32.9 Examples of Matrix Operations

In this subsection, we provide detailed examples of Gauß-Jordan elimination and matrix multiplication to illustrate these fundamental matrix operations.

32.9.1 Example of Matrix Multiplication

Consider the matrices A and B given by:

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 0 & -2 \end{pmatrix} \in \mathbb{R}^{2\times 3} \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 \\ -1 & 3 \\ 4 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 2}$$

To compute the product $C = AB \in \mathbb{R}^{2\times 2}$, we calculate each entry c_{ij} using the formula:

$$c_{ij} = \sum_{k=1}^{3} a_{ik}b_{kj}$$

Let's calculate each entry of C:

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ &= 2\cdot 1 + 3\cdot(-1) + 1\cdot 4 \\ &= 2 - 3 + 4 = 3 \end{aligned}$$

$$\begin{aligned} c_{12} &= a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ &= 2\cdot 2 + 3\cdot 3 + 1\cdot 0 \\ &= 4 + 9 + 0 = 13 \end{aligned}$$

$$\begin{aligned} c_{21} &= a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ &= 1\cdot 1 + 0\cdot(-1) + (-2)\cdot 4 \\ &= 1 + 0 - 8 = -7 \end{aligned}$$

$$\begin{aligned} c_{22} &= a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \\ &= 1\cdot 2 + 0\cdot 3 + (-2)\cdot 0 \\ &= 2 + 0 + 0 = 2 \end{aligned}$$

Therefore, the product $C = AB$ is:

$$C = AB = \begin{pmatrix} 3 & 13 \\ -7 & 2 \end{pmatrix}$$

Let's verify that matrix multiplication is not generally commutative by attempting to compute BA:

Since B is a 3×2 matrix and A is a 2×3 matrix, the product BA would be a 3×3 matrix. However, this calculation cannot be performed since the number of columns in B (which is 2) does not equal the number of rows in A (which is 2). Thus, BA is undefined, demonstrating that matrix multiplication is not always commutative.

32.9.2 Example of Gauß-Jordan Elimination

We'll use Gauß-Jordan elimination to solve the linear system:

$$\begin{gathered} 2x + y - z = 8 \\ -3x - y + 2z = -11 \\ x + y + z = 3 \end{gathered}$$

First, we set up the augmented matrix:

$$\begin{pmatrix} 2 & 1 & -1 & | & 8 \\ -3 & -1 & 2 & | & -11 \\ 1 & 1 & 1 & | & 3 \end{pmatrix}$$

Now we apply Gauß-Jordan elimination to transform this into reduced row echelon form:

Step 1: We'll choose the first element in the first row as our pivot. Let's first swap row 1 and row 3 to get a simpler pivot:

$$\begin{pmatrix} 1 & 1 & 1 & | & 3 \\ -3 & -1 & 2 & | & -11 \\ 2 & 1 & -1 & | & 8 \end{pmatrix}$$

Step 2: Eliminate the first elements in rows 2 and 3:

Row 2 + 3 × Row 1:

$$\begin{pmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 2 & 5 & | & -2 \\ 2 & 1 & -1 & | & 8 \end{pmatrix}$$

Row 3 - 2 × Row 1:

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

Step 3: Make the pivot in row 2 equal to 1 by dividing the entire row by 2:

$$\begin{pmatrix} 1 & 1 & 1 & | & 3 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & -1 & -3 & | & 2 \end{pmatrix}$$

Step 4: Eliminate the second element in rows 1 and 3:

Row 1 - Row 2:

$$\begin{pmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & -1 & -3 & | & 2 \end{pmatrix}$$

Row 3 + Row 2:

$$\begin{pmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & -\frac{1}{2} & | & 1 \end{pmatrix}$$

Step 5: Make the pivot in row 3 equal to 1 by multiplying the entire row by -2:

$$\begin{pmatrix} 1 & 0 & -\frac{3}{2} & | & 4 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & 1 & | & -2 \end{pmatrix}$$

Step 6: Eliminate the third element in rows 1 and 2:
Row 1 + $\frac{3}{2}\times$ Row 3:

$$\begin{pmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & \frac{5}{2} & | & -1 \\ 0 & 0 & 1 & | & -2 \end{pmatrix}$$

Row 2 - $\frac{5}{2}\times$ Row 3:

$$\begin{pmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & -1+5=4 \\ 0 & 0 & 1 & | & -2 \end{pmatrix}$$

The matrix is now in reduced row echelon form:

$$\begin{pmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & 4 \\ 0 & 0 & 1 & | & -2 \end{pmatrix}$$

This corresponds to the system:

$$x = 1$$
$$y = 4$$
$$z = -2$$

Therefore, the solution to the original system is $x = 1$, $y = 4$, and $z = -2$.

32.9.3 Example of Finding the Inverse of a Matrix using Gauß-Jordan Elimination

Let's find the inverse of the matrix:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

We form the augmented matrix $[A|I_3]$:

$$\begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 3 & 2 & 1 & | & 0 & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix}$$

Now we apply Gauß-Jordan elimination:
Step 1: Make the first pivot equal to 1 by dividing the first row by 2:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 3 & 2 & 1 & | & 0 & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix}$$

Step 2: Eliminate the first element in rows 2 and 3:
Row 2 - 3 $\times$ Row 1:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & | & -\frac{3}{2} & 1 & 0 \\ 2 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix}$$

Row 3 - 2 $\times$ Row 1:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & | & -\frac{3}{2} & 1 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix}$$

Step 3: Make the second pivot equal to 1 by multiplying the second row by 2:

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & | & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & | & -3 & 2 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix}$$

Step 4: Eliminate the second element in row 1 and the third element in row 2:

Row 1 - $\frac{1}{2}\times$ Row 2:

$$\begin{pmatrix} 1 & 0 & 1 & | & 2 & -1 & 0 \\ 0 & 1 & -1 & | & -3 & 2 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix}$$

Row 2 + Row 3:

$$\begin{pmatrix} 1 & 0 & 1 & | & 2 & -1 & 0 \\ 0 & 1 & 0 & | & -4 & 2 & 1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix}$$

Step 5: Eliminate the third element in row 1:
Row 1 - Row 3:

$$\begin{pmatrix} 1 & 0 & 0 & | & 3 & -1 & -1 \\ 0 & 1 & 0 & | & -4 & 2 & 1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix}$$

The right side of the augmented matrix now gives us A^{-1}:

$$A^{-1} = \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

To verify, we can check that $AA^{-1} = I_3$:

$$\begin{aligned} AA^{-1} &= \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & -1 & -1 \\ -4 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2\cdot 3 + 1\cdot(-4) + 1\cdot(-1) & 2\cdot(-1) + 1\cdot 2 + 1\cdot 0 & 2\cdot(-1) + 1\cdot 1 + 1\cdot 1 \\ 3\cdot 3 + 2\cdot(-4) + 1\cdot(-1) & 3\cdot(-1) + 2\cdot 2 + 1\cdot 0 & 3\cdot(-1) + 2\cdot 1 + 1\cdot 1 \\ 2\cdot 3 + 1\cdot(-4) + 2\cdot(-1) & 2\cdot(-1) + 1\cdot 2 + 2\cdot 0 & 2\cdot(-1) + 1\cdot 1 + 2\cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} 6-4-1 & -2+2+0 & -2+1+1 \\ 9-8-1 & -3+4+0 & -3+2+1 \\ 6-4-2 & -2+2+0 & -2+1+2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3 \end{aligned}$$

This confirms that we have correctly found the inverse of matrix A.

32.10 The Determinant of a Matrix

The determinant of a square matrix $A \in \mathbb{R}^{n\times n}$, denoted $\det(A)$ or $|A|$, is a scalar value that provides important information about the matrix, including whether it is invertible and the volume scaling factor of the linear transformation represented by A. The determinant can be computed using various methods, including the Laplace expansion, row reduction, or the Leibniz formula. The determinant of a 2×2 matrix is given by:

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

For a 3×3 matrix, the determinant can be computed using the rule of Sarrus or the cofactor expansion:

$$\det(A) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh$$

The determinant of larger matrices can be computed using cofactor expansion along any row or column:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

where A_{ij} is the $(n-1)\times(n-1)$ submatrix obtained by deleting the i-th row and j-th column of A. The determinant has several important properties:

- $\det(A) = 0$ if and only if A is singular (not invertible).
- $\det(AB) = \det(A) \cdot \det(B)$ for any square matrices A and B of the same size.
- $\det(A^T) = \det(A)$.
- If a row (or column) of A is multiplied by a scalar α, then $\det(A)$ is multiplied by α.
- If two rows (or columns) of A are swapped, then $\det(A)$ changes sign.

$$\det(a, b, c) = -\det(b, a, c)$$

- If a row (or column) of A is added to another row (or column), then $\det(A)$ remains unchanged.
- If one of the columns is a linear combination of the others, then $\det(A) = 0$.
- The determinant of the identity matrix I_n is 1.
- The determinant of can splitet into the sum of more determinants:

$$\det(a, b, c + d) = \det(a, b, c) + \det(a, b, d)$$

- The determinant of a diagonal matrix is the product of its diagonal entries.
- The determinant of a triangular matrix (upper or lower) is the product of its diagonal entries.
- The determinant of a matrix is a multilinear function of its rows (or columns).
- The determinant is a continuous function of the entries of the matrix.
- The determinant can be computed using the Leibniz formula:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

  where S_n is the set of all permutations of $\{1, 2, \ldots, n\}$ and $\operatorname{sgn}(\sigma)$ is the sign of the permutation σ.

Laplace's Method (Cofactor Expansion)

Here's how to find the determinant of a matrix using Laplace's method:

Laplace's method, also known as cofactor expansion, allows you to compute the determinant of a square matrix by expanding along any row or column.

Steps:

1.**Choose a Row or Column:** Select any row or column of the matrix. It's often easiest to choose one with many zeros.
2.**For Each Element:** For each element, a_{ij}, in the chosen row or column:
Find the Minor, M_{ij}: The minor M_{ij} is the determinant of the submatrix formed by deleting the i-th row and the j-th column of the original matrix.
Find the Cofactor, C_{ij}: The cofactor C_{ij} is the minor multiplied by a sign factor:

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

The term $(-1)^{i+j}$ gives a checkerboard pattern of signs:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

3.**Calculate the Determinant:** The determinant of the matrix, A, is the sum of the products of the elements in the chosen row or column and their corresponding cofactors.

Expansion along the i-th row:

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

Expansion along the j-th column:

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Both expansions give the same result.

Example (3×3 Matrix):

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Expanding along the first row:
1.a_{11}: $M_{11} = \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$, $C_{11} = +M_{11}$
2.a_{12}: $M_{12} = \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$, $C_{12} = -M_{12}$
3.a_{13}: $M_{13} = \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$, $C_{13} = +M_{13}$
Therefore,

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

32.10.1 Example of Determinant Calculation

Let's calculate the determinant of the matrix:

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Using the rule of Sarrus for 3×3 matrices:

$$\begin{aligned} \det(A) &= 2 \cdot 0 \cdot 1 + 1 \cdot 2 \cdot 3 + 3 \cdot 1 \cdot 1 - (3 \cdot 0 \cdot 0 + 1 \cdot 2 \cdot 2 + 2 \cdot 1 \cdot 1) \\ &= 0 + 6 + 3 - (0 + 4 + 2) \\ &= 9 - 6 = 3 \end{aligned}$$

Thus, the determinant of matrix A is $\det(A) = 3$.
Now consider the following 4×4 matrix:

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

For a 4×4 matrix, we can use Laplace Method along the first row:

$$\begin{aligned} \det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} + a_{14}C_{14} \\ = 2 \cdot \det \begin{pmatrix} 0 & -1 & 3 \\ 3 & 1 & 5 \\ -1 & 0 & 2 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 4 & -1 & 3 \\ -2 & 1 & 5 \\ 1 & 0 & 2 \end{pmatrix} \\ + 3 \cdot \det \begin{pmatrix} 4 & 0 & 3 \\ -2 & 3 & 5 \\ 1 & -1 & 2 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 4 & 0 & -1 \\ -2 & 3 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{aligned}$$

$$\det(A) = 145$$

32.11 Linear Maps as matrices and Their Properties

Let V and W be vector spaces over the field K.
Let $v_1, \ldots, v_n \in V$ and $w_1, \ldots, w_n \in W$. If $(v_1, \ldots, v_n)$ forms a basis of V, then there exists a unique $f \in \operatorname{Hom}(V, W)$ with $f(v_i) = w_i$, $1 \leq i \leq n$. The map f has the following properties:

1. $\operatorname{Im}(f) = \operatorname{span}(f(v_1), \ldots, f(v_n))$.
2. f is injective $\Leftrightarrow$ $w_1, \ldots, w_n$ are linearly independent.

Let V and W be two K-vector spaces, $B_V = (v_1, \ldots, v_n)$ a basis of V and $B_W = (w_1, \ldots, w_m)$ a basis of W, and let $f : V \to W$ be linear. Then there exists a unique matrix $M_{B_V}^{B_W}(f) = (a_{ij}) \in K^{m \times n}$ with

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \forall j = 1, \ldots, n$$

32.12 Example of an exercise

1. Determination of the kernel 2. Determination of the dimension of the kernel 3. Determination of the rank (dimension formula) 4. Determination of the image

Example 4.51: Given

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 - x_2 + x_3 \\ 4x_1 - x_2 + 2x_3 \end{pmatrix}.$$

It should be shown that f is linear, and $\ker(f)$, $\operatorname{Im}(f)$ and their dimensions should be determined.

A direct proof of linearity or by means of Remark 4.12.2 is easily possible. Instead, we give the transformation matrix A. The images of the (canonical) basis vectors are

$$f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, \quad f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

We obtain

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{pmatrix}.$$

But now it must be shown that indeed $f(x) = Ax \quad \forall x \in \mathbb{R}^3$ holds, by, for example, calculating both Ax and $f(x)$ for a general x and showing equality: Here, with $x = (x_1, x_2, x_3)^T$

$$A \cdot x = \begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 4 & -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 - x_2 + x_3 \\ 4x_1 - x_2 + 2x_3 \end{pmatrix}.$$

This obviously corresponds to $f(x)$, so that by Theorem 4.36, the map f is linear. To determine the kernel, one has to solve the system of linear equations

$$\begin{array}{cccc} 2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 4 & -1 & 2 & 0 \end{array}$$

which corresponds to the equation $Ax = f(x) = 0$. Gaussian elimination yields $x_3 = \lambda'$; $x_2 = \frac{2}{3} \cdot \lambda'$; $x_1 = -\frac{1}{3} \cdot \lambda'$, thus with $\lambda = \frac{1}{3}\lambda'$:

$$\ker(f) = \left\{ x = \lambda \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix} \middle| \lambda \in \mathbb{R} \right\}$$

It follows that $\dim(\ker(f)) = 1$ and because of $\dim(V) = 3$ from the dimension formula $\dim(\operatorname{Im}(f)) = 2$. By a corallary, $\operatorname{Im}(f)$ corresponds to the linear span of the columns of the matrix. One chooses consequently $\dim(\operatorname{Im}(f))$ column vectors, e.g., the first ones, and tests if they are linearly independent. In the concrete case, this is obvious, because the second column is not a multiple of the first. It follows therefore

$$\operatorname{Im}(f) = \left\{ x \middle| x = \lambda \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \lambda, \mu \in \mathbb{R} \right\}$$

32.13 Eigenvectors and Eigenvalues

Eigenvector An eigenvector of a square matrix A is a non-zero vector $\vec{v}$ that, when multiplied by A, results in a vector that is a scalar multiple of itself. In other words, the direction of the vector $\mathbf{v}$ remains unchanged (up to scaling) when the linear transformation represented by A is applied to it.

Eigenvalue The scalar multiple, denoted by λ, is called the eigenvalue associated with the eigenvector $\mathbf{v}$. It represents the factor by which the eigenvector is scaled when transformed by the matrix A.

Mathematically, the relationship between a square matrix A, an eigenvector $\vec{v}$, and its corresponding eigenvalue λ is expressed by the following equation:

$$A\vec{v} = \lambda\vec{v}$$

32.13.1 How to find the Eigenvectors and Eigenvalues

To find the eigenvalues and eigenvectors of a square matrix A, we solve the eigenvalue equation:
1.**Form the characteristic equation:** Rewrite the equation $A\mathbf{v} = \lambda\mathbf{v}$ as $(A - \lambda I)\mathbf{v} = \mathbf{0}$, where I is the identity matrix of the same size as A. To have a non-trivial solution for $\mathbf{v}$, the matrix $(A - \lambda I)$ must be singular, which means its determinant must be zero. Thus, we have the characteristic equation:

$$\det(A - \lambda I) = 0$$

2.**Solve for the eigenvalues:** Solve the characteristic equation for λ. The solutions $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of the matrix A.
3.**Find the eigenvectors:** For each eigenvalue λ_i, substitute it back into the equation $(A - \lambda_i I)\mathbf{v} = \mathbf{0}$ and solve for the vector $\mathbf{v}$. The non-zero solutions for $\mathbf{v}$ are the eigenvectors corresponding to the eigenvalue λ_i.

32.13.2 How to diagonalize a matrix

Diagonalizing a matrix involves finding a diagonal matrix that is similar to the given matrix. A square matrix A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. The process of diagonalization is as follows:
1. **Find the eigenvalues and eigenvectors of A.**
2. **Form the matrix P:** Create a matrix P whose columns are the linearly independent eigenvectors of A.
3. **Form the diagonal matrix D:** Create a diagonal matrix D whose diagonal entries are the eigenvalues of A, corresponding to the order of the eigenvectors in P. That is, if the i-th column of P is the eigenvector corresponding to the eigenvalue λ_i, then the i-th diagonal entry of D is λ_i.
4. **Verify the diagonalization:** Check that $P^{-1}AP = D$.

A matrix A is diagonalizable if and only if it has n linearly independent eigenvectors, where n is the size of the matrix.

Example

Consider the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

1. **Find the eigenvalues:** The characteristic equation is

$$\det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0$$

Solving for λ, we get $\lambda_1 = 1$ and $\lambda_2 = 3$.

2. **Find the eigenvectors:** For $\lambda_1 = 1$:

$$(A - I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives $x + y = 0$, so an eigenvector is $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

For $\lambda_2 = 3$:

$$(A-3I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives $-x + y = 0$, so an eigenvector is $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

3. **Diagonalize the matrix:** Let $P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Then $P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

$$P^{-1}AP = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = D$$

Thus, A is diagonalized as $P^{-1}AP = D$.