Development of the Number Systems

Labix

October 28, 2024

Abstract

These notes will act as a development for the number systems from the natural numbers to the complex numbers.

Contents

1 Natural Numbers

1.1 Order of Natural Numbers

Definition 1.1.1: Natural Numebers

The set of natural numbers $\mathbb{N}$, formulated in ZFC set theory via peano axioms, is the set

$$\mathbb{N} = \{0, 1, 2, \dots\}$$

of natural numbers. Addition and multiplication is also defined.

Definition 1.1.2: Order

Let $a, b \in \mathbb{N}$. We say $a < b$ if there exists some $c \in \mathbb{N}$ and $c \neq 0$ so that $a + c = b$. $<$ is a relation in the set $\mathbb{N}$.

Proposition 1.1.3: Trichotomy

Let $a, b \in \mathbb{N}$. Then either $a = b$ or $a < b$ or $b < a$.

Proof. From set theory, we have that $a \in b$ or $a = b$ or $b \in a$, which corresponds to $a < b$, $a = b$, $b < a$ respectively. □

Proposition 1.1.4

Suppose that $a, b, c \in \mathbb{N}$. The relation $<$ has the below properties.

- $a < b$ and $b < c \implies a < c$
- $a < b \implies a + c < b + c$
- $a < b \implies ac < bc$

Proof. We prove the three using the definition of $<$.

- Suppose $a < b$ and $b < c$. There exists some $x, y \in \mathbb{N}$ such that $a + x = b$ and $b + y = c$. Then $a + x + y = c$ thus $a < c$
- Suppose $a < b$. Then $a + c = b$ for some $c \in \mathbb{N}$. Then $b + d = (a + c) + d = (a + d) + c$ thus $a + d < b + d$.
- Suppose $a < b$. Then $a + d = b$ for some $d \in \mathbb{N}$. Then $ac + dc = bc$ thus $ac < bc$

□

Definition 1.1.5: Less Than or Equal

Let $a, b \in \mathbb{N}$. We say $a \leq b$ if either $a < b$ or $a = b$.

Theorem 1.1.6

The relation $\leq$ in the natural numbers are partial order.

Proof. Recall from set theory that a partial order is a relation such that it is reflexive, antisymmetric and transitive.

- Since we have $a \leq a$, $\leq$ is reflexive.
- $a \leq b$ and $b \leq a \implies a = b$ by the trichotomy of natural numbers.
- $a \leq b$ and $b \leq c \implies a \leq c$ from the properties of the relation $<$.

□

Theorem 1.1.7

The set of natural numbers is totally ordered.

Proof. For any two numbers in $\mathbb{N}$, we have the trichotomy, thus we have either $a \leq b$ or $b \leq a$ for all $a, b \in \mathbb{N}$. □

2 Integers

2.1 Introduction to Integers

Lemma 2.1.1

Define a relation R on $\mathbb{N} \times \mathbb{N}$ by $(r, s)R(p, q)$ if $r + q = s + p$. Then R is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.

Definition 2.1.2: Integers

Define the set of integers $\mathbb{Z}$ to be the set of equivalence classes of $\mathbb{N} \times \mathbb{N}$ under R. Each equivalence class $E_{(0,n)}$ is denoted by $-n$.

Definition 2.1.3: Operations on the Integers

We define subtraction of $a, b \in \mathbb{Z}$ to be $a - b$, which is $a + (-b)$.

Proposition 2.1.4

The set of integers is totally ordered.

2.2 Divisibility

Definition 2.2.1: Divisibility

Let $a, b \in \mathbb{Z}$. We define the relation

$$a|b$$

if and only if there exists some $k \in \mathbb{Z}$ such that $b = ak$. We say that a divides b in this case.

Proposition 2.2.2

Let $d, m, n \in \mathbb{Z}$. The relation $|$ has the following properties and thus is a partial order in $\mathbb{N}$.

- Reflexive: $n|n$
- Anti-symmetric: $m|n$ and $n|m \implies m = n$
- Transitive: $d|n$ and $n|m \implies d|m$
- Linear: $d|n$ and $d|m \implies d|(an + bm)$ for any $a, b \in \mathbb{Z}$
- $1|n$
- $n|0$

Proof. We prove antisymmetry and transitivity and leave the others for the reader. Let $m, n, d \in \mathbb{Z}$.

- (Antisymmetry) If $m|n$ and $n|m$ then there exists some $k_1, k_2 \in \mathbb{N}$ such that $n = k_1 m$ and $m = k_2 n$ thus $n = k_1 k_2 n$. Then $k_1 k_2 = 1 \implies k_1 = k_2 = 1$ and $m = n$
- (Linearity) If $d|n$ and $n|m$ then there exists $k_1 k_2 \in \mathbb{N}$ such that $n = k_1 d$ and $m = k_2 n$. Then $m = k_2 k_1 d$ thus $d|m$

□

These properties will come up again and again and will be the foundation of number theory. It is safe to say that number theory is built upon the notion of divisibility.

2.3 The Division Algorithm

This section is dedicated to develop the Euclidean algorithm, a means to find the greatest common divisor. The gcd is a central notion in number theory as well.

Definition 2.3.1: Greatest Common Divisor

Suppose that $m, n \in \mathbb{Z}$. A number $d \in \mathbb{N}$ such that
- $d \geq 0$
- $d|m$ and $d|n$
- $e|a$ and $e|b \implies e|d$

is called the greatest common divisor of m and n, denoted $\gcd(m, n)$.

In contrast to the greatest common divisor, we also have the lowest common multiple. Although they work as a pair, we often see the notion of gcd come up more than lcm.

Definition 2.3.2: Lowest Common Multiple

Suppose that $m, n \in \mathbb{Z}$. A number $l \in \mathbb{N}$ such that
- $l \geq 0$
- $m|l$ and $n|l$
- $m|e$ and $n|e \implies l|e$

is called the lowest common multiple of m and n, denoted $\operatorname{lcm}(m, n)$.

Beware that both of these definitions does not imply the uniqueness of such a number. However, with a little work, we will see that both of them are indeed unique. Readers should think about whether the existence of these numbers is guaranteed as well.

Proposition 2.3.3

Let $m, n \in \mathbb{Z}$. Then the numbers $\gcd(m, n)$ and $\operatorname{lcm}(m, n)$ are unique.

Proof. By the third property of both numbers, we must have if c, d are $\gcd(m, n)/\operatorname{lcm}(m, n)$, then $c|d$ and $d|c$ thus $c = d$ and $\gcd(m, n)/\operatorname{lcm}(m, n)$ is unique. □

We will see more on gcd and lcm when we deal with factorization. For now, we turn our heads to the division algorithm. This algorithm proves to us that upon dividing two integers, as long as they are not divisible by one or the other, you can always guarantee a remainder smaller than the divident.

Theorem 2.3.4: The Division Algorithm

Let $a \in \mathbb{N}$ and $b \in \mathbb{Z}$ with $b \neq 0$. Then there exists unique $q, r \in \mathbb{Z}$ such that

$$b = aq + r$$

with $0 \leq r < a$.

Proof. We prove existence first by considering three cases.
Cases 1: b is divisible by a. If b is divisible by a then there exists $k \in \mathbb{Z}$ such that $b = ka$ thus $k = q$ and $r = 0$.

Case 2: b is positive and a does not divide b. Let

$$S = \{b - ka \in \mathbb{N} | k \in \mathbb{N}\}$$

Then $S \subseteq \mathbb{N}$ thus we can apply the well-ordering principle to S. Let r be the least natural number in S. Then $r \in S$ implies $r = b - ka$ for some $k \in \mathbb{N}$. Thus $b = ka + r$ for some k and r. We show that $r < a$. Suppose for a contradiction that $r \geq a$. Then $u = r - a \in \mathbb{N}$ and

$$b = ka + r \implies b = ka + (u - a) \implies b = (k - 1)a + u$$

thus $u \in S$ and $u < r$, contradicting the fact that r is the least element in S. Thus $r \leq a$. If $r = a$, then

$$b = ka + a \implies b = (k + 1)a$$

which means that $a|b$ which is false in our case. Thus we must have $r < a$.

Case 3: b is negative and a does not divide b. Then apply the exact same argument to the number $-b$ to get $(-b) = ka + r$ and $b = -ka - r$. Let $k' = -k - 1$ and $r' = -r + a$. Then

$$b = -ka - r = k'a + a + r' - a = k'a + r'$$

Since we have $0 \leq r < a$, we have $-a < -r \leq 0$ and $0 < r' \leq a$. Again $r' \neq a$ or else $a|b$ which contradicts our assumption.

We now prove uniqueness. Suppose that $b = aq_1 + r_1$ and $b = aq_2 + r_2$. Then $r_1 - r_2 = a(q_2 - q_1)$. We know that $-a < r_1 - r_2 < a$ thus $-a < a(q_2 - q_1) < a$ and $-1 < q_2 - q_1 < 1$ which is impossible for integers q_1, q_2 unless $q_1 = q_2$. If $q_1 = q_2$ then $r_1 = r_2$ and we are done. □

The division algorithm does not require b to be larger than a. In fact, if a is larger than b, then the division algorithm simply gives a itself as the remainder. Before we reach our conclusion, we need one more proposition.

Proposition 2.3.5

Suppose that $m \geq n > 0$ are natural numbers with $m = qn + r$ for some $q, r \in \mathbb{N}$. Then

$$\gcd(m, n) = \gcd(n, r)$$

Proof. Suppose that $d = \gcd(m, n)$. Then we know that $d < n$ from definition. We want to show that d satisfies the three results of a gcd but in terms of n and r. Since $d|n$ and $d|m$, by linearity we must have $d|r$.

Now suppose for a contradiction that there exists e such that e is a common divisor of n and r and $e > d$. Then $e|n$ and $e|r$ by definition thus $e|m$ by linearity. $e|m$ and $e|n$ implies that e is a larger common divisor of m and n than d. However this is not possible since d is assumed to be the largest among the common divisors. This is a contradiction thus $d = \gcd(n, r)$ and we are done. □

Theorem 2.3.6: Euclid's Algorithm

Suppose that $m \geq n > 0$ are natural numbers. We have the following inequalities.

$$m = nq_1 + r_1 \text{ with } 0 < r_1 < n$$

$$n = r_1q_2 + r_2 \text{ with } 0 < r_2 < n$$

$$r_1 = r_2q_3 + r_3 \text{ with } 0 < r_3 < n$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$r_{k-2} = r_{k-1}q_k + r_k \text{ with } 0 < r_k < r_{k-1}$$

$$r_{k-1} = r_kq_{k-1}$$

From this, we have $r_k|r_{k-1}$, $r_k|r_{k-2} \ldots r_k|n$ and $r_k|m$.

Proof. The first part of the results is due to the repeated use of the division algorithm. For the second part, we have

$$\gcd(m, n) = \gcd(n, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{k-1}, r_k) = r_k$$

and we are done. □

Lemma 2.3.7: Bezout's Lemma

Let $a, b \in \mathbb{Z}$ such that they are not both 0. Then there exists $x, y \in \mathbb{Z}$ such that

$$ax + by = \gcd(a, b)$$

Proof. Reconstruct x and y using the Euclidean Algorithm. This is possible since $\gcd(m, n) = r_k$ and every $r_1, \ldots, r_{k-1}$ has a factor of r_k in it. □

Lemma 2.3.8

Let $a, b \in \mathbb{Z}$ such that they are not both 0. Then the equation

$$ax + by = \gcd(a, b)$$

has an infinite number of integer solutions.

Proof. Using Bezout's Lemma, we conclude that (x_0, y_0) is a solution to the equation. But then

$$(x_0 - bt, y + at)$$

are also solutions for $t \in \mathbb{Z}$ since

$$a(x_0 - bt) + b(y + at) = ax + by = \gcd(a, b)$$

□

Corollary 2.3.9

Let $a, b \in \mathbb{Z}$ such that they are not both 0 and $d \in \mathbb{Z}$. Then d divides a and b if and only if $d | \gcd(a, b)$.

2.4 Unique Factorization

Definition 2.4.1: Prime Numbers

We say that $n \in \mathbb{N}$ is a prime number if and only if it has exactly two factors, which is 1 and n. Else n is composite.

Lemma 2.4.2

Every integer is divisible by a prime.

Lemma 2.4.3

Every integer $n > 1$ can be written as a product of primes.

Theorem 2.4.4

There is an infinite number of primes.

Proposition 2.4.5: Euclid's Lemma

Suppose that $p, m, n \in \mathbb{N}$, with p prime and $m, n > 1$. Suppose that $p|mn$. Then p divides at least one of m or n.

Proposition 2.4.6

Suppose that p is a prime such that $p|a_1a_2 \cdots a_k$. Then $p|a_i$ for some $i \in \{1, 2, \ldots, k\}$

Theorem 2.4.7: Fundamental Theorem of Arithmetic

Suppose that $n \neq 0$ is a natural number. Then there exists exactly one prime factorization for every n, meaning that the decomposition

$$n = \prod_{k=1}^{n} p_k^{s_k}$$

where p_k is prime exists and is unique.

Theorem 2.4.8

Suppose that $m, n \in \mathbb{N}$. Suppose that

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$$

$$n = p_1^{\beta_1} p_2^{\beta_2} \cdots p_q^{\beta_q}$$

with $p_1 = 2$, $p_2 = 3$, $p_3 = 5 \ldots$. Without loss of generality $r \leq q$. Then

$$\gcd(m, n) = p_1^{\min(\alpha_1, \beta_1)} p_2^{\min(\alpha_2, \beta_2)} \cdots p_q^{\min(\alpha_q, \beta_q)}$$

$$\operatorname{lcm}(m, n) = p_1^{\max(\alpha_1, \beta_1)} p_2^{\max(\alpha_2, \beta_2)} \cdots p_q^{\max(\alpha_q, \beta_q)}$$

Proposition 2.4.9

Suppose that m and n are natural numbers. Then

$$\gcd(m, n) \times \operatorname{lcm}(m, n) = m \times n$$

Proof. Since $\min\{a, b\} \cdot \max\{a, b\} = ab$, from the above theorem, we have that $\gcd(m, n) \times \operatorname{lcm}(m, n) = m \times n$ and we are done. □

3 Rational Numbers

3.1 Introduction to Rationals

Lemma 3.1.1

Define a relation R on $\mathbb{Z} \times \mathbb{Z}$ by $(a, b)R(c, d)$ if $ad = bc$. Then R is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.

Definition 3.1.2: Rational Numbers

Define the set of rational numbers $\mathbb{Q}$ to be the set of equivalence classes of $\mathbb{Z} \times \mathbb{Z}$ under R. Each equivalence class $E_{(a,b)}$ is denoted by $\frac{a}{b}$.

Definition 3.1.3: Operations on the Rationals

We define the four basic operators on rationals as follows.

- $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$
- $\frac{a}{b} - \frac{c}{d} = \frac{ad-bc}{bd}$
- $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$
- $\frac{a}{b} / \frac{c}{d} = \frac{ad}{bc}$

Proposition 3.1.4

The additive inverse of $\frac{a}{b}$ is $-\frac{a}{b}$. Its multiplicative inverse is $\frac{b}{a}$

Definition 3.1.5: Reduced Form

Suppose that $a \in \mathbb{Q}$. $a = \frac{r}{s}$ is a reduced form if

- $s > 0$
- $\gcd(r, s) = 1$

Theorem 3.1.6

For every $x \in \mathbb{Q}$, x has exactly one reduced form.

4 Real Numbers

4.1 Dedekind Cuts

Definition 4.1.1: Dedekind Cuts

A dedekind cut is a partition of $\mathbb{Q}$ into two subsets A, B such that
- A is non-empty
- $A \neq \mathbb{Q}$
- $x, y \in \mathbb{Q}$ and $x < y$ and $y \in A \implies x \in A$
- $x \in A \implies$ there exists a $y \in A$ such that $y > x$

We use A to denote this cut since B is determined by A

Definition 4.1.2: Order

If A, B are dedekind cuts then we say that $A < B$ if and only if $A \subset B$.

Definition 4.1.3: Real Numbers

We define the set of real numbers $\mathbb{R}$ as the set of all dedekind cuts of $\mathbb{Q}$.

Proposition 4.1.4

Define addition, subtraction, multiplication, division as follows. Then the resulting set is also a dedekind cut.
- $A + B = \{a + b : a \in A \text{ and } b \in B\}$
- $A - B = \{a - b : a \in A \text{ and } b \in \mathbb{Q} \setminus B\}$
- $A \times B = \{a \times b : a \in A \text{ and } b \in B\}$ if $A, B \geq 0$ or $A, B \leq 0$. If at one of $A, B < 0$ then use the identity $-(-A \times B)$ or $-(A \times -B)$ depending on whether $A < 0$ or $B < 0$ respectively.
- $A/B = \{a/b : a \in A \text{ and } b \in \mathbb{Q} \setminus B\}$ if $A, B \geq 0$ or $A, B \leq 0$. Use the similar approach as multiplication when one of $A, B < 0$.

Proposition 4.1.5

The set of all rational numbers $\mathbb{Q}$ is a subset of the real numbers $\mathbb{R}$.

Proof. We define $A = \{x \in \mathbb{Q} : x < q\}$ for every $q \in \mathbb{Q}$. Thus the set A satisfies a dedekind cut. □

Definition 4.1.6: Irrational Numbers

Any dedekind cut which is not a rational number is called an irrational number.

Theorem 4.1.7

There exists an irrational number.

Proof. We want to show that there is an irrational number represented by a dedekind cut such that its square is 2. Consider the set $A = \{x \in \mathbb{Q} : x < 0 \text{ or } x^2 < 2\}$. A is non-empty since $0 \in A$. $A \neq \mathbb{Q}$ since $3 \notin A$. Suppose $p \in A$. We then need to show that $q \in A$ whenever $q < p$. When $0 \leq q < p$, we have $0 \leq q^2 < p^2$ from ordering of rationals. When $q < 0$, then $q \in A$ by definition of A. Thus this is true. Now we need to show that there is always a rational q larger than p which is in A. Choose $q = \frac{2p+2}{p+2}$, then $p < q$ and $q^2 < 2$. Thus A is a dedekind cut.

Now consider $A \times A$. We have $A \times A \leq 2$ since for all $x, y \in A$, we have $x^2 < 2$ and $y^2 < 2$ and thus $xy < 2$ whenever $xy \geq 0$. Thus the set
$A \times A = \{r \in \mathbb{Q} : r < 0 \text{ or } r = xy \text{ for some } x, y \in A \text{ and } x, y > 0\}$ is less than or equal to 2. We know that $A \times A$ is a dedekind cut. But we want to know if $A \times A$ represents the number 2. Suppose that $u \in A \times A$. Then we know that from A, there exists a number $v \in A$ such that $u < v^2 < 2$. And this applies for every u. Then we know that $A \times A = 2$ since $A \times A = \{x \in \mathbb{Q} : x < 2\}$, which is our definition of rational numbers with dedekind cut.

We have proved that there exists a dedekind cut such that its square is 2. But is that dedekind cut irrational? We now represent A with $\sqrt{2}$. Suppose that $\sqrt{2}$ is rational. Then we can write it is as $\frac{m}{n}$ in reduced form. Then we have $2n^2 = m^2$. Then $2|m^2$ thus $2|m$. Let $m = 2k$ for some $k \in \mathbb{N}$. Then $2k^2 = n^2$ which similarly implies that $2|n$. This contradicts the fact that $\frac{m}{n}$ is in reduced form, thus $\sqrt{2}$ is in fact not rational, and is an irrational number. □

Proposition 4.1.8

Suppose that A, B, C are dedekind cuts.
- (O1) $A < B$ or $A = B$ or $B < A$
- (O2) $A < B$ and $A < C \implies A < C$
- (O3) $A < B \implies A + C < B + C$
- (O4) $A < B$ and $z > 0 \implies AC < BC$

Proposition 4.1.9

Let x, y, $z \in \mathbb{R}$.
- (A1) $x + y \in \mathbb{R}$
- (A2) $(x + y) + z = x + (y + z)$
- (A3) $\exists 0 \in \mathbb{R}$ such that $x + 0 = 0 = 0 + x$
- (A4) $x + y = y + x$
- (A5) $\exists (-x) \in \mathbb{R}$ such that $x + (-x) = 0 = (-x) + x$
- (M1) $xy \in \mathbb{R}$
- (M2) $(xy)z = x(yz)$
- (M3) $\exists 1 \in \mathbb{R}$ such that $x \cdot 1 = x = 1 \cdot x$
- (M4) $xy = yx$
- (M5) $\exists (x^{-1}) \in \mathbb{R}$ such that $x(x^{-1}) = 1 = (x^{-1})x$
- (D1) $x(y + z) = xy + xz$
- (O1) $x < y$ or $x = y$ or $y < x$
- (O2) $x < y$ and $y < z \implies x < z$
- (O3) $x < y \implies x + z < y + z$
- (O4) $x < y$ and $z > 0 \implies xz < yz$

The absolute value is an important function when it comes to defining useful concepts such as distances in the field of real.

Definition 4.1.10

[The Absolute Value] The absolute value of a real number x is defined by

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

The absolute value has some properties that are extremely useful in certain circumstances, notably number 4 and 5.

Proposition 4.1.11

The absolute Value has the folowing properties

1. $|x| \geq 0$
2. $|xy| = |x||y|$
3. $|\frac{x}{y}| = \frac{|x|}{|y|}$
4. $|x+y| \leq |x| + |y|$
5. $||x| - |y|| \leq |x-y|$

Proof. I left out the proofs of (2) and (3) since they are simplay obtained via case by case analysis.

1. When $x \geq 0$ we have $|x| = x \geq 0$. When $x < 0$ we have $|x| = -x > 0$
2. We start by squaring the left hand side of the inequality.

$$\begin{aligned} |x+y|^2 &= (x+y)^2 \\ &= x^2 + 2xy + y^2 \\ &\leq |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2 \end{aligned}$$

   Since the both sides of the inequality is non-negative, we can take the square root on both sides, thus obtaining $|x+y| \leq |x| + |y|$.

3. Choose x to be $x - y$ in (4) and we obtain $|x| - |y| \leq |x-y|$. Similarly, choosing y to be $y - x$ in (4), we find that $|y| - |x| \leq |y-x| = |x-y|$. Thus we have $||x| - |y|| \leq |x-y|$.

□

4.2 The Binomial Theorem

Definition 4.2.1: The Binomial Coefficient

Let $n, r \in \mathbb{N}$ with $n > 0$. We define the binomial coefficient $\binom{n}{r}$ to mean the number $\frac{n!}{r!(n-r)!}$ when $r \leq n$. When $r > m$ then $\binom{n}{r} = 0$.

Proposition 4.2.2

Let $n, r \in \mathbb{N}$ with $0 < r < n$, we have $\binom{n}{r} = \binom{n}{n-r}$.

Proposition 4.2.3

Let $n, r \in \mathbb{N}$ with $0 < r < n$, we have $\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r}$.

Theorem 4.2.4: The Binomial Theorem

Suppose $a, b \in \mathbb{R}$. Then

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

Theorem 4.2.5

[Vandermonde's Theorem] Suppose that $a, b, n \in \mathbb{N}$. Then

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

5 Complex Numbers

5.1 Introduction to Complex Numbers

Definition 5.1.1: Complex Numbers

Define the number $i = \sqrt{-1}$. Define $z = a + bi$ to be a complex number. Every complex number is uniquely determined by an ordered pair $(a, b) \in \mathbb{R}^2$. $a = \text{Re}(z)$ is called the real part of z. $b = \text{Im}(z)$ is called the imaginary part of z. The set of all complex numbers is denoted $\mathbb{C}$.

Definition 5.1.2: Equality in Complex Numbers

We define the relation equality in $\mathbb{C}$ as $z_1 = z_2$ with $z_1 = a + bi$ and $z_2 = c + di$ if and only if $a = c$ and $b = d$.

Definition 5.1.3: Addition and Multiplication

Let $z = a + bi$, $w = c + di$. We define the $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ in $\mathbb{C}$ as follows.
- $z + w = (a + c) + (b + d)i$
- $zw = (ac - bd) + (ad + bc)i$

Proposition 5.1.4

The set of complex numbers $(\mathbb{C}, +, \cdot)$ is a field. In particular, let $x, y, z \in \mathbb{C}$. Then

$(\mathbb{C}, +)$ is a commutative group.
- $x + y \in \mathbb{C}$
- $(x + y) + z = x + (y + z)$
- $\exists 0 \in \mathbb{C}$ such that $x + 0 = 0 = 0 + x$
- $\exists (-x) \in \mathbb{C}$ such that $x + (-x) = 0 = (-x) + x$
- $x + y = y + x$

$(\mathbb{C}/\{0\}, \cdot)$ is a commutative group.
- $xy \in \mathbb{C}$
- $(xy)z = x(yz)$
- $\exists 1 \in \mathbb{C}$ such that $x \cdot 1 = x = 1 \cdot x$
- $\exists (x^{-1}) \in \mathbb{C}$ such that $x(x^{-1}) = 1 = (x^{-1})x$ when $x \neq 0$
- $xy = yx$

The distributive laws hold.
- $x(y + z) = xy + xz$
- $(x + y)z = xz + yz$

5.2 Conjugates

Definition 5.2.1: Conjugation

For every complex number $z = a + bi$ there exists a conjugate $\bar{z} = a - bi$

Proposition 5.2.2

Suppose that $z, w \in \mathbb{C}$.
- $\overline{\overline{z}} = z$
- $\overline{z + w} = \overline{z} + \overline{w}$
- $\overline{zw} = \overline{z}\,\overline{w}$

5.3 Modulus and Argument

Definition 5.3.1: Modulus

Let $z \in \mathbb{C}$. Define the modulus of $z = a + bi$ to be

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

Proposition 5.3.2

Suppose that $z, w \in \mathbb{C}$. Then the following are true for the modulus.

- $|z|^2 = |z||\bar{z}|$
- $|\bar{z}| = |z|$
- $|zw| = |z||w|$
- $z\bar{z} = |z|^2$
- $|z + w| \leq |z| + |w|$
- $|z - w| = ||z| - |w||$

Definition 5.3.3: Argument

Let $z = a + bi$. Define the argument of z to be

$$\arg(z) = \{\theta \in \mathbb{R} | z = \cos(\theta) + i\sin(\theta)\}$$

The principal argument of z is defined to be the $\theta \in \arg(z)$ such that $-\pi < \theta \leq \pi$, denoted $\text{Arg}(z)$.

Proposition 5.3.4

Suppose that $z, w \in \mathbb{C}$. Then the following are true for the argument of a complex number.

- $\arg(z) = \{\text{Arg}(z) + 2\pi k | k \in \mathbb{Z}\}$
- $\arg(zw) = \arg(z) + \arg(w) = \{\theta + \phi | \theta \in \arg(z), \phi \in \arg(w)\}$
- $\arg(\bar{z}) = -\arg(z)$

Definition 5.3.5: Polar Form

Using the modulus and the argument, a complex number can be uniquely determined by $|z|$ and $\arg z$. It can be written as $z = r(\cos(\theta) + i\sin(\theta))$ where $r = |z|$ and $\cos(\theta) = \frac{a}{a^2+b^2}$ and $\sin(\theta) = \frac{b}{a^2+b^2}$ with $-\pi < \theta \leq \pi$ This is the polar form of a complex number.

Proposition 5.3.6

Suppose that $z = r(\cos(\theta) + i\sin(\theta))$ and $w = s(\cos(\phi) + i\sin(\phi))$.

- $zw = rs(\cos(\theta + \phi) + i\sin(\theta + \phi))$
- $\frac{1}{z} = \frac{1}{r}(\cos(-\theta) + i\sin(-\theta))$
- $\frac{z}{w} = \frac{r}{s}(\cos(\theta - \phi) + i\sin(\theta - \phi))$
- $\bar{z} = r(\cos(-\theta) + i\sin(-\theta))$

Theorem 5.3.7: De Moivre's Theorem

Suppose that $r \in \mathbb{Q}$. Then $(\cos\theta + i\sin\theta)^r = \cos(r\theta) + i\sin(r\theta)$

5.4 Roots of Complex Numbers

Definition 5.4.1: nth Roots

Let $w \neq 0$ be a complex number and n a positive integer. A number z is called the nth root of w if and only if $z^n = w$.

Theorem 5.4.2: Roots of Unity

Suppose that $z = re^{i\theta}$. Then the nth roots of z are

$$r^{\frac{1}{n}}\left[\cos\left(\frac{(\theta+2\pi k)i}{n}\right)+i\sin\left(\frac{(\theta+2\pi k)i}{n}\right)\right]$$

where $k = 0, 1, \ldots, n-1$.

6 Algebraic Inequalities

Theorem 6.0.1: Bernoulli's Inequality

For all $x \geq -1$ and $n \in \mathbb{N}$,

$$(1+x)^n \geq 1 + nx$$

Proof. We prove the inequality by induction on n. In the case of $n = 1$, we have $1 + x \geq 1 + x$, which is true for all x. Now suppose that the inequality works for some $n \in \mathbb{N}$. We have

$$\begin{aligned}(1+x)^{n+1} &\geq (1+x)(1+nx) && \text{(Induction Hypothesis and } x \geq -1)\\ &= 1 + (n+1)x + nx^2 \\ &\geq 1 + (n+1)x && \text{(since } x^2 \geq 0)\end{aligned}$$

Thus we have the Bernoulli's Inequality by the principle of mathematical induction. □

Theorem 6.0.2: Weierstrass' Inequality

Let $a_1, \ldots, a_n$ be positive numbers. Then when $n \geq 2$,

$$(1+a_1)\ldots(1+a_n) > 1 + a_1 + \cdots + a_n$$

Theorem 6.0.3: AMGM

Let $a_1, \ldots, a_n$ be positive numbers. Then

$$\frac{a_1 + \cdots + a_n}{n} \geq (a_1 a_2 \ldots a_n)^{\frac{1}{n}}$$

Theorem 6.0.4: Cauchy-Schwarz Inequality

Let $x_1, \ldots, x_n, y_1 \ldots, y_n \in \mathbb{R}$. Then

$$\left(\sum_{k=1}^{n} x_k y_k\right) \leq \left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right)$$

Theorem 6.0.5: Tchbychef's Inequality

Let $x_1, \ldots, x_n, y_1 \ldots, y_n \in \mathbb{R}$ such that $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$. Then

$$n\left(\sum_{k=1}^{n} x_k y_k\right) \geq \left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{n} y_k\right)$$

7 Basics of Matrices

7.1 Matrices and its Operations

Definition 7.1.1: Matrix

A rectangular array of $m \times n$ real numbers, called the elements, or entries,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called an $m \times n$ matrix over $\mathbb{R}$. For $i = 1, \ldots, m$, let

$$r_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdot & a_{in} \end{pmatrix}$$

and for $j = 1, \ldots, n$, let

$$c_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

then r_i is called the ith row of A and c_j is called the jth row of A. The element of A at the intersection of the ith row and jth column is called the (i, j)th entry of A. The set of all $m \times n$ matrices over $\mathbb{R}$ is denoted by $M_{m\times n}(\mathbb{R})$. We sometimes denote A as $(a_{i,j})_{m\times n}$

Definition 7.1.2: Matrix Addition

Let A, B be $m \times n$ matrices. We define the binary operation $+ : M_{m\times n}(\mathbb{R}) \times M_{m\times n}(\mathbb{R}) \to M_{m\times n}(\mathbb{R})$ to be

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Proposition 7.1.3

Let $A, B, C \in M_{m\times n}(\mathbb{R})$. Then
- $(A + B) + C = A + (B + C)$
- $(A + B) = (B + A)$
- $A + 0 = 0 + A = A$
- There exists a unique $M \in M_{m\times n}(\mathbb{R})$ such that $A + M = M + A = 0$

Proof. Addition is associative, commutative and has an identity in $\mathbb{R}$. □

Definition 7.1.4: Scalar Multiplication

Let $A = (a_{i,j})_{m\times n}$ and $\lambda \in \mathbb{R}$. We define the scalar multiple $\cdot : \mathbb{R} \times M_{m\times n}(\mathbb{R}) \to M_{m\times n}(\mathbb{R})$ as $\lambda A = (\lambda a_{i,j})_{m\times n}$.

Proposition 7.1.5

Let $A, B \in M_{m\times n}(\mathbb{R})$ and $\lambda, \mu \in \mathbb{R}$. Then
- $(\lambda\mu)A = \lambda(\mu A)$

- $\lambda(A+B) = \lambda A + \lambda B$
- $(\lambda + \mu)A = \lambda A + \mu A$

Proof. Simple proof by using the definition of scalar multiplication directly. □

Definition 7.1.6: Matrix Multiplication

Let $A \in M_{m\times p}(\mathbb{R})$, $B \in M_{p\times n}(\mathbb{R})$. We define matrix multiplication as $\cdot : M_{m\times p}(\mathbb{R}) \times M_{p\times n}(\mathbb{R}) \to M_{m\times n}(\mathbb{R})$ where

$$A \cdot B = (c_{i,j})_{m\times n}$$

with

$$c_{i,j} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

Proposition 7.1.7

Let A, B, C be matrices over R, with matrix multiplication assumed possible below.

- $(AB)C = A(BC)$
- $(A+B)C = AB + AC$
- $A(B+C) = AB + AC$

Proof. Once again an easy proof exploiting the definition of matrix multiplication □

Definition 7.1.8: Invertible Matrices

A square matrix A is said to be invertible or non-singular if there is a square matrix B such that $AB = BA = I$. In this case B is the inverse of A. A matrix that is not-invertible is a singular matrix.

Theorem 7.1.9

If A is invertible, then it has a unique inverse.

Proof. Suppose that A is invertible. Then there exists B such that $AB = I$. Thus the inverse is exactly B. Suppose that C is also an inverse of A. Then $AB = I = AC$. Thus $BAB = BAC$ implies $B = C$. □

Definition 7.1.10: Upper Triangular Matrices

A matrix is called upper triangular if all of its entries below the main diagonal are zero.

Definition 7.1.11: Diagonal Matrices

A square matrix is said to be a diagonal matrix if $d_{ij} = 0$ whenever $i \neq j$

Definition 7.1.12: Transpose

Let $A = (a_{ij})_{m\times n}$. The transpose of A is the $n \times m$ matrix denoted by A^T obtained by interchanging the row and columns of A, that is, $A^T = (a_{ji})_{n\times m}$

7.2 Elementary Matrices

Definition 7.2.1

[Recombine Matrix] The $n \times n$ recombine matrix is given by

$$R_{i,j,a} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & a & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

where the diagonals are all 1 and other elements that are not shown is 0.

Definition 7.2.2

[Scale Matrix] The $n \times n$ scale matrix is given by

$$R_i(a) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & a & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

where the diagonals are all 1 except for the i, ith element and other elements that are not shown is 0.

Definition 7.2.3

[Transposition Matrix] The $n \times n$ transposition matrix is given by

$$R_{i,j} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & 1 & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix}$$

where the diagonals are all 1 except for the i, ith and j, jth element and other elements that are not shown is 0.

Theorem 7.2.4

The inverse of the three elementary matrices exists and are also their respective elementary matrices.

Proof. Note that

$$R_{i,j,a}R_{i,j,-a} = I$$

and

$$R_i(a)R_i(a^{-1}) = I$$

and

$$R_{i,j}R_{i,j} = I$$

□

7.3 Row Operations

Definition 7.3.1

[Row Operations] Let $A_{m\times n}$ with rows $r_1, \ldots, r_m$. There are three types of row operations available on A.

- For some $i \neq j$, add a multiple of r_j to r_i
- Interchange r_i and r_j
- Multiply a row by a non-zero scalar

Theorem 7.3.2

The three row operations are in fact matrix left multiplications of the elementary matrices. Namely, the recombine matrix, the scale matrix and the transposition matrix corresponds to the above row operations respectively.

Proof. It suffices to check for yourselves that each row operation corresponds to an elementary matrix by simple matrix multiplication. □

Theorem 7.3.3

Let $A_{m\times n}$ be a matrix and $P_{m\times m}$ a product of $m \times m$ elementary matrices. Then the equations $Ax = 0_m$ and $(PA)x = 0_m$ has the same solution set.

Proof. Since P is invertible, we have that $Ax = 0_m$ if and only if $(PA)x = 0_m$. Thus they have the same solutions. □

Definition 7.3.4

[Upper Echelon Form] A matrix satisfying the below properties is said to be in upper echelon form.

- All zero rows are below all non-zero rows.
- The first non-zero entry of a row is to the right of the first non-zero entry of the row above.
- The first non-zero entry of every row is 1

Definition 7.3.5

[Row-reduced Form] We say that a matrix in upper echelon form is in row-reduced form if above and below every first non-zero entry of a row, all entries are 0.

Definition 7.3.6

[Reduction Procedure] Let $A = (a_{ij})_{m\times n}$. Start with a_{11}.

1. If a_{ij} and all entries below are 0, move on pivot to the right to $a_{i,j+1}$ and repeat step 1, or terminate if $j = n$
2. If $a_{ij} = 0$ but $(a_k j) \neq 0$ for some $k > i$, apply $R_{i,j}$

3. If $a_{ij} \neq 1$, apply $R_{a_{i,j}^{-1}}$
4. If for any $k \neq i$, $a_{kj} \neq 0$, apply $R_{k,i,-a_{kj}}$
5. If $i = m$ or $j = n$ then terminate, else move pivot to $i+1, j+1$ and go back to step 1.

Theorem 7.3.7

Every matrix is row equivalent to one and only one matrix in row reduced form.

Proof. The above reduction procedure provides the existence of a row reduced form. We prove by induction on the number of columns of A the uniqueness of the matrix. Let A be an $m - \times n$ matrix. When $n = 1$, there are only two possible row reduced forms. $a_{i1} = 0$ for all $i > 1$, and a_{11} i either 0 or 1. We have $a_{11} = 0$ if and only if the original matrix is the 0 matrix. So any non-zero matrix $m \times 1$ has only one possible row reduced form.

Now suppose that $n > 1$, and the theorem is true for smaller n. Let A' be the $m \times (n-1)$ matrix obtained by deleting the last column from A. This means that $A = (A'|\mathbf{k})$. By induction the row reduced form of A' is unique. Now if any sequence of row operations that places A into row reduced form, it also places A' into row reduced form, so if B and C are two row reduced form of A, they differ only by the last column. Now since row operations conserve the set of solutions to $A\mathbf{x} = 0$, we have that if $\mathbf{c}$ is the solution, then $B\mathbf{c} = 0$ and $C\mathbf{c} = 0$ and $(B - C)vbc = 0$. Since the first $n - 1$ columns of B and C are the same, if $B \neq C$, we have $B - C$ is of the form $(\mathbf{0}_{m,n-1}|\mathbf{u})$ and $\mathbf{u} \neq 0$. Since the last column is nonzero, there must be at least one element in $\mathbf{u}$ nonzero, meaning there is at least one row in the form $(\mathbf{0}|p)$ for some $p \neq 0$.

Now $(B - C)\mathbf{c} = 0 \implies (\mathbf{0}|\mathbf{u})\mathbf{c} = 0$ and thus $(\mathbf{0}|p) \cdot \mathbf{c} = 0$ and $pc_n = 0$. If $p \neq 0$ then naturally $c_n = 0$ by cancellation law. This implies that there is a leading one in the nth column of B. Because if this is not true, any choice of $\alpha \in \mathbb{R}$ and setting $c_n = \alpha$ could lead to a solution to $B\mathbf{c} = 0$, contradicting the fact that $c_n = 0$.

Now the leading one in the nth column must occur in the first zero row of A' in both B and C. And since other every other entry in the column of a leading one is zero, we finally have that the nth column of B and C is equal. Thus $B = C$. □

7.4 Determinants

We borrow notations from group theory.

Definition 7.4.1

[Odd Even Permutations] A permutation is said to be even, and to have sign $+1$ if ϕ is a composition of an even number of transpositions, and ϕ is said to be odd, and to have sign -1 if ϕ is a composition of an odd number of transpositions.

Definition 7.4.2

[Determinants] The determinant of a $n \times n$ matrix $A = (a_{ij})$ is the scalar quantity

$$\det(A) = \sum_{\phi \in S_n} \text{sign}(\phi) a_{1\phi(1)} a_{2\phi(2)} \dots a_{n\phi(n)}$$

Lemma 7.4.3

$\det(I_n) = 1$.

Proof.

$$\begin{aligned}\det(I_n) &= \sum_{\phi \in S_n} \operatorname{sign}(\phi) a_{1\phi(1)} a_{2\phi(2)} \dots a_{n\phi(n)} \\ &= a_{11} a_{22} \dots a_{nn} \\ &= 1\end{aligned}$$

□

Proposition 7.4.4

If A has two equal rows then $\det(A) = 0$

Proposition 7.4.5

Applying elementary row operations does the following to the determinant of a matrix.

- $\det(R_{i,j,a}A) = \det(A)$
- $\det(R_i(a)A) = a\det(A)$
- $\det(R_{i,j}) = -\det(A)$

Proposition 7.4.6

If $A = (a_{ij})_{n\times n}$ is upper triangular, then

$$\det(A) = a_{11} a_{22} \dots a_{nn}$$

Proposition 7.4.7

Let $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$. Then

- $\det\left(A^T\right) = \det(A)$
- $\det(AB) = \det(A)\det(B)$

7.5 Inverses of Matrices

Definition 7.5.1

[Minor] Let $A \in M_{n\times n}(\mathbb{R})$. The minor M_{ij} of the element a_{ij} of A is the determinant of the submatrix obtained by deleting the ith row and the jth column of A.

$$M_{ij} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,n} \end{vmatrix}$$

Definition 7.5.2

[Cofactor] The cofactor of a_{ij} in $A \in M_{n\times n}(\mathbb{R})$ is defined as

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

Definition 7.5.3

[Adjoint] Let $A \in M_{n\times n}(\mathbb{R})$. The adjoint of A is defined as

$$\text{adj}(A) = (A_{ij})^T_{m\times n}$$

Theorem 7.5.4

Let $A \in M_{n\times n}(\mathbb{R})$. Then

$$A(\text{adj}(A)) = (\text{adj}(A))A = \det(A)I$$

Proposition 7.5.5

Let $A \in M_{n\times n}(\mathbb{R})$. Then $\det(A) = \sum_{k=1}^{n} a_{ik}A_{ik}$ for $i = 1, 2, 3$ and $\det(A) = \sum_{k=1}^{n} a_{kj}A_{kj}$ for $j = 1, 2, 3$

Theorem 7.5.6

[Inverse of a Matrix] Let $A \in M_{n\times n}(\mathbb{R})$. Then A is invertible if and only if $\det(A) \neq 0$. In this case,

$$A^{-1} = \frac{1}{det(A)}\,\text{adj}(A)$$

Proposition 7.5.7

Let $A \in M_{n\times n}(\mathbb{R})$. If A is invertible then $\det\left(A^{-1}\right) = \frac{1}{det(A)}$

Theorem 7.5.8

Suppose that A, B are invertible.

- AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$
- A^n is also invertible where $n \in \mathbb{N}$ and $(A^n)^{-1} = (A^{-1})^n$
- A^{-1} is also invertible and $(A^{-1})^{-1} = A$
- A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$

7.6 System of Linear Equations

Definition 7.6.1

[System of Linear Equations] We say that

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

is a system of linear equations in n unknowns $x_1, \ldots, x_n$.

7.7 System of Linear Equations

Definition 7.7.1

olution Sets A solution of a system of equations is a list of numbers $x_1, \ldots, x_n$ that makes all of the equations true simultaneously. The solution set of a system of equations is the collection of all solutions.

Definition 7.7.2: Consistent Systems

We say that the system of linear equations is consistent if it has at least one solution. We say that the system of linear equations is inconsistent if it has no solutions.

Definition 7.7.3: Homogenous Systems

We say that the system of linear equations is a homogenous system if all the constant coefficients are 0.

Definition 7.7.4: Representation of System of Linear Equations

The matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

and

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

can represent a system of linear equations by $Ax = B$ with $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. A is said to be the coefficient matrix. X is said to be a solution if it satisfies the above m equations simultaneously.