Lecture 1

Divisibility of integers. Greatest common divisor. Euklid's algorithm.
Relativety prime numbers. Prime numbers. The fundamental theorem of arithmetic.

Divisisbility of integers.

The set { 1, 2, 3, . . . } is called the set of *natural numbers*, and will be denoted by **N.** The set {. . . ,-3, -2, -1, 0, 1, 2, 3, . . . } is called the set of *integers*, and will be denoted by **Z**.

Definition 1.1
An integer a is called a multiple of an integer $b \neq 0$, written $a \vdots b$, if there exist some integer c that $a = b \cdot c$. In this case we also say that b is a divisor of a.

Corollary 1.2

1. If $a, b, c \in \mathbf{Z}$ and $a \vdots c$, $b \vdots c$ then $(a \pm b) \vdots c$.
2. If $a, b, c \in \mathbf{Z}$ and $a \vdots c$, then $ab \vdots c$.
3. If $a \in \mathbf{Z}, a \neq 0$, then $0 \vdots a$.
4. If $a \in \mathbf{Z}$, then $a \vdots 1$.
5. If $a \in \mathbf{Z}, 1 \vdots a$, then $a = \pm 1$.
6. If $a \vdots b$, $b \vdots a$, then $a = \pm b$.
7. If $a \vdots b, a \neq 0$, then $|a| \geq |b|$.

Well-Ordering Principle 1.3
Let n_0 be any fixed integer. Then any nonempty set of integers $\geq n_0$ has a least element.

We will use this principle to prove one of the basic propertiess of integers - the division property, which is well known from arithmetic.

Theorem 1.4 (Division property)
For any integers a and b, with $b \neq 0$, there exist unique integers q (the quotient) and r (the remainder) such that $a = bq + r$, with $0 \leq r < |b|$.

Proof. Consider the set of integers $S = \{a - k \cdot b | k \in \mathbf{Z}\}$:

$$\ldots,\ a-2b, a-b, a, a+b, a+2b, \ldots\ .$$

Let S^+ be the subset of all non-negative integers of $S, S^+ \subset S$. The set S^+ is non-empty, so by well-ordering principle has a least element $r = a - qb \geq 0$. Claim that $r < |b|$. If not, then $r - |b| \geq 0$, and also

$$r - |b| = (a - qb) - |b| = a - (q \pm 1)\, b \geq 0$$

(with the sign depending on the sign of b), contradiction.
To show uniqueness, suppose that both

$$\begin{array}{ll} a = bq + r, & 0 \leq r < |b| \\ a = bq' + r', & 0 \leq r' < |b| \end{array}$$

and $r \geq r'$. Subtract to find

$$r - r' = (q' - q)\, b$$

Thus $r - r'$ is a multiple of b. But since $r - r' < |b|$ we have $r = r'$ and then $q = q'$.

$\triangle$

Division algorithm

Given:	$a, b \in \mathbf{N}$
Received:	$0 < q = a \text{ div } b, 0 < r = a \bmod b : a = bq + r$
1.	$Q := 0, R := a$
2.	If $R < b$, then $q = Q, r = R$
	If $R \geq b$,then 3.
3.	$R := R - b, Q := Q + 1$ and 2.

Greatest common divisor

Definition 1.5

If a and b are not both zero, then $d > 0$ is the greatest common divisor of a and b, writen $gcd(a, b)$ or (a, b), if

1. $a \vdots d$ and $b \vdots d$.

2. Whenever $a \vdots c$ and $b \vdots c$, then $d \vdots c$.

Remark 1.6
For convenience we define $(a, 0) = |a|$ for any a, and $\gcd(0.0) = 0$.

The uniqueness of the greatest common divisor follows from property 2 and the fact that it is positive (see property 6 of Corollary 1.2).

The following special charakterization of the greatest common divisor of two integers is fundamental.

Theorem 1.7

If a and $b$are integers, not both zero, then there exist integers x_0 and y_0 such that $gcd\,(a, b) = ax_0 + by_0$ is least positive integer of the form $ax + by$ with $x, y \in \mathbf{Z}$.

Proof. Consider the set of integers $M = \{ax + by | x, y \in \mathbf{Z}\}$

Let M^+ be the subset of all positive integers of $M, M^+ \subset M$. The set M^+ is non-empty, so by well-ordering principle has a least element $d = ax_0 + by_0 > 0$. First we show that any divisor c of a and b divides d. Let $a \vdots c$ and $b \vdots c$. Then

$$d = \underbrace{\underbrace{a \cdot x_0}_{\text{is a multiple of } c} + \underbrace{b \cdot y_0}_{\text{is a multiple of } c}}_{\text{is a multiple of } c} .$$

and d satisfies property 2 of Definition 1.5.

We will show that d also satisfies property 1. Let $a = dq + r, 0 \leq r < d$. Then

$$0 \leq r = a - dq = a - (ax_0 + by_0)\, q = (1 - x_0 q)\, a + (-y_0 q)\, b \in M$$

Since $r < d$ and since d is the smallest positive integer so $r = 0$. Therefore $a \vdots d$, and similarly $b \vdots d$.

$\triangle$

Remark 1.8

The greatest common divisor (a, b) is expressible as $ax + by$, but integers x and y are not unique. For instance,

$$5 = (15, 35) = 15 \cdot (-2) + (35) \cdot (1) = 15 \cdot (5) + (35) \cdot (-2)\,.$$

Euklid's Algorithm.

The greatest common divisor of two numbers can be computed by using a procedure known as the Euclidean algorithm.

The main observation for the Euklidean algorithm is that if $a = bq + r$, then $(a, b) = (b, r)$. Thus given integers $a > b > 0$, the Euklidean algorithm uses a sequence of divisions as follows.

Let $a_0 = a$ and $a_1 = b$, then

$$\begin{array}{cc} a_0 = a_1 q_1 + a_2 & 0 < a_2 < a_1 \\ a_1 = a_2 q_2 + a_3 & 0 < a_3 < a_2 \\ \cdots & \cdots \\ a_{n-2} = a_{n-1} q_{n-1} + a_n & 0 < a_n < a_{n-1} \\ a_{n-1} = a_n q_n & \end{array}$$

Since $a_1 > a_2 > a_3 > \cdots > a_{n-1} > a_n > 0$, the remainders get smaller and smaller, and after a finite number of steps the process stops. Thus $gcd\,(a, b) = (a_0, a_1) = (a_1, a_2) = \cdots = (a_n, 0) = a_n$.

Euklid's Algorithm

Given:	$a, b \in \mathbf{N}. a \geq b$
Received:	$d = \gcd(a, b)$
1.	$A := a, B := b$
2.	$R := A \bmod B$ and 3.
3.	If $R = 0$, then $d = B$
	If $R \neq 0$, then 4.
4.	$A := B, B := R$ and 2.

Extended Euklidean Algorithm

The greatest common divisor of two numbers a, b can be expressible in the form $ax + by$, with x,y$\in Z$, by using the Extended Euclidean algorithm.

$$\begin{array}{c} a_0 = a \cdot 1 + b \cdot 0 \\ a_1 = a \cdot 0 + b \cdot 1 \\ a_2 = a \cdot x_2 + b \cdot y_2 \\ a_3 = a \cdot x_3 + b \cdot y_3 \\ \cdots \\ a_k = a \cdot x_k + b \cdot y_k \end{array}$$

$$
\begin{gathered}
a_j = a_{j-2} - a_{j-1}q_{j-1} = \\
(a \cdot x_{j-2} + b \cdot y_{j-2}) - (a \cdot x_{j-1} + b \cdot y_{j-1})\, q_{j-1} = \\
a\,(x_{j-2} - q_{j-1}x_{j-1}) + b\,(y_{j-2} - q_{j-1}y_{j-1}) \\
x_j = x_{j-2} - q_{j-1}x_{j-1} \\
y_j = y_{j-2} - q_{j-1}y_{j-1}
\end{gathered}
$$

Relativety prime numbers.

Definition 1.9
Two integers a and b are relatively prime if $(a, b) = 1$.

Theorem 1.10
Two integers a and b are relatively prime if and only if there exist integers x and y such that $1 = ax + by$, with some $x, y \in \mathbf{Z}$.

Proof. The first part follows using Theorem 1.7: if two integers a and b are relatively prime , $gcd\,(a, b) = 1$, then $1 = ax + by$.
On the other side, if $ax + by = 1$ and $\gcd(a, b) = d$, then

$$
1 = \underbrace{\underbrace{a \cdot x}_{\text{is a multiple of } d} + \underbrace{b \cdot y}_{\text{is a multiple of } d}}_{\text{is a multiple of } c\ d},
$$

Thus $1 \vdots d$ and $d = 1$.

△

Proposition 1.11(Property of relatively prime numbers)
Let $a_1, ..., a_m$ *and* $b_1, ..., b_n$ be integers. If $(a_i, b_j) = 1$ *with* $1 \le i \le m, 1 \le j \le n$, then $(a_1 \cdots a_m, b_1 \cdots b_n) = 1$.
Without proof.

△

Examples 1.12
1. Let $\frac{a}{b}$ be any rational irreducible fraction $\frac{a}{b}$, so that $(a, b) = 1$. Then $\frac{a^n}{b^n}$ is also an irreducible fraction with any $n \in \mathbf{N}$.
2. Let c and n be positive integers. Then $\sqrt[n]{c}$ is either an integer or an irrational number.

Theorem 1.13
If a is a multiple of relatively prime numbers b and c, then a is a multiple of $b \cdot c$.

Proof. If a is a multiple of b, then $a = b \cdot m$, with some $m \in \mathbf{Z}$. Since integers b and c are relatively prime implies $bx + cy = 1$. Then

$$\underbrace{\underbrace{mbx}_{\text{is a multiple of } c} + \underbrace{mcy}_{\text{is a multiple of } c}}_{\text{is a multiple of } c} = m,$$

Thus m is a multiple of c, $m = c \cdot n$, and $a = n \cdot m \cdot c$, i.e. *a is a multiple of $b \cdot c$.*

△

Prime numbers

Definition 1.14
An integer $p > 1$ is called a prime number if its only divisors are ± 1 and $\pm p$.

Proposition 1.15
Every nonzero integer $a \neq \pm 1$is a multiple of a prime number.
Without proof.

△

Theorem 1.16 (Euklid)
There exist infinitely many prime numbers.

Proof. The proof is indirect. Suppose that there is a finite number of primes $p_1, p_2, ..., p_s$. Consider the number $n = p_1 p_2 \cdots p_s + 1$, which will have a prime factor p. If p were one of the primes p_i, then

$$1 = \underbrace{\underbrace{n}_{\text{is a multiple of } p_i} - \underbrace{p_1 p_2 \cdots p_s}_{\text{is a multiple of } p_i}}_{\text{is a multiple of } p_i},$$

i.e. $1 \vdots p_i$, which is impossible because $p_i > 1$.

$\triangle$

Remark 1.17

1. Note that the first numbers in the form $p_1 p_2 \cdots p_s + 1$ are prime: $2+1 = 3, 2 \cdot 3 + 1 = 7, 2 \cdot 3 \cdot 5 + 1 = 31, 2 \cdot 3 \cdot 5 \cdot 7 + 1 = 211, 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 + 1 = 2311$. However, $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031 = 59 \cdot 509$.

2. Let $\pi(x)$ be the number of primes less than or equal to x. There is a proven, that $\lim\limits_{x \to \infty} \frac{\pi(x)}{\frac{x}{\ln x}} = 1, i.e.\ \pi(x)$ is asymptotic to $\frac{x}{\ln x}$.

Note, that if $x \geq 3$, then $\frac{Ax}{\ln x} < \pi(x) < \frac{Bx}{\ln x}$, with $A = \frac{1}{2}, B = 2$.

If $x \geq 17$, then $\pi(x) > \frac{x}{\ln x}$, and if $x > 1$, then $\pi(x) < 1,255506 \frac{x}{\ln x}$.

3. The Mersenne numbers M_p defined by $M_p = 2^p - 1$ are sometimes prime. It is known, that if $M_p = 2^p - 1$ is prime, then p is prime. Note the fact that the converse is false. For example, $2^{11} - 1 = 23 \cdot 89$ is not prime even thout 11 is. The largest known prime today is the 12978189 digit Mersenne prime $2^{43112609} - 1$ found in August 2008.

Proposition 1.18

(Property of prime numbers) *An integer $p > 1$ is prime if and only if it satisfies the following property: If $ab \vdots p$ for integers a and b, then either $a \vdots p$ or $b \vdots p$.*

Without proof.

$\triangle$

Proposition 1.19

If a is a multiple of prime numbers p_1 and p_2, then a is a multiple of $p_1 p_2$.

This is the conclution of the theorem 1.13.

$\triangle$

Theorem 1.20 (Fundamental Theorem of Arithmetic)

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

$$n = p_1^{k_1} \cdot p_2^{k_2} \cdot \cdots \cdot p_s^{k_s},$$

with positive integer exponents and primes $p_1 < p_2 < \cdots < p_s$.

Without proof.

$\triangle$

Lecture 2

Congruences. Congruence classes. $\mathbf{Z}_m$. Finite fields.

Congruences

Definition 2.1
Let m be a positive integer. Integers a anb b are said to be congruent modulo m if m divides $a-b$, written $a \equiv b \pmod{m}$.

Examples 2.2
1. $a \equiv b \pmod{1}$ for all integers a and b.
2. $a \equiv b \pmod{2}$ if and only if integers a and b are both odd or even.

The next proposition presents basic properties of congruences.

Proposition 2.3
1. *Reflexivity: $a \equiv a \pmod{m}$ for all $a \in Z$*
2. Symmetry: if $a \equiv b \pmod{m}$ then $b \equiv a \pmod{m}$.
3. Transitivity: if $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$ then $a \equiv c \pmod{m}$
4. If $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$ then $a \pm b \equiv c \pm d \pmod{m}$.
5. If $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$ then $a \cdot b \equiv c \cdot d \pmod{m}$.
6. Cancellation property: if $ab \equiv ac \pmod{m}$, then $b \equiv c \left(\mod \frac{m}{d}\right)$, where $d = (a, m)$.
7. If $ab \equiv ac \pmod{m}$ and $(a, m) = 1$, then $b \equiv c \pmod{m}$.

Remark 2.4
From $3 \equiv 15 \pmod{6}$ we obtain $1 \not\equiv 5 \pmod{6}$ because $(3, 6) \neq 1$.

Proposition 2.5
Let a and $m > 1$ be integers. There exist an unique integer $r \in \{0, 1, \ldots, m-1\}$ such that $a \equiv r \pmod{m}$.

Proof. Write an integer a as $a = mq + r$ with $0 \leq r \leq m-1$. Then $a \equiv r \pmod{m}$.

△

Congruence classes

Definition 2.6

Let a and m be integers and $m > 1$. The congruence class of an integer a modulo m, denoted ${}_mK_a$ (or $\bar{a}$, with only implicit reference to m), is the set of all integers equal to $a \bmod m$:

$$ {}_mK_a = \{b \in \mathbf{Z} | b \equiv a \,(\mathrm{mod}\, m)\}\,. $$

We say that $\mathbf{Z}_m$ is the set of all congruence classes.

Examples 2.7

1. $m = 2, a = 0 : K_0 = \bar{0} = \{b \in \mathbf{Z} | b \equiv 0 \,(\mathrm{mod}\, 2)\} = 2\mathbf{Z}$ is the set of all even integers.

2. $m = 2, a = 1 : K_1 = \bar{1} = \{b \in \mathbf{Z} | b \equiv 1 \,(\mathrm{mod}\, 2)\}$ is the set of all odd integers.

3. $m = 2, a = 2 : K_2 = \bar{2} = \{b \in \mathbf{Z} | b \equiv 2 \,(\mathrm{mod}\, 2)\} = 2\mathbf{Z}$ is the set of all even integers. The next proposition presents basic properties of congruence classes.

Proposition 2.8

1. Reflexivity:: if $a \in Z$ then $a \in {}_mK_a$.

2. Symmetry: if $a \in K_b$ then $b \in {}_mK_a$.

3. Transitivity: if $a \in K_b$ and $b \in {}_mK_c$ then $a \in {}_mK_c$.

Write an integer a as $a = mq + r$ with $0 \leq r \leq m - 1$. The Proposition 2.5 says that the congruence class of an integer a, $\bar{a}$, is precisely the set of integers with the same remainder r, i.e. $\bar{a} = \bar{r}$. For some $K \in \mathbf{Z}_m$ a choise of ordinary integers a so that $\bar{a} = K$ is a representative for the congruence class K. The remainders $0, 1, \ldots, m - 1$ are representatives of the congruence classes: $\mathbf{Z}_m$ is simply the set $\left\{\overline{0}, \overline{1}, ..., \overline{m-1}\right\}$. Moreover, the set $\left\{\overline{0}, \overline{1}, ..., \overline{m-1}\right\}$ is a decomposition of the set of all integers $\mathbf{Z}$, i.e.

$$ (i)\ \mathbf{Z} = \overline{0} \cup \overline{1} \cup ... \cup \overline{m-1}, $$
$$ (ii)\ \text{if } \overline{a} \cap \overline{b} \neq \emptyset \text{ then } \overline{a} = \overline{b}\,. $$

It is possible to do arithmetic with congruence classes. To add two congruence classes modulo m, we just select any element a from the first class and any element b from the second class, and then compute $a + b$ as we would for normal integers.

The sum of the two congruence classes is then defined to be equal to the congruence class containing the sum $a + b$.

Multiplication of congruence classes behaves in a similar manner: to multiply two congruence classes, we select any elements a and b from each of the classes and multiply them together. The product of the congruence classes is then defined to be the congruence class containing the product $a \cdot b$.

Definition 2.9

If $K', K'' \in Z_m$ *and* $a \in K', b \in K''$ *then* $K' + K'' = \overline{a + b}$ *and* $K' \cdot K'' = \overline{a \cdot b}$

Remark 2.10

The Proposition 2.3 says that the sum and the product of two congruence classes are independent of the representatives: if $a, a' \in K'$ and $b, b' \in K''$, i.e $a \equiv a' \pmod{m}$ and $b \equiv b' \pmod{m}$ then $\overline{a + b} = \overline{a' + b'}$ and $\overline{a \cdot b} = \overline{a' \cdot b'}$.

Example 2.11

1. Arithmetic mod 3

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

2. Multiplication mod 6

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{0}$	$\overline{2}$	$\overline{4}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{2}$	$\overline{0}$	$\overline{4}$	$\overline{2}$
5	$\overline{0}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

The set of congruence classes $\mathbf{Z}_m$ inherit many basic properties from ordinary arithmetic.

Proposition 2.12

Fix the modulus m. Let $\bar{a}, \bar{b}, \bar{c} \in Z_m$

1. *Associativity of addition:* $\left(\bar{a}+\bar{b}\right)+\bar{c}=\bar{a}+\left(\bar{b}+\bar{c}\right)$.
2. *Commutativity of addition:* $\bar{a}+\bar{b}=\bar{b}+\bar{a}$.
3. *Property of 0:* $\bar{a}+\bar{0}=\bar{a}$.
4. *Additionative inverse:* $\bar{a}+\overline{(-a)}=\bar{0}$.
5. *Distributivity:* $\begin{array}{l}\bar{a}\cdot\left(\bar{b}+\bar{c}\right)=\bar{a}\cdot\bar{b}+\bar{a}\cdot\bar{c}\\ \left(\bar{a}+\bar{b}\right)\cdot\bar{c}=\bar{a}\cdot\bar{c}+\bar{b}\cdot\bar{c}\end{array}$
6. *Associativity of multiplication :* $\left(\bar{a}\cdot\bar{b}\right)\cdot\bar{c}=\bar{a}\cdot\left(\bar{b}\cdot\bar{c}\right)$.
7. *Commutativity of multiplication:* $\bar{a}\cdot\bar{b}=\bar{b}\cdot\bar{a}$.
8. *Property of 1:* $\bar{a}\cdot\bar{1}=\bar{a}$.

In this context, an additive inverse exist for any congruence class $\bar{a} \in \mathbf{Z}_m$, but a multiplicative inverse modulo m, defined as the solution to equation $\bar{a}\cdot\overline{x}=\overline{1}$ in $\mathbf{Z}_m$, or congruence equation $ax \equiv 1 \pmod{m}$, does not always exist.

Theorem 2.13

A congruence class ${}_mK_a = \overline{a}$ has a multiplikative inverse modulo m if and only if $(a, m) = 1$.

Proof. If a multiplicative inverse $\bmod m$ to a congruence class $\overline{a}$ is a congruence class $\bar{b}$ then $\bar{a}\cdot\bar{b}=\overline{1}$ in $\mathbf{Z}_m$ and $ab \equiv 1 \pmod{m}$, i.e. $ab-1=mt$ and $ab+m(-t)=1$with some integer t, witch implies that $(a, m) = 1$.

On the other side, if $(a, m) = 1$ then $ab + mt = 1$, with some $b, t \in \mathbf{Z}$. Thus $ab \equiv 1 \pmod{m}$ and $\bar{a}\cdot\bar{b}=\overline{1}$ in $\mathbf{Z}_m$.

△

Finite fields

Here we introduce some basic construction of an algebraic system.

A *ring* R is a set with two operations - addition and multiplication ($+$ and $\cdot$) with two special elements: 0 (additive identity) and 1 (multiplicative identity), and with the properties:

1. *Associativity of addition:* $(a+b)+c=a+(b+c)$ for all $a, b, c \in R$.
2. *Commutativity of addition:* $a+b=b+a$.
3. *Property of 0:* $a+0=a$ for all $a \in R$
4. *Additionative inverse: if* $a \in R$ then a has a additionative inverse $b \in R$: $a+b=0$.
5. *Distributivity:* $\begin{array}{l}a\cdot(b+c)=a\cdot b+a\cdot c\\ (a+b)\cdot c=a\cdot c+b\cdot c\end{array}$ for all $a, b, c \in R$.

6. Associativity of multiplication : $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.

Very often, a concrete ring has some additional properties:

7. Commutativity of multiplication: $a \cdot b = b \cdot a$ for all $a, b \in R$, then the ring is called commutative ring.

8. Property of 1 : $a \cdot \overline{1} = \bar{a}$.for all $a \in R,$, then the ring is called a ring with unit.

In a commutative ring R with unit, for a given $a \in R$, if there is $a^{-1} \in R$ so that $a \cdot a^{-1} = 1$, then a^{-1} is said to be multiplicative inverse, and a is to said to have a multiplicative inverse.

9. A commutative ring in witch every nonzero element has multiplicative inverse is a field.

Example 2.14

1. The integers $\mathbf{Z}$ with the usual addition and multplication is a commutative ring with unit.

2. The set $\mathbf{Z}_m$ with addition and multiplcation modulo m is a commutaive ring with unit.

3. The even integers $2\mathbf{Z}$ with the usual addition and multplication is a commutative ring without unit.

4. The rational numbers $\mathbf{Q}$ and the real numbers $\mathbf{R}$ with the usual addition and multplication are all fields. This fields have an infinite number of elements. However, there do exist fields with finite numbers of elements.We are able to show that if an integer p is prime then the set of congruence classes $\mathbf{Z}_p$ is a field.

Theorem 2.15

$\mathbf{Z}_m$ *is a finite field if and only if* m *is prime.*

Proof. If m is a prime number p, then $(1, m) = (2, m) = \cdots = (m-1, m) = 1$. Thus all the nonzero classes of $\mathbf{Z}_p$ have multiplicative inverses, and hence, $\mathbf{Z}_p$ is a field. On the other hand, if m is not a prime, then $m = a \cdot b$ with $1 < a, b < m$. Thus $(a, m) = a > 1$ and the nonzero classses $\bar{a}$ and $\overline{b}$ have not multiplicative inverses and $\mathbf{Z}_m$ is not a field.

△

Example 2.16

1 .Consider $p = 7$. Then $\mathbf{Z}_7$ is a field because all the nonzero elements $\overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}$ have inverses, witch are $\overline{1}, \overline{4}, \overline{5}, \overline{2}, \overline{3}, \overline{6}$ respectively.

2. Consider $p = 6$. Then the nonzero elements $2, 3, 4$ has not inverses, because $2 \cdot 3 \equiv 3 \cdot 4 \equiv 0 \,(\mathrm{mod}\, 6)$ and hence $\mathbf{Z}_6$ is not a field.

Lecture 3

Euler's function. Euler's theorem. Fermat's little theorem. Wilson's theorem. Chinese remainder theorem

Euler's function

Definition 3.1

We say that U_m *is the set of all congruence classes modulo* m*, witch have multiplicative inverses:*

$$U_m = \{\bar{a} |\, (a,m) = 1, 0 \leq a \leq m-1\}.$$

We say that the set U_m *has* $\varphi(m)$ *elements, where* φ *is the Euler function, i.e.* $\varphi(m)$ *is a cardinality of* U_m*.*

Remark 3.2

For a positive integer m the Euler function $\varphi(m)$ is the number of integers a so that $0 \leq a \leq m-1$ and $(a,m) = 1$.

The following property of the Euler function is fundamental.

Proposition 3 3 (multiplicative function)

If $(m,n) = 1$*, then* $\varphi(m \cdot n) = \varphi(m) \cdot \varphi(n)$ ($\varphi(m)$ *is a multiplicative function).*

Without proof.

Example 3.4

1. If p is a prime number, then $\varphi(p) = p-1$, because all numbers $1, 2, ..., p-1$ are relatively prime with p.

2. Let a be an integer and p be a prime number . Then $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha \left(1 - \frac{1}{p}\right)$, because if $(a,p) = 1$ and $0 \leq a \leq p^\alpha - 1$ then $a = t \cdot p$ with $0 \leq t \leq p^{\alpha-1} - 1$.

3. Let m is a product of primes $m = p_1^{k_1} \cdot p_2^{k_2} \cdot \cdots \cdot p_s^{k_s}$. Then using Propositiom 3.3 we have $\varphi(m) = m\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_s}\right)$.

4. $\varphi(120) = \varphi(8 \cdot 3 \cdot 5) = \varphi(2^3) \cdot \varphi(3) \cdot \varphi(5) = (2^3 - 2^2)(3-1)(5-1) = 32$.
$\varphi(120) = 120 \cdot \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right) = 32$.

Proposition 3.5

1.The set U_m is closed under multiplication in this sence that $a, b \in U_m$ implies $ab \in U_m$.

2. The set U_m is closed under inverses in this sence that $a \in U_m$ implies $a^{-1} \in U_m$.

Proof is obvious.

Euler's theorem. Fermat's little theorem

Definition 3.6

Let m be a positive integer and $s = \varphi(m)$. A set of integers $r_1, r_2, ..., r_s$ is called a prime residue system modulo m if $U_m = \{\overline{r_1}, \overline{r_2}, ..., \overline{r_s}\}$

Lemma 3.7

Let $r_1, r_2, ..., r_s$ be a prime residue system modulo m and an integer a is relatively prime with m. Then the system $ar_1, ar_2, ..., ar_s$ is also a prime residue system modulo m.

Proof. The integers $r_1, r_2, ..., r_s$ and a are relatively prime with m. Then, by Proposition 1.11, we have that for any $i, 1 \leq i \leq s$, $(ar_i, m) = 1$. Thus the integers $ar_1, ..., ar_s$ are relatively prime with m. Now if $ar_i \equiv ar_j \pmod{m}$ then, by cancellation property of Proposition 2.7 , $r_i \equiv r_j \pmod{m}$ and the integers $ar_1, ar_2, ..., ar_s$ is a prime residue system modulo m.

△

Corollary 3.8

Let $r_1, r_2, ..., r_s$ be a prime residue system modulo m and an integer a is relatively prime with m. Then

$$a^s r_1 \cdots r_s \equiv r_1 \cdots r_s \pmod{m}.$$

Proof. From Lemma 3.7 we have that

$$U_m = \{\overline{r_1}, ..., \overline{r_s}\} = \{\overline{ar_1}, ..., \overline{ar_s}\},$$

and hence

$$\overline{ar_1} \cdot \cdot \cdot \cdot \cdot \overline{ar_s} = \overline{r_1} \cdot ... \cdot \overline{r_s},$$
$$ar_1 \cdots ar_s \equiv r_1 \cdots r_s \,(\mathrm{mod}\, m)\,,$$
$$a^s r_1 \cdots r_s \equiv r_1 \cdots r_s \,(\mathrm{mod}\, m)\,.$$

△

Theorem 3.9 (Euler)
If $(a, m) = 1$, *then* $a^{\varphi(m)} \equiv 1 \,(\mathrm{mod}\, m)$.

Proof Let $r_1, r_2, ..., r_s$ be a prime residue system modulo m . From Corollary 3.8 we have

$$a^s r_1 \cdots r_s \equiv r_1 \cdots r_s \,(\mathrm{mod}\, m)$$

and since the integers are relatively prime to m, we can apply cancellation property of Proposition 2.7 to obtain

$$a^s \equiv 1 \,(\mathrm{mod}\, m)$$

and

$$a^{\varphi(m)} \equiv 1 \,(\mathrm{mod}\, m)$$

△

Corollary 3.10
If $(a, m) = 1$, *then* $a^{-1} \equiv a^{\varphi(m)-1} \,(\mathrm{mod}\, m)$.

Theorem 3.11(Fermat little theorem)
If p is a prime and a is an integer not divisible by p, then $a^{p-1} \equiv 1 \,(\mathrm{mod}\, p)$

Proof is obvious.

Corollary 3.12
If p is a prime and a is an integer not divisible by p, then $a^{-1} \equiv a^{p-2} \,(\mathrm{mod}\, p)$.

The next theorem is a generalization of Fermat's little theorem.

Defintion 3.13
A natural number is squarefree if it is the product of distinct primes.

Theorem 3.14
If a natural number m is squarefree, that is $m = p_1 \cdot \cdots \cdot p_r$ for distinct primes $p_1, ..., p_r$, then for all integers a, $a^{\varphi(m)+1} \equiv a \pmod{m}$.

Proof. We have that

$$\varphi(m) = \varphi(p_1 \cdot \cdots \cdot p_r) = (p_1 - 1) \cdot \cdots \cdot (p_r - 1).$$

Let a be an integer. For each prime p_i, $1 \leq i \leq r$, we have two cases.
First consider the case $(a, p_i) = 1$.From Fermat Theorem we have.

$$a^{p_i - 1} \equiv 1 \pmod{m}$$
$$a^{\varphi(m)} = (a^{p_i - 1})^{\frac{\varphi(m)}{p_i - 1}} \equiv 1 \pmod{p_i}$$

and

$$a^{\varphi(m)+1} \equiv a \pmod{p_i}.$$

Next,consider the case $(a, p_i) \neq 1$. Then a is divisible by p_i , i.e.

$$a \equiv 0 \pmod{p_i}$$

and

$$a^{\varphi(m)+1} \equiv 0 \equiv a \pmod{p_i}.$$

So we have that in both cases the integer $a^{\varphi(m)+1} - a$ is divisible by prime p,
$1 \leq i \leq r$ and from Proposition 1.19 we have that $a^{\varphi(m)+1} - a$ is divisible by $p_1 \cdot \cdots \cdot p_r = m$:

$$a^{\varphi(m)+1} \equiv a \pmod{m}$$
$$\bar{a}^{\varphi(m)+1} = \bar{a}.$$

△

Wilson's theorem

Theorem 3.15 (Wilson)
An integer m *is prime if and only if* $(m-1)! \equiv -1 \pmod{m}$

Proof. We have two cases. First consider the case where m is prime p.Thus all the nonzero classes of $\mathbf{Z}_p$ have multiplicative inverses:

$$\overline{1}^{-1} = \overline{i_1}, ..., \overline{(p-1)}^{-1} = \overline{i_{p-1}}.$$

If $\overline{a}^{-1} = \overline{a}$ then $\overline{a}^2 = \overline{a} \cdot \overline{a}^{-1} = \overline{1}$ and

$$\begin{gathered} a^2 \equiv 1 \pmod{p} \\ (a^2 - 1) \equiv 0 \pmod{p} \\ (a-1)(a+1) \equiv 0 \pmod{p}. \end{gathered}$$

Thus, $(i-1)(i+1)$ is divisible by p This means that either $i-1$ is divisible by p and $i-1=0$, i.e. $i=1$, or $i+1$ is divisible by p and $i+1=p$, i.e. $i=p-1$.

Hence, we have

$$\begin{gathered} \overline{(p-1)!} = \overline{1} \cdot \overline{2} \cdot \cdots \cdot \overline{(p-2)} \cdot \overline{(p-1)} \\ = \overline{1} \cdot (\overline{2} \cdot \overline{i_2}) \cdot \cdots \cdot (\overline{(p-2)} \cdot \overline{i_{p-2}}) \cdot \overline{(p-1)} \\ = \overline{1} \cdot \overline{(p-1)} = \overline{(p-1)} = \overline{(-1)} \end{gathered}$$

and

$$(p-1)! \equiv -1 \pmod{p}.$$

Next, consider the case where m is the product $a \cdot b$, $1 < a, b < m$. We have three cases:

1. If $1 < a < b < m$, then $(m-1)! = 1 \cdots a \cdots b \cdots (m-1)$ is divisible by $a \cdot b$ and $(m-1)! = 0 \neq -1 \pmod{m}$.

2. If $m = a \cdot a = a^2$ and $a > 2$, $m > 4$, then $(m-1)! = 1 \cdots (1 \cdot a) \cdots (2 \cdot a) \cdots (m-1)$ is divisible by a^2 and $(m-1)! = 0 \neq -1 \pmod{m}$.

3. If $m = 4$, then $(4-1)! = 3! = 6 \equiv 2 \pmod{4} \neq -1 \pmod{4}$.

△

Remark 3.16
Using Wilson's theorem, we have

$$f(m) = \sin\left(\frac{\pi \cdot ((m-1)! + 1)}{m}\right) = \left[\begin{array}{c} 0, \text{ with prime } m \\ \neq 0, \text{ with composite } m \end{array}\right. .$$

However, to use $f(m)$ as a primality test is not practical, because we have to calculate a very large number $(m-1)!$ even for small m.

Chinese remainder theorem

Theorem 3.17 (Chinese remainder theorem).
Let $m_1, m_2, ..., m_k$ be positive integers, with $(m_i, m_j) = 1$,when $i \neq j$. Then the system of congruences

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\cdots \\ x &\equiv a_k \pmod{m_k} \end{aligned}$$

has a solution. Moreover, any two solutions are congruent modulo $m_1 \cdot m_2 \cdot \cdots \cdot m_k$.

Proof of theorem is based on the following Lemmas.

Lemma 3.18
Let $m_1, m_2, ..., m_k$ be positive integers, with $(m_i, m_j) = 1$,when $i \neq j$. If $a \equiv b \pmod{m_i}$, $1 \leq i \leq k$, then $a \equiv b \pmod{m_1 m_2 \cdots m_k}$.

Proof The proof uses induction. We consider the case when $k = 2$: we have $a \equiv b \pmod{m_1}$ and $a \equiv b \pmod{m_2}$, with $(m_1, m_2) = 1$. By Theorem 1.10 there exist x and y such that $xm_1 + ym_2 = 1$. Multiplying through by $(a-b)$, we obtain

$$\underbrace{\underbrace{(a-b)}_{\vdots m_2} x m_1}_{\vdots m_1 m_2} + \underbrace{\underbrace{(a-b)}_{\vdots m_1} y m_2}_{\vdots m_1 m_2} = a - b$$

Thus, we have that $a - b$ is divisible by $m_1 m_2$.

△

Lemma 3.19

If $a \equiv b \pmod{m}$ and m is divisible by d, then $a \equiv b \pmod{d}$.

Proof is is obvious (I hope).

$\triangle$

Now we can prove Chinese remainder theorem.

Proof of Theorem 3.17. Define $M = m_1 \cdot m_2 \cdot \cdots \cdot m_k, M_i = \dfrac{M}{m_i}$ with $i = 1, 2, ..., k$. From $(M_i, m_i) = 1$ there exist N_i such that

$$M_i N_i \equiv 1 \pmod{m_i}, \ i = 1, 2, ..., k.$$

N_i is the multiplicative inverse of M_i modulo m_i.
Consider the number

$$x^* = M_1 N_1 a_1 + \cdots + M_k N_k a_k.$$

Then we have

$$x^* = M_1 N_1 a_1 + \cdots + M_k N_k a_k \equiv M_i N_i a_i \equiv a_i \pmod{m_i},$$

hence x^* is a solution of the system of congruences. Therefore, if x is a solution of the system of congruences

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\cdots \\ x &\equiv a_k \pmod{m_k} \end{aligned}$$

then

$$\begin{aligned} x &\equiv x_0 \pmod{m_1} \\ x &\equiv x_0 \pmod{m_2} \\ &\cdots \\ x &\equiv x_0 \pmod{m_k} \end{aligned} .$$

and using Lemma 3.18 and Lemma 3.19 we have

$$x = x^* \pmod{m_1 m_2 \cdots m_k}.$$

△

Example 3.20
Solve the system of congruences:

$$\begin{array}{l} x \equiv 2 \bmod 3 \\ x \equiv 3 \bmod 5 \\ x \equiv 2 \bmod 7 \end{array} .$$

1. $M = 3 \cdot 5 \cdot 7 = 105$
2. $M_1 = \frac{105}{3} = 35, M_2 = \frac{105}{5} = 21; M_3 = \frac{105}{7} = 15.$
3. $N_1 \equiv 35^{-1} \pmod 3 = 2; N_2 \equiv 21^{-1} \pmod 5 = 1; N_3 = 15^{-1} \pmod 7 = 1.$
4. $x^* = 35 \cdot 2 \cdot 2 + 21 \cdot 1 \cdot 3 + 15 \cdot 1 \cdot 2 = 233.$

Finally, all solutions of the system of congruences is $x = 233 + 105t, t \in \mathbf{Z}$, i.e. $x \in_{105} K_{233}$.

Lecture 4

Determinants: definitions.

Determinants: definitions

The line l in the xy-plane can be represented by an equation of the form $a_1x + a_2y = b$, where a_1, a_2 and b are real constants and $a_1^2 + a_2^2 > 0$, a_1 and a_2 are not both zero. Consider the two lines l_1 and l_2 :

$$\begin{array}{ll} a_{11}x + a_{12}y = b_1 & (a_{11}^2 + a_{12}^2 > 0) \\ a_{21}x + a_{22}y = b_2 & (a_{21}^2 + a_{22}^2 > 0) \end{array} \qquad (1)$$

The solutions of the system of equations correspond to points of intersection of l_1 and l_2. Add a_{22} times the first equation to $-a_{12}$ times the second and add $-a_{21}$ times the first equation to a_{11} times the second to obtain

$$\begin{array}{l} (a_{11}a_{22} - a_{21}a_{12})\, x = b_1a_{22} - b_2a_{12} \\ (a_{11}a_{22} - a_{21}a_{12})\, y = b_2a_{11} - b_1a_{21} \end{array}$$

There are three possibilities:

1) if $a_{11}a_{22} - a_{21}a_{12} = 0$ and $b_1a_{22} - b_2a_{12} = 0$ then this system has infinitely many solutions and the lines l_1 and l_2 overlap completely;

2) if $a_{11}a_{22} - a_{21}a_{12} = 0$ and $b_1a_{22} - b_2a_{12} \neq 0$ then this system no solution and the lines l_1 and l_2 are parallel and do not intersect;

3) if $a_{11}a_{22} - a_{21}a_{12} \neq 0$ then this system has exactly one solution

$$\begin{array}{l} x = \dfrac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \\ y = \dfrac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{21}a_{12}} \end{array}$$

and the lines l_1 and l_2 intersect at one point.

Definition 4.1

1. The table $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *is called the coefficient matrix of the system* (1) .

2. Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *be a matrix. We define the determinant of* A*, denoted by* $\det A$*, to be scalar*

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

The notation $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ is also used for the determinat of A. The determinat's formula is obtained by multiplying the entries on the rightwarf arrow and subtracting the product on the the leftward arrow:

$$\text{positive product } \begin{pmatrix} a_{11} & & a_{12} \\ & \searrow & \\ a_{21} & & a_{22} \end{pmatrix} \qquad \text{negative product } \begin{pmatrix} a_{11} & & a_{12} \\ & \nearrow & \\ a_{21} & & a_{22} \end{pmatrix}.$$

If $\det A \neq 0$, where A is the coefficient matrix of the system (1), then the system has a unique solution and this solution is

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \text{ ir } y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

This formula is known as Cramer's Rule.

The plane p in the xyz-space can be represented by an equation of the form $a_1x + a_2y + a_3z = b$, where a_1, a_2, a_3 and b are real constants and $a_1^2 + a_2^2 + a_3^2 > 0$, a_1, a_2 and a_3 are not both zero. Consider the three planes p_1, p_2 and p_3 :

$$\begin{array}{ll} a_{11}x + a_{12}y + a_{13}z = b_1 & (a_{11}^2 + a_{12}^2 + a_{13}^2 > 0) \\ a_{21}x + a_{22}y + a_{23}z = b_2 & (a_{21}^2 + a_{22}^2 + a_{23}^2 > 0) \\ a_{31}x + a_{32}y + a_{33}z = b_3 & (a_{31}^2 + a_{32}^2 + a_{33}^2 > 0) \end{array} \qquad (2)$$

The solutions of the system of equations correspond to points of intersection of p_1, p_2 and p_3.

Add $(a_{22}a_{33} - a_{32}a_{23})$ times the first equation to $(a_{32}a_{13} - a_{12}a_{33})$ times the second equation and then add the obtained sum to $(a_{12}a_{23} - a_{22}a_{13})$ times the third equation

$$\begin{array}{rcl} a_{11}x + a_{12}y + a_{13}z = b_1 & \big| \cdot & a_{22}a_{33} - a_{32}a_{23} \\ a_{21}x + a_{22}y + a_{23}z = b_2 & \big| \cdot & a_{32}a_{13} - a_{12}a_{33} \\ a_{31}x + a_{32}y + a_{33}z = b_3 & \big| \cdot & a_{12}a_{23} - a_{22}a_{13} \end{array}$$

to obtain

$$\begin{gathered} (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13})\, x = \\ b_1a_{22}a_{33} + b_2a_{32}a_{13} + b_3a_{12}a_{23} - b_1a_{32}a_{23} - b_2a_{12}a_{33} - b_3a_{22}a_{13}. \end{gathered}$$

Add $(a_{23}a_{31} - a_{33}a_{21})$ times the first equation to $(a_{33}a_{11} - a_{13}a_{31})$ times the second equation and then add the obtained sum to $(a_{13}a_{21} - a_{23}a_{11})$ times the third equation

$$\begin{array}{rcl} a_{11}x + a_{12}y + a_{13}z = b_1 & \big| \cdot & a_{23}a_{31} - a_{33}a_{21} \\ a_{21}x + a_{22}y + a_{23}z = b_2 & \big| \cdot & a_{33}a_{11} - a_{13}a_{31} \\ a_{31}x + a_{32}y + a_{33}z = b_3 & \big| \cdot & a_{13}a_{21} - a_{23}a_{11} \end{array}$$

to obtain

$$\begin{gathered} (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13})\, y = \\ a_{11}b_2a_{33} + a_{21}b_3a_{13} + a_{31}b_1a_{23} - a_{11}b_3a_{23} - a_{21}b_1a_{33} - a_{31}b_2a_{13}. \end{gathered}$$

Add $(a_{21}a_{32} - a_{31}a_{22})$ times the first equation to $(a_{31}a_{12} - a_{11}a_{32})$ times the second equation and then add the obtained sum to $(a_{11}a_{22} - a_{21}a_{12})$ times the third equation

$$\begin{array}{rcl} a_{11}x + a_{12}y + a_{13}z = b_1 & \big| \cdot & a_{21}a_{32} - a_{31}a_{22} \\ a_{21}x + a_{22}y + a_{23}z = b_2 & \big| \cdot & a_{31}a_{12} - a_{11}a_{32} \\ a_{31}x + a_{32}y + a_{33}z = b_3 & \big| \cdot & a_{11}a_{22} - a_{21}a_{12} \end{array}$$

to obtain

$$\begin{gathered} (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13})\, z = \\ a_{11}a_{22}b_3 + a_{21}a_{32}b_1 + a_{31}a_{12}b_2 - a_{11}a_{32}b_2 - a_{21}a_{12}b_3 - a_{31}a_{22}b_1. \end{gathered}$$

If $(a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1) \neq 0$, then

$$x = \frac{d_1b_2c_3 + d_2b_3c_1 + d_3b_1c_2 - d_1b_3c_2 - d_2b_1c_3 - d_3b_2c_1}{a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13}}$$

$$y = \frac{a_1d_2c_3 + a_2d_3c_1 + a_3d_1c_2 - a_1d_3c_2 - a_2d_1c_3 - a_3d_2c_1}{a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13}}$$

$$z = \frac{a_1b_2d_3 + a_2b_3d_1 + a_3b_1d_2 - a_1b_3d_2 - a_2b_1d_3 - a_3b_2d_1}{a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{11}a_{32}c_{23} - a_{21}a_{12}c_{33} - a_{31}a_{22}a_{13}}$$

and the planes p_1, p_2 and p_3 intersect at one point.

Definition 4.2

1. The table $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ *is called the coefficient matrix of the system* (2) .

2. Let $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ *be a matrix. We define the determinant of* A*, denoted by* $\det A$*, to be scalar*

$$\det A = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

The notation $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ is also used for the determinat of A. The determinat's formula is obtained by recopying the first and second rows as shown in (3) . The determinant is computed by summning the products on the rightward arrows and subtracting the products on the leftward arrows.

$$\begin{matrix} \text{positive} \\ \text{products} \end{matrix} : \begin{pmatrix} a_1 & & b_1 & & c_1 \\ & \searrow & & & \\ a_2 & & b_2 & & c_2 \\ & \searrow & & \searrow & \\ a_3 & & b_3 & & c_3 \\ & \searrow & & \searrow & \\ a_1 & & b_1 & & c_1 \\ & & & \searrow & \\ a_2 & & b_2 & & c_2 \end{pmatrix}, \quad \begin{matrix} \text{negative} \\ \text{products} \end{matrix} : \begin{pmatrix} a_1 & & b_1 & & c_1 \\ & & & \nearrow & \\ a_2 & & b_2 & & c_2 \\ & \nearrow & & \nearrow & \\ a_3 & & b_3 & & c_3 \\ & \nearrow & & \nearrow & \\ a_1 & & b_1 & & c_1 \\ & \nearrow & & & \\ a_2 & & b_2 & & c_2 \end{pmatrix}$$

(3)

If $\det A \neq 0$, where A is the coefficient matrix of the system (1), then the system has a unique solution and this solution is

$$x = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \; y = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \; z = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

This formula is known as Cramer's Rule.

Definition 4.3
A permutation of the set of integers $\{1, 2, ..., n\}$ *is an arrangement of the numbers* $1, 2, ..., n$, *denoted* $(i_1, i_2, ..., i_n)$. *An inversion occurs in the permutation* $(i_1, i_2, ..., i_n)$ *when* $i_r > i_s$ *but* $r < s$. *The total number of inversions in the permutation*$(i_1, i_2, ..., i_n)$ *is denoted* $\text{inv}(i_1, i_2, ..., i_n)$ *A permutaion is called even if* $\text{inv}(i_1, i_2, ..., i_n)$ *is an even integer and is called odd if* $\text{inv}(i_1, i_2, ..., i_n)$ *is an odd integer.*

The set $\{1, 2, ..., n\}$ have $n!$ different permutations.

Example 4.4
The table of the permutations of $\{1, 2\}$:

even permutation	odd permutation
$(1, 2)$	$(2, 1)$

The table of the permutations of $\{1, 2, 3\}$:

even permutations	odd permutations
$(1, 2, 3)$	$(3, 2, 1)$
$(2, 3, 1)$	$(2, 1, 3)$
$(3, 1, 2)$	$(1, 3, 2)$

.

Definiton 4.5

The determinat of an $n \times n$ *matrix* $\begin{pmatrix} a_{11} & \cdots & .a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *is the sum of* $n!$ *signed products* $(-1)^{\text{inv}(i_1,i_2,...,i_n)} a_{1i_1} a_{2i_2} \cdots a_{ni_n}$, *where* $(i_1, i_2, ..., i_n)$ *is a permutation of* $\{1, 2, ..., n\}$:

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{(i_1,...,i_n)} (-1)^{\text{inv}(i_1,i_2 \ ...,i_n)} a_{1i_1} a_{2i_2} \cdots a_{ni_n}$$

where $\sum_{(i_1,...,i_n)}$ *is the sum over all permutations* $(i_1, i_2, ..., i_n)$.

Lecture 5

Determinants: properties.

Determinants: properties

Definition 5.1
Supose that A is an $n \times n$ matrix. The (i,j)- minor of A is defined to be the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and the j^{th} column. We will denote this by $M_{ij}(A) = M_{ij}$.

Proposition 5.2
Let A be a 3×3 matrix, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Then

$$\det A = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}.$$

Proof.
$a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$
$a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} = \det A.$

△

This proposition is a special case of the folowing theorem, which we state without proof.

Theorem 5.3
Let $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$. Then we can expand the determinant of A by along any row or column:

$$\det A = (-1)^{i+1} a_{i1}M_{i1} + (-1)^{i+2} a_{i2}M_{i2} + \cdots + (-1)^{i+n} a_{in}M_{in}$$

for each $1 \leq i \leq n$, *the* i*-th row expansion, and*

$$\det A = (-1)^{1+j}\, a_{1j} M_{1j} + (-1)^{2+j}\, a_{2j} M_{2j} + \cdots + (-1)^{n+j}\, a_{nj} M_{nj}$$

for each $1 \leq i \leq n$, *the* j*-th column expansion.*
Without proof.

We shall present now some of the fundamental properties of the determinant.

Property D1

If the matrix $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *has zero row, i.e.* $a_{i1} = a_{i2} = \cdots = a_{in} = 0$ *for some*$1 \leq i \leq n$ *, then* $\det A = 0$. *If the matrix* A *has zero column, i.e.* $a_{1j} = a_{2j} = \cdots = a_{nj} = 0$ *for some*$1 \leq j \leq n$ *, then* $\det A = 0$.

Property D2

The determinant of a upper triangular matrix is the product of the entries on the diagonal: if $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *and* $a_{i,j} = 0$, *with* $1 \leq i < j \leq n$, *then*

$$\det A = a_{11} a_{22} \cdots a_{nn}.$$

Property D3

The determinant of a lower triangular matrix is the product of the entries on the diagonal: if $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *and* $a_{i,j} = 0$, *with* $n \geq i > j \geq 1$, *then*

$$\det A = a_{11} a_{22} \cdots a_{nn}.$$

Property D4

If $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *and* $A^T = \begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{pmatrix}$, *the transpose of the matrix* A *, then*

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

Property D5
If two rows of the matix A are equal, then $\det A = 0$.

Property D6
If two columns of the matix A are equal, then $\det A = 0$.

Property D7
If two rows or two columns of a matrix are interchanged, then the determinant changes sign.

Property D8.
The determinant has a linearity of each row and column:

$$\begin{vmatrix} a_{11}+a'_{11} & \cdots & a_{1n}+a'_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a'_{11} & \cdots & a'_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{vmatrix} a_{11}+a'_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+a'_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a'_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{vmatrix} ta_{11} & \cdots & ta_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = t \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$\begin{vmatrix} ta_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ ta_{n1} & \cdots & a_{nn} \end{vmatrix} = t \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Property D9
If a multiple of one row is added to another row, the determinant of the new matrix is the same as the old one.

Example 5.4
Evaluate the determinant $\begin{vmatrix} 1 & 3 & 2 \\ 4 & 2 & 10 \\ 8 & 19 & 6 \end{vmatrix}$.

Keep the first row, multiply the first row with (-4) and add to the second row, then multiply the first row with (-8) and add to the third row

$$\begin{vmatrix} 1 & 3 & 2 \\ 4 & 2 & 10 \\ 8 & 19 & 6 \end{vmatrix} \begin{matrix} (-4) & (-8) \\ \downarrow & | \\ & \downarrow \end{matrix}$$

We get

$$\begin{vmatrix} 1 & 3 & 2 \\ 4 & 2 & 10 \\ 8 & 19 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -10 & 2 \\ 0 & -5 & -10 \end{vmatrix}.$$

Expand along the first column, gives

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & -10 & 2 \\ 0 & -5 & -10 \end{vmatrix} = (1)\begin{vmatrix} -10 & 2 \\ -5 & -10 \end{vmatrix} - (0)\begin{vmatrix} 2 & 3 \\ -5 & -10 \end{vmatrix} + (0)\begin{vmatrix} 2 & 3 \\ -10 & 2 \end{vmatrix} = \begin{vmatrix} -10 & 2 \\ -5 & -10 \end{vmatrix}.$$

Keep the first row, multiply the first row with $(-\frac{1}{2})$ and add to the second row

$$\begin{vmatrix} -10 & 2 \\ -5 & -10 \end{vmatrix} \begin{matrix} (-\frac{1}{2}) \\ \downarrow \end{matrix}$$

We get

$$\begin{vmatrix} 1 & 3 & 2 \\ 4 & 2 & 10 \\ 8 & 19 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -10 & 2 \\ 0 & -5 & -10 \end{vmatrix} = \begin{vmatrix} -10 & 2 \\ -5 & -10 \end{vmatrix} = \begin{vmatrix} -10 & 2 \\ 0 & -11 \end{vmatrix} \overset{\text{property } D2}{=} (-10)(-11) = 110.$$

Example 5.5

Evaluate the determinant $\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 7 & 2 \\ 3 & 8 & 1 & 2 \\ 1 & 4 & 3 & 1 \end{vmatrix} = -15.$

$$\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 7 & 2 \\ 3 & 8 & 1 & 2 \\ 1 & 4 & 3 & 1 \end{vmatrix} \begin{matrix} (-2) & (-3) & (-1) & \\ \downarrow & | & | & \\ & \downarrow & | & \\ & & \downarrow & \end{matrix} = \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & -5 & 5 & 0 \\ 0 & -1 & -2 & -1 \\ 0 & 1 & 2 & 0 \end{vmatrix} =$$

$$(-5) \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & -2 & -1 \\ 0 & 1 & 2 & 0 \end{vmatrix} \begin{matrix} (1) & (-1) \\ \downarrow & | \\ & \downarrow \end{matrix} = (-5) \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -3 & -1 \\ 0 & 0 & 3 & 0 \end{vmatrix} \begin{matrix} \\ \\ \\ (1) \\ \\ \downarrow \end{matrix} =$$

$$(-5) \begin{vmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -3 & -1 \\ 0 & 0 & 0 & -1 \end{vmatrix} = (-5)\,(1)\,(1)\,(-1)\,(-3) = -15.$$

Lecture 6

Permutations: Definitions. Cycles. Transpositions. Sign of the permutations

Definitions. Cycles

Definition 6.1
Let $\Sigma_n = \{1, 2, ..., n\}$ *be the set of integers from* 1 *to* n. *A function* $\pi : \Sigma_n \longrightarrow \Sigma_n$ *is called a permutation of* Σ_n *if* π *is bijection on* Σ_n. *The set of all permutations of* Σ_n *will be denoted by* S_n.

We can describe a permutation π by listing the integers $i \in \Sigma_n$ and the corresponding values $\pi(i)$:

$$\begin{pmatrix} 1 & 2 & ... & n \\ \pi(1) & \pi(2) & ... & \pi(n) \end{pmatrix}.$$

The identity permutation, denoted ε, is the permutation which represents i to i :

$$\varepsilon = \begin{pmatrix} 1 & 2 & ... & n \\ 1 & 2 & ... & n \end{pmatrix}$$

Definition 6.2
A permutation $\sigma \in S_n$ *is a* k*-cycle if there exist* k *distinct integers* $i_1, i_2, ..., i_k$ *from* Σ_n *such that* $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \ldots, \sigma(i_k) = i_1$ *and* $\sigma(i) = i$ *for all* $i \notin \{i_1, i_2, ..., i_k\}$. *We will write* $(i_1, i_2, ..., i_k)$ *to denote the cycle* σ.

Let $\sigma = (i_1, i_2, ..., i_k)$ be the k-cycle . The notation for σ we can also write in k different ways:

$$\begin{gathered} \sigma = (i_2, i_3, ..., i_k, i_1) \\ \sigma = (i_3, ..., i_k, i_1, i_2) \\ \cdots \\ \sigma = (i_k, i_1, i_2 ..., i_{k-1}). \end{gathered}$$

Definition 6.3
Two cycles $\sigma_1 = (i_1, i_2, ..., i_k)$ *and* $\sigma_2 = (j_1, j_2 \ldots, j_l)$ *are disjoint if* $i_r \neq i_s$ *for all* r *and* s.

Theorem 6.4
Any permutation in S_n *can be expressed as a product of disjoint cycles.*

Proof. Let $\pi \in S_n$, $i_1 \in \Sigma_n$ and define the finite set

$$C_1 = \{i_1, \pi(i_1), \pi^2(i_1), ...\},$$

where $\pi^t(i) = \pi(\pi^{t-1}(i))$ for all positive integers t .
Now let i_2 be the integer in Σ_n that is not in C_1 and define the finite set

$$C_2 = \{i_2, \pi(i_2), \pi^2(i_2), ...\}.$$

Now let i_3 be the integer in Σ_n that is not in $C_1 \cup C_2$ and define the finite set

$$C_3 = \{i_3, \pi(i_3), \pi^2(i_3), ...\}.$$

We continue in this way to define finite disjoint sets $C_4, ...$ Since Σ_n is a finite set, this process will end and there will be a finite number of these sets: $C_1, C_2 \ldots, C_m$. Define the disjoint cycles

$$\sigma_1 = (i_1, \pi(i_1), \pi^2(i_1), ...), \sigma_2 = (i_2, \pi(i_2), \pi^2(i_2), ...), \ldots, \sigma_m = (i_m, \pi(i_m), \pi^2(i_m), ...).$$

Then $\pi = \sigma_1\sigma_2 \cdots \sigma_m$. This expression is called the disjoint cycle decomposition of π.

△

Example 6.5
Let $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 3 & 2 & 7 & 8 & 6 & 5 \end{pmatrix}$.
We start a cycle with 1, since $\pi(1) = 4, \pi(4) = 2, \pi(2) = 1$ we obtain the cycle $(1, 4, 2)$. Because the integer 3 not used so far, we obtain the cycle (3) because $\pi(3) = 3$. The integer 5 not used so far, therefore we obtain the cycle $(5, 7, 6, 8)$, since $\pi(5) = 7, \pi(7) = 6, \pi(6) = 8, \pi(8) = 5$.
The disjoint cycle decomposition of π is $\pi = (1, 4, 2)(3)(5, 7, 6, 8)$.

.

Definition 6.6

Let π and ρ be permutations in S_n. A composition of the functions π and ρ is called the product of π and ρ, written $\pi\rho$. That is $\pi\rho(i) = \pi(\rho(i))$, for all $i \in \Sigma_n$.

Remark 6.7

If π and ρ are disjoint cycles, then $\pi\rho = \rho\pi$.

The next example shows that the multiplication on S_n is not commutative.

Example 6.8

If $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ and $\rho = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, then

$$\rho\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \pi\rho.$$

We summarize some properties of permutations which hold under multiplication.

Proposition 6.8

Let π, ρ and τ be permutations in S_n.
S1 Associativity of multiplication : $\tau(\rho\pi) = (\tau\rho)\pi$:

$$(\tau(\rho\pi))(i) = \tau((\rho\pi)(i)) = \tau(\rho(\pi(i))) = (\tau\rho)(\pi(i)) = ((\tau\rho)\pi)(i).$$

S2 *Property of identity* : $\varepsilon\pi = \pi\varepsilon = \pi$
S3 Multiplicative inverse : if

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

then

$$\pi^{-1} = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix},$$

and

$$\pi\pi^{-1} = \pi^{-1}\pi = \varepsilon.$$

Transpositions. Sign of the permutations

Definition 6.9
A 2-cycle (i_1, i_2) is called a transposition.

Theorem 6.10
If $n \geq 2$, then any permutation in S_n can be written as a product of transpositions.

Proof. Any cycle can be written as a product of transpositions as follows:

$$(i_1, i_2, i_3, ..., i_{k-1}, i_k) = (i_1, i_2)(i_2, i_3)\cdots(i_{k-1}, i_k).$$

Since each permutaton is a product of cycles, we can obtain each permutation as a product of transpositions.

△

As we can see, there is no unique way to represent permutation as the product of transpositions. For instance, $(1,2) = (1,2)(1,2)(1,2)$.

Proposition 6.11
If $\pi \in S_n$ can be written as a product of r transpositions and if the same π can be written as a product of s transpositions, then r and s is either both even or both odd. A permutation π is called even if π is a product of an even number of transpositions and odd if it is a product of an odd number of transpositions.
The proposition is more difficult to prove and we omit the proof.

Definition 6.12
If the permutation $\pi \in S_n$ can be written as a product of transpositions as follow $\pi = \tau_1\tau_2\cdots\tau_k$, then the sign of the permutation π, written $sgn(\pi)$, is defined by $sgn(\pi) = (-1)^k$.

The *sgn* function satisfies the following properties.

Corollary 6.13
1. $sgn(\varepsilon) = 1$ and every transposition (i_1, i_2) has sign -1, $sgn(i_1, i_2) = -1$.

2. If a permutation π is even, then $sgn(\pi) = 1$, and if a permutation π is odd, then $sgn(\pi) = -1$.

3. For π and $\rho \in S_n$, $sgn(\pi\rho) = sgn(\pi)\, sgn(\rho)$.

4. For all $\pi \in S_n$, $sgn(\pi) = sgn(\pi^{-1})$.

5. The sign of the k-cycle σ is $sgn(\sigma) = (-1)^{k+1}$.

Without proof.

Definition 6.14

Let A_n denote the set of all even permutation. The set A_n is called the alternating set on Σ_n.

Proposition 6.15

For $n \geq 2$, the number of even permutations in S_n is equal to the number of odd permutations. So that the number of even permutations is $\frac{n!}{2}$(the alternating set A_n has $\frac{n!}{2}$ elements).

Proof. Let B_n be the set odd permutations in S_n. Fix a transposition τ in S_n.Denote the set

$$U_\tau = \{\pi \in S_n | \pi = \rho\tau, \rho \in A_n\}.$$

The set U_τ is the subset of B_n and let $\pi_1 = \rho_1\tau$ and $\pi_2 = \rho_2\tau$ be two permutations of U_τ. Suppose that $\pi_1 = \pi_2$. Then $\rho_1\tau = \rho_2\tau$ and so

$$\rho_1 = \rho_1\tau\tau = \rho_2\tau\tau = \rho_2.$$

Therefore, the number of even permutations in S_n, denote m, is not more than the number of all odd permutations in $S_n : m \leq n - m$ and $2m \leq n$.

Conversely denote the set

$$V_\tau = \{\pi \in S_n | \pi = \rho\tau, \rho \in B_n\}.$$

The set V_τ is the subset of A_n and let $\pi_1 = \rho_1\tau$ and $\pi_2 = \rho_2\tau$ be two permutations of V_τ. Suppose that $\pi_1 = \pi_2$. Then $\rho_1\tau = \rho_2\tau$ and so

$$\rho_1 = \rho_1\tau\tau = \rho_2\tau\tau = \rho_2.$$

Therefore, the number of odd permutations in S_n is not more than the number of all even permutations in $S_n : n - m \leq m$ and $n \leq 2m$.

Summarizing, we have $n = 2m$. So that $m = \frac{n!}{2}$.

△

Lecture 7

Systems of linear equations. Gaussian elimination. Solving linear equations

Systems of linear equations

We begin with a definition.

Definition 7.1
A linear equation in n unknowns $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1+a_2x_2+\cdots+a_nx_n= b,$$

where $\alpha_1, \alpha_2, ..., \alpha_n$ and b are real constants.

A system of m linear equations in in n unknowns $x_1, x_2, \ldots, x_n$ is a set of linear equations

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21} + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \qquad (1)$$

If all the $b_i = 0$ then the system (1) is called *homogeneous*.

We say that a system (1) has a solution if there exist a sequence of numbers $\alpha_1, \alpha_2, ..., \alpha_n$ which satisfy each of the equations:

$$\begin{array}{c} a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n = b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n = b_2 \\ \cdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \cdots + a_{mn}\alpha_n = b_m \end{array} .$$

A system of equations is called *consistent* if it has a solution. Otherwise the system is called *inconsistent*.

The matrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

is called the coefficient matrix of the system. The matrix

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ & & \cdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \qquad (2)$$

is called the augmented matrix of the system.

Definition 7.2 (Elementary row operations)

There are three types of elementary row operations of an augmented matrix (2) *corresponding to the equations in the associated system* (1)*:*

1. Interchanging two rows of (2) *corresponds to interchanging two equations in* (1) *.*

2. Multiplying the row by a nonzero constant correspomds to multiplying the equation in (1) *by the same nonzero constant.*

3. Adding a multiple of one row to another row in (2) *corresponds to adding the same multiple of the respective equation to another respective equation.*

Matrix A is row equivalent to matrix B if B is obtained from A by a sequence of elementary row operations.

Proposition 7.3

Elementary row operations do not change the set of solutions of the system: if A and B are row equivalented matrices of two system of linear equations, then the two systems have the same set of solutions.

Without proof.

Definition 7.4 (Row-echelon form)

1. The zero matrix of any size is in row-echelon form

2. A nonzero matrix

$$\begin{array}{c} \\ \\ \left(\begin{array}{ccccccccccccc} & & & j_1 & & \cdots & j_2 & \cdots & & j_r & & & \\ 0 & \cdots & 1 & \cdots & & a_{1j_2} & \cdots & & a_{1j_r} & & & a_{1n} \\ 0 & \cdots & 0 & \cdots & 0 & 1 & \cdots & & & & & a_{2n} \\ \cdots & & \cdots & & & \cdots & & & & & & \\ 0 & \cdots & & & & & & 0 & 1 & \cdots & & a_{rn} \\ 0 & \cdots & & & & & & & 0 & \cdots & 0 & 0 \\ \cdots & & \cdots & & & & & \cdots & & & & \\ 0 & \cdots & & & & & & & 0 & \cdots & 0 & 0 \end{array}\right) \end{array} \qquad (3)$$

where

$$\textit{1)}\ 1 \le j_1 < j_2 < ... < j_r \le n;$$
$$\textit{2)}\ a_{1j_1} =, a_{2j_2} = \cdots = a_{rj_r} = 1;$$
$$\textit{3)}\ a_{is_i} = 0,\ \textit{with}\ s_i < j_i\ ,\ 1 \le i \le r;$$
$$\textit{4)}\ a_{ti} = 0,\ \textit{with}\ t > r,\ 1 \le i \le n,$$

is in row-echelon form.

3. *Let a matrix A be row equivalent to the matrix in row-echelon form* (3). *Then* rank$A = r$.

Example 7.5

The follwing matrices are in row-echelon form:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The matrix $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 5 & -1 \\ 0 & 0 & 1 & 6 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ is not in row-echelon form.

Proposition 7.6

If a matrix A is in row echelon form, then

$$\left(\begin{array}{c|c} 0 \cdots 0 & A \end{array} \right)$$

and

$$\begin{pmatrix} 1 & | & * \\ 0 & | & A \\ \vdots & & \\ 0 & & \end{pmatrix}$$

are in row-echelon form.
Without proof.

Gaussian elimination

Theorem 7.7 (Gaussian elimination)
Let A be a given matrix . Then A is row equivalent to matrix B which is in row-echelon form.

Proff. Let A is a matrix that has n columns. We will use induction by n.
For $n = 1$ the matrix A is

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$$

Consider two cases.
(1) If $a_{11} = a_{21} = \cdots = a_{m1} = 0$ then the matrix A is in row echelon form.
(2) If $a_{i1} \neq 0$ for some $i, 1 \leq i \leq m$, then we use the following steps:
(i) interchange the top row with i-th row.
(ii) multiply the first row of the preceding matrix by $\frac{1}{a_{i1}}$.
(iii) for all $j, 2 \leq j \leq m$, add $(-a_{j1})$ times the first row of the preceding matrix to the j-th row.
After these steps we have the following matrix B in row-echelon form:

$$B = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Asume now that the theorem is true for $n-1$. We will show that it is true for the matrix A with n rows:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Consider two cases.

(1) If $a_{11} = a_{21} = \cdots = a_{m1} = 0$ then

$$A = \left(\begin{array}{c|ccc} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & & A' & \\ 0 & & & \end{array}\right),$$

where A' is the matrix with $n-1$ rows and by induction there exist such matrix B' in row-echelon form that A' is row equivalent to B'. Then, by Proposition 7.6, the matrix A is row equivalent to the matrix B wich is in row-echelon form:

$$B = \left(\begin{array}{c|ccc} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & & B' & \\ 0 & & & \end{array}\right).$$

(2) If $a_{i1} \neq 0$ for some $i, 1 \leq i \leq m$, then we use the following steps:

(i) interchange the top row with i-th row.

(ii) multiply the first row of the preceding matrix by $\frac{1}{a_{i1}}$.

(iii) for all $j, 2 \leq j \leq m$, add $(-a_{j1})$ times the first row of the preceding matrix to the j-th row.

After these steps we have the following matrix A_1:

$$A_1 = \left(\begin{array}{c|ccc} 1 & a_{i2} & \cdots & a_{in} \\ 0 & & & \\ \vdots & & A'_1 & \\ 0 & & & \end{array}\right),$$

where A'_1 is the matrix with $n-1$ rows and by induction there exist such matrix B' in row-echelon form that A' is row equivalent to B'. Then, by Proposition 7.6, the matrix A is row equivalent to the matrix B wich is in row-echelon form:

$$B = \left(\begin{array}{c|ccc} 1 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right).$$

△

Solving linear equations

We describe the Gauss algorithm to solve a system of linear equations.

This algorithm starts with the augmented matrix A (2) of the system of linear equations (1) and using Gaussian eliminations reduces to the matrix B in row-echelon form, wich is row-equivalent to A :

$$\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\!\!\begin{array}{c} \begin{array}{cccccccccccccc} & & & j_1 & \cdots & & j_2 & \cdots & & j_r & \cdots & & n & & \end{array} \\ \left(\begin{array}{cccccccccccccc} 0 & \cdots & \mathbf{1} & \cdots & & a'_{1j_2} & \cdots & & & a'_{1j_r} & \cdots & & a'_{1n} & | & b'_1 \\ 0 & \cdots & 0 & & 0 & \mathbf{1} & \cdots & & & a'_{2j_r} & \cdots & & a'_{2n} & | & b'_2 \\ \cdots & & \cdots & & & \cdots & & & & & & & & | & \\ 0 & \cdots & & & & & & & 0 & \mathbf{1} & \cdots & & a'_{rn} & | & b'_r \\ 0 & \cdots & & & & & & & & 0 & \cdots & 0 & 0 & | & b'_{r+1} \\ \cdots & & \cdots & & & & & & \cdots & & & & & | & \\ 0 & \cdots & & & & & & & & 0 & \cdots & 0 & 0 & | & 0 \end{array}\right) \end{array}.$$

The corresponding system of equations is

$$\begin{array}{rcccccc} x_{j_1}+ & & & \cdots & +a'_{1n}x_n & = & b'_1 \\ & x_{j_2}+ & & \cdots & +a'_{2n}x_n & = & b'_2 \\ & & & \cdots & & & \\ & & x_{j_r}+ & \cdots & +a'_{rn}x_n & = & b'_r \\ & & & & 0 & = & b_{r+1} \end{array} \qquad (4)$$

Consider three caces.

Case 1. If $b'_{r+1} \neq 0$, then the system is inconsistent: the last equation is

$$0x_1 + \cdots + 0x_n = b'_{r+1}$$

which has no solution.

Case 2. If $c_{r+1} = 0$ and $r = n$, then the system is consistent and has a unique solution:

$$\begin{array}{c} x_n = b'_n, \\ x_{n-1} = b'_{n-1} - a'_{n-1,n}b'_n, \\ \cdots, \\ x_1 = b'_1 - \cdots - a'_{1n}b'_n. \end{array}$$

Case 3. If $b'_{r+1} = 0$ and $r < n$, then the system is consistent and has more than one solution: the dependent unknowns $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ are expressed in terms of the remaining independent unknowns $x_1, \ldots, x_{j_1-1}, x_{j_1+1}, \ldots, x_{j_r-1}, x_{j_r+1}, \ldots, x_n$:

$$x_{j_1} = b'_1 + a''_{1j_1+1}x_{j_1+1} + \ldots + a''_{1j_r-1}x_{j_r-1} + a''_{1j_r+1}x_{j_r+1} + \ldots + a''_{1n}x_n$$
$$\ldots$$
$$x_{j_r} = b'_r + a''_{rj_r+1}x_{j_r+1} + \ldots + a''_{rn}x_n.$$

Corollary 7.8

A homogeneous system of m linear equations in n unknowns always is consistent and has a non-trivial solution if $m < n$.

Theorem 7.9 (L.Kronecker-A.Capelli)

The system of linear equations (1) *is consistent if and only if the rank of the coefficient matrix is equal to the rank of the augmented matrix* (2) :

$$\operatorname{rank}\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \operatorname{rank}\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ & \cdots & & \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

Without proof.

Corollary 7.10

1. If the system of m linear equations in n unknowns (1) *is consistent and the rank of the coefficient matrix is equal n, then the system has a unique solution.*

2. If the system of m linear equations in n unknowns (1) *is consistent and the rank of the coefficient matrix is less than n, then the system has more than one solution solution.*

Lecture 8

Matrices: definitions, arithmetic operations on matrices, inverses

Arithmetic operations on matrices

Definition 8.1
A matrix over a field $\mathbf{F}$ *is a rectangu twolar array of elements from* $\mathbf{F}$.

We shall use capital letters to denote matrices. The equation

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

means that the matrix A has m rows and n columns, A is called an $m \times n$ matrix, and the element in the i-th row and j-th column of the matrix A equals a_{ij} (also written $a_{ij} = (A)_{ij}$). We will denote by $\mathbf{M}_{m\times n}(\mathbf{F})$ the set of all $m \times n$ matrices over $\mathbf{F}$. A matrix A with n rows and n columns is called a *square matrix of order* n.

Definition 8.2
Two matrices A *and* B *are said to be equal if they have the same size, that is* $A, B \in M_{m\times n}(\mathbf{F})$, *and* $(A)_{ij} = (B)_{ij}$ *for* $1 \leq i \leq m, 1 \leq j \leq n$.

We shall consider the arithmetic operations on matrices.

Definition 8.3
Let $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ *and* $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$ *be two matrices, both have* m *rows and* n *columns. Then the sum* $A + B$ *is the matrix* $A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$, *i.e* . $(A + B)_{ij} = (A)_{ij} + (B)_{ij}$ *and*

the difference $A-B$ *is the matrix* $A-B=\begin{pmatrix} a_{11}-b_{11} & \cdots & a_{1n}-b_{1n} \\ \vdots & & \vdots \\ a_{m1}-b_{m1} & \cdots & a_{mn}-b_{mn} \end{pmatrix}$, *i.e.* $(A-B)_{ij}=(A)_{ij}-(B)_{ij}$.

Definition 8.4

Let $A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M_{m\times n}(\mathbf{F})$ *and* $a \in F$. *Then the product of the matrix* A *by the scalar* aA *is the matrix* $aA=\begin{pmatrix} aa_{11} & \cdots e & aa_{1n} \\ \vdots & & \vdots \\ aa_{m1} & \cdots & aa_{mn} \end{pmatrix}$, *i.e.* $(aA)_{ij}=a(A)_{ij}$.

Definition 8.5

The matrix O *in* $M_{m\times n}(\mathbf{F})$, *all of whose elements are zero, is called the zero matrix, i.e.* $(O)_{ij}=0$.

Proposition 8.6

Suppose that A,B,C *are matrices in* $M_{m\times n}(\mathbf{F})$ *and* a,b *are scalars in* $\mathbf{F}$. *Then*
V1.(Associativity of addition) $(A+B)+C=A+(A+C)$.
V2.(Commutativity of addition) $A+B=B+A$.
V3.(Property of zero) $O+A=A$.
V4.(Additionative inverse) $A+(-A)=O$.
V5.(Distributivity) $(a+b)A=aA+bA$.
V6.(Distributivity) $a(A+B)=aA+aB$.
V7. (Associativity)$(ab)A=a(bA)$.
V8.(*Property of identity*) $1\cdot A=A$.
V9.(Property of zero) $A0=0A=O$.
Without proof.

The system of linear equations

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

is equivalent to the matrix equation:

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

Definition 8.7

Let $A = (a_{ij}) \in M_{m\times n}(\mathbf{F})$ *and* $B = (b_{ij}) \in M_{n\times r}(\mathbf{F})$ *(the number of colums of the first matrix* A *must be equal to the number of rows of the second matrix* B*). Then the product* AB *is the* $m \times r$ *matrix* $AB = \begin{pmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & & \vdots \\ c_{m1} & \cdots & c_{mr} \end{pmatrix}$ *where* $c_{ij} = (AB)_{ij} = \sum_{s=1}^{n} (A)_{is} (B)_{sj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$ *for all* $i = 1, \ldots, m$ *and* $j = 1, ..., r$.

Example 8.8

1. $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 3 & -4 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 3 + 3\cdot 1 & 1\cdot(-1) + 2\cdot(-4) + 3\cdot 3 \end{pmatrix} = \begin{pmatrix} 10 & 0 \end{pmatrix}$.

2. The product $\begin{pmatrix} 1 & -1 \\ 3 & -4 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ is not defined.

3. $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 4 & 7 \end{pmatrix} = \begin{pmatrix} 11 & 41 \\ 5 & 16 \end{pmatrix}$

4. $\begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -3 & 2 \\ 4 & 7 \end{pmatrix} \neq \begin{pmatrix} -3 & 2 \\ 4 & 7 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -7 & -11 \\ 19 & 34 \end{pmatrix}$.

The system of linear equations

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

is equivalent to the matrix equation:

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

Proposition 8.9

Suppose that A, B, C, D, E *are matrices and* a, *be a scalar in* $\mathbf{F}$. *Then*

S1. (Associativity) $a(AB) = (aA)B = A(aB)$ if $A \in M_{m\times n}(\mathbf{F})$ *and* $B \in M_{n\times r}(\mathbf{F})$.

S2. (Associativity of multiplication) $(AB)C = A(BC)$ if $A \in M_{m\times n}(\mathbf{F})$, $B \in M_{n\times r}(\mathbf{F})$ and $C \in M_{r\times s}(\mathbf{F})$

S3. (Distributivity) $(A+D)B = AB + DB$ if $A, D \in M_{m\times n}(\mathbf{F})$ *and* $B \in M_{n\times r}(\mathbf{F})$.

S4. (Distributivity) $E(A+D) = EA + ED$ if $E \in M_{k\times m}(\mathbf{F})$ and $A, D \in M_{m\times n}(\mathbf{F})$.

Proof of S2.

$$((AB)C)_{ij} = \sum_{u=1}^{r} (AB)_{iu} \cdot (C)_{uj} = \sum_{u=1}^{r} \left(\sum_{v=1}^{n} (A)_{iv} (B)_{vu} \right) \cdot (C)_{uj} =$$
$$\sum_{\mathbf{u=1}}^{\mathbf{r}} \left(\sum_{v=1}^{n} ((A)_{iv} (B)_{vu}) \cdot (C)_{uj} \right) = \sum_{\mathbf{v=1}}^{\mathbf{n}} \left(\sum_{u=1}^{r} (A)_{iv} \cdot \left((B)_{vu} (C)_{uj} \right) \right) =$$
$$\sum_{\mathbf{v=1}}^{\mathbf{n}} \left((A)_{iv} \cdot \sum_{u=1}^{r} \left((B)_{vu} (C)_{uj} \right) \right) = \sum_{\mathbf{v=1}}^{\mathbf{n}} \left((A)_{iv} \cdot (BC)_{vj} \right) = (A(BC))_{ij}.$$

We used the equality of numbers:

$$\sum_{\mathbf{u=1}}^{\mathbf{r}} \sum_{v=1}^{n} d_{uv} = \sum_{\mathbf{u=1}}^{\mathbf{r}} (d_{u1} + d_{u2} + \cdots + d_{un}) =$$
$$(d_{11} + d_{12} + \cdots + d_{1n}) + (d_{21} + d_{22} + \cdots + d_{2n}) + \cdots + (d_{r1} + d_{r2} + \cdots + d_{rn}) =$$
$$(d_{11} + d_{21} + \cdots + d_{r1}) + (d_{12} + d_{22} + \cdots + d_{r2}) + \cdots + (d_{1n} + d_{2n} + \cdots + d_{rn}) =$$
$$\sum_{\mathbf{v=1}}^{\mathbf{n}} (d_{1v} + d_{2v} + \cdots + d_{rv}) = \sum_{\mathbf{v=1}}^{\mathbf{n}} \sum_{u=1}^{r} d_{uv}.$$

△

Let $y_1, y_2, ..., y_k$ be the linear combinations of $x_1, x_2, ..., x_n$:

$$
\begin{array}{c}
y_1 = b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\
y_2 = b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
y_k = b_{k1}x_1 + b_{k2}x_k + \cdots + b_{kn}x_n
\end{array}
$$

and $z_1, z_2, ..., z_m$ be the linear combinations of $y_1, y_2, ..., y_k$:

$$
\begin{array}{c}
z_1 = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1k}y_k \\
z_2 = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2k}y_k \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
z_m = a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mk}y_k \ .
\end{array}
$$

Then

$$
\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_k \end{pmatrix} = B \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} \text{ ir } \begin{pmatrix} z_1 \\ z_2 \\ \cdots \\ z_m \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_k \end{pmatrix},
$$

where $A = (a_{ij}) \in \mathbf{M}_{m\times k}$ and $B = (b_{ij}) \in \mathbf{M}_{k\times n}$
and $z_1, z_2, ..., z_m$ be the linear combinations of $x_1, x_2, ..., x_n$:

$$
\begin{array}{c}
z_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1k}x_n \\
z_2 = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2k}x_n \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
z_m = c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mk}x_n
\end{array}
$$

where the matrix $C = \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mk} \end{pmatrix}$ is equal to the product AB :

$$
\begin{pmatrix} z_1 \\ z_2 \\ \cdots \\ z_m \end{pmatrix} = C \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = AB \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.
$$

Inverses

Definition 8.10

The square matrix $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ *in* $M_{n\times n}(F)$ *is called the identity matrix of order* n.

Proposition 8.11

1. *If* $A \in M_{n\times m}(\mathbf{F})$, *then*

$$A \cdot I_m = I_n \cdot A = A \ .$$

2. If $A \in M_{n\times n}(\mathbf{F})$, *then*

$$A \cdot I_n = I_n \cdot A = A \ .$$

Theorem 8.12

If $I_n(\mathbf{F}) = \{a \cdot I_n \in \mathbf{M}_n(\mathbf{F}) \mid a \in \mathbf{F}\}$ *and the function* $e : F \to I_n(\mathbf{F})$ *is defined by* $e(a) = a \cdot I_n$. *Then* e *is a bijective function and* $e(a+b) = e(a) + e(b), e(a \cdot b) = e(a) \cdot e(b)$.

We leave **the proof** to the reader.

Thus, the set of matrices $\mathbf{M}_n(\mathbf{F})$ becomes an extension of the field $\mathbf{F} : \mathbf{F} \subset \mathbf{M}_n(\mathbf{F})$. The set $\mathbf{M}_n(\mathbf{F})$ is a commutative ring, but whether it is a field? We have the following question: is it posible for given square matrix A to find a square matrix B such that $AB = BA = I_n$?

Now answer this question.

Definition 8.13

A square matrix $A \in M_{n\times n}(\mathbf{F})$ *is called inversible if there exists a matrix* $B \in M_{n\times n}(\mathbf{F})$ *such that* $AB = BA = I_n$. *We say that* B *is the inverse of* A *and write* $B = A^{-1}$.

Proposition 8.14 (inverse is unique)

If B *and* C *are inverses of* A, *then* $B = C$.

Proof. Since B and C are inverses of A, then

$$AB = BA = I_n \text{ and } AC = CA = I_n.$$

Thus

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

△

The next theorem list the main properties of the inversible matrices.

Theorem 8.15

If A is a matrix in $M_{n\times n}(\mathbf{F})$, then the following statements are equivalent.

1. $\text{rank}A = n$.

2. The homogeneous system of linear equations $AX = O$, where $X = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}$ and $O = \begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix}$, has the unique solution.

3. $\det A \neq 0$.

4. A is inversible.

5. *The system of linear equations $AX = B$, where $X = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}$ and $B = \begin{pmatrix} b_1 \\ \cdots \\ b_n \end{pmatrix}$, has the solution.*

We need the following proposion.

Proposition 8.16

If $A \in M_{m\times n}(\mathbf{F})$ and $B \in M_{n\times r}(\mathbf{F})$, then

$$\begin{aligned} \text{rank}(AB) &\leq \text{rank}A, \\ \text{rank}(AB) &\leq \text{rank}B. \end{aligned}$$

Without proof.

Proof of theorem 8.16.

From Corollary 7.10.1 we have that $1 \Leftrightarrow 2$.

Now prove that $1 \Leftrightarrow 3$. Assume that the row-echelon form of A is B, so that A can be reduced to B by the following sequence of elementary row operations: k times interchanges two rows; multiplying the rows by the non-zero constants $a_1, ..., a_l$; t times adding a multiple of one row to another row. Thus

$$\det B = (-1)^k \alpha_1 \cdot \cdots \cdot \alpha_l \det A.$$

Now we have the sequence of equivalences:

$$\text{rank}A = n \Leftrightarrow \text{rank}B = n \Leftrightarrow \det B \neq 0 \Leftrightarrow \det A \neq 0.$$

Finally, we shall prove the sequence of implications $1 \Rightarrow 5 \Rightarrow 4 \Rightarrow 1$.
$1 \Rightarrow 5$.
If $\text{rank}A = n$, then

$$n = \text{rank}A \leq \text{rank}(A|B) \leq n.$$

Thus

$$\text{rank}A = \text{rank}(A|B) = n$$

and by Kronecker-Capelli's theorem 7.9 the system of linear equations $AX = B$ has the solution.

$5 \Rightarrow 4$. Let the column X_1 be the solution of the system of linear equations $AX = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}$, the column X_2 be the solution of the system of linear equations $AX = \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}$, ...,the column X_n be the solution of the system of linear equations $AX = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}$ and the square matrix Y be a matrix of the form $Y = (X_1|X_2|\cdots|X_n)$. Thus,

$$AY = (AX_1|AX_2|\cdots|AX_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n.$$

Now we shal show that $YA = I_n$, i.e. $Y = A^{-1}$.
By Proposition 8.17 we know that

$$n = \text{rank}I_n = \text{rank}AY \leq \text{rank}Y \leq n$$

and rank$Y = n$. By the above there exist the matrix Z such that $YZ = I_n$. We thus get

$$YA = YAI_n = YAYZ = YI_nZ = YZ = I_n.$$

$4 \Rightarrow 1$. If A is inversible, then

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

Thus

$$\text{rank}AA^{-1} = \text{rank}I_n = n$$

and

$$n = \text{rank}AA^{-1} \leq \text{rank}A \leq n,$$

finally, we have rank$A = n$ which completes the proof of theorem.

$\triangle$

Lecture 9

Complex numbers 1

Complex numbers

The set of complex numbers $\mathbf{C}$ can be defined with matrices from $M_2(\mathbf{R})$.

Definition 9.1
A complex number z is a matrix of the form

$$\mathbf{z}=\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

where a and b are real numbers.
The real complex number $[a]=\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ is called the real part of z and the real complex number $[b]=\begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix}$ the imaginary part of $\mathbf{z}$. These numbers are denoted by $Re(\mathbf{z})$ and $\mathrm{Im}(\mathbf{z})$.

The complex number $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is denoted by i. Thus we migth write

$$\begin{gathered}\mathbf{z}=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}=\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}+\begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}= \\ =\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}+\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} b & 0 \\ 0 & b \end{pmatrix} \\ =[a]+i[b]\end{gathered}$$

and

$$i^2=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}=\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}=[-1].$$

Definition 9.2
Two complex numbers $[a]+i\,[b]$ *and* $[c]+i\,[d]$ *are defined to be equal,* $[a]+i\,[b] = [c]+i\,[d]$*, if* $a = c$ *and* $b = d$*.*

The sum and product of two real complex numbers are also real complex number

$$[a] + [b] = [a + b] \text{ and } [a]\,[b] = [ab]$$

The set of real complex numbers is a field under the matrix addition and multiplication:

1. *Associativity of addition:* $([a] + [b]) + [c] = [a] + ([b] + [c])$.
2. *Commutativity of addition:* $[a] + [b] = [b] + [a]$.
3. *Property of 0:* $[a] + [0] = [a]$.
4. *Additionative inverse:* $[a] + [-a] = [0]$.
5. *Associativity of multiplication :* $([a]\,[b])\,[c] = [a]\,([b]\,[c])$.
6. *Commutativity of multiplication:* $[a]\,[b] = [b]\,[a]$.
7. *Property of 1* $[a]\,[1] = [a]$.
8. *Multiplicative inverse:* If $a \neq 0$ then $[a]\,[a^{-1}] = [1]$.
9. *Distributivity:* $[a]\,([b] + [c]) = [a]\,[b] + [a]\,[c]$.
10. $[0] \neq [1]$.

Notice that

$$[a] - [b] = [a] + (-\,[b]) = [a] + [-b] = [a - b] \text{ and}$$
$$\tfrac{[a]}{[b]} = [a]\,[b]^{-1} = [a]\,[b^{-1}] = [ab^{-1}] = \left[\tfrac{a}{b}\right].$$

So the real complex number $[a]$ can be identified with the real number a and we write the complex number $\mathbf{z} = [a] + i\,[b]$ as $\mathbf{z} = a + ib$.

The sum and difference of two complex numbers is the complex numbers:

$$(a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i\,(b_1 + b_2)$$
$$(a_1 + ib_1) - (a_2 + ib_2) = (a_1 - a_2) + i\,(b_1 - b_2)$$

The product of complex numbers is defined so that the usual commutative and distributive laws hold:

$$\begin{gathered}(a_1+ib_1)(a_2+ib_2)\\ =a_1(a_2+ib_2)+(ib_1)(a_2+ib_2)=a_1a_2+a_1(ib_2)+(ib_1)a_2+(ib_1)(ib_2)\\ =a_1a_2+ia_1b_2+ib_1a_2+i^2b_1b_2=(a_1a_2+(-1)b_1b_2)+i(a_1b_2+b_1a_2)\\ =(a_1a_2-b_1b_2)+i(a_1b_2+b_1a_2).\end{gathered}$$

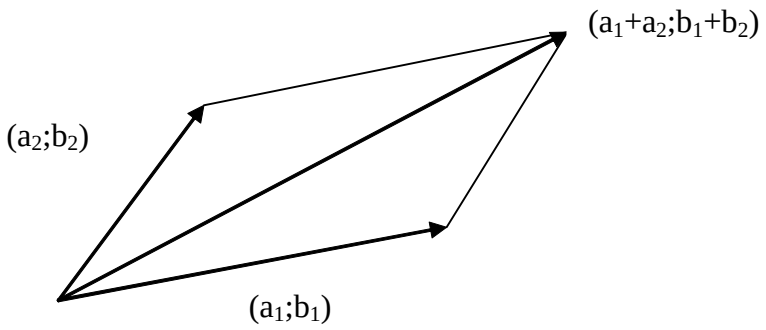
Thus the set of complex numbers $\mathbf{C}$ is a field under the matrix addition and multiplication:

1. *Associativity of addition:* $(\mathbf{z}_1+\mathbf{z}_2)+\mathbf{z}_3=\mathbf{z}_1+(\mathbf{z}_2+\mathbf{z}_3)$.
2. *Commutativity of addition:* $.\mathbf{z}_1+\mathbf{z}_2=\mathbf{z}_2+\mathbf{z}_1$
3. *Property of 0:* $\mathbf{z}+0=\mathbf{z}$.
4. *Additionative inverse:* $\mathbf{z}+(-\mathbf{z})=0$.
5. *Associativity of multiplication :*$(\mathbf{z}_1\mathbf{z}_2)\mathbf{z}_3=\mathbf{z}_1(\mathbf{z}_2\mathbf{z}_3)$.
6. *Commutativity of multiplication:* $\mathbf{z}_1\mathbf{z}_2=\mathbf{z}_2\mathbf{z}_1$.
7. *Property of 1* : $\mathbf{z}\cdot 1=\mathbf{z}$.
8. *Multiplicative inverse:* If $\mathbf{z}=a+ib\neq 0$ then $\mathbf{z}^{-1}=\frac{a}{a^2+b^2}+i\frac{-b}{a^2+b^2}$ and $\mathbf{z}\mathbf{z}^{-1}=1$.
9. *Distributivity:* $(\mathbf{z}_1+\mathbf{z}_2)\mathbf{z}_3=\mathbf{z}_1\mathbf{z}_3+\mathbf{z}_2\mathbf{z}_3$.
10. $[0]\neq[1]$.

Geometric representation of complex numbers

The complex number $\mathbf{z}=a+ib$ can also be represented by the vector $(a;b)$ in the plane $\mathbf{R}^2$(called the Argand plane). Thus, the complex number $i=0+1\cdot i$ is identified with the vector $(0;1)$. Just as vectors in $\mathbf{R}^2$ are added or subtracted by adding or subtracting corresponding components, so complex numbers are added or subtracted by adding or subtracting their real parts and their imaginary parts:

$$(a_1;b_1)\pm(a_2;b_2)=(a_1\pm a_2;b_1\pm b_2)$$



If a complex number $\mathbf{z} = a + ib$ is viewed as a vector in $\mathbf{R}^2$, then the length of the vector, written $|\mathbf{z}|$, is called the modulus of $\mathbf{z} : |\mathbf{z}| = \sqrt{a^2 + b^2}$.

Let $\mathbf{z} = a + ib$ be a non-zero complex number, $r = |\mathbf{z}| = \sqrt{a^2 + b^2} > 0$. Then we have $a = r\cos\varphi, b = r \cdot \sin\varphi$, where φ is the angle fom the positive real axis to the vector $\mathbf{z}$. Thus, the complex number $\mathbf{z}$ can be written as

$$z = [r, \varphi] = r\cos\varphi + ir\sin\varphi = r\left(\cos\varphi + i\sin\varphi\right).$$

where $r = |\mathbf{z}| = \sqrt{a^2 + b^2}$ and $\tan\varphi = \dfrac{b}{a}$. This representation of $\mathbf{z}$ is called *trigonometric form of* $\mathbf{z}$.

The angle φ is called *the argument* of $\mathbf{z}$ and is denoted by $\varphi = \arg\mathbf{z}$. The argument of $\mathbf{z}$ is not uniquely; any two arguments of $\mathbf{z}$ differ an integer multiple of 2π. The argument of $\mathbf{z}$ that satisfies $-\pi < \varphi \leq \pi$ is called the principal argument of $\mathbf{z}$ and is denoted Arg $\mathbf{z}$:

$$\text{Arg } \mathbf{z} = \arctan\frac{b}{a} + \pi k,$$

where

$$\begin{gathered} k = 0 \text{ if } a > 0 \text{ and } b > 0, \\ k = 1 \text{ if } a < 0 \text{ and } b > 0, \\ k = -1 \text{ if } a < 0 \text{ and } b < 0, \end{gathered}$$

Definition 9.2

Two non-zero complex numbers $\mathbf{z}_1 = [r_1, \varphi_1]$ *and* $\mathbf{z}_2 = [r_2, \varphi_2]$ *are defined to be equal,* $\mathbf{z}_1 = \mathbf{z}_2$*, if* $r_1 = r_2 \quad ir \quad \varphi_1 = \varphi_2 + k \cdot 2\pi, k \in \mathbf{Z}$.

We show how trigonometric form can be used to give geometric interpretations of multiplication of complex numbers.

Let

$$\begin{array}{c} \mathbf{z}_1 = [r_1, \varphi_1] = r_1 \cos \varphi_1 + i r_1 \sin \varphi_1 \\ \mathbf{z}_2 = [r_2, \varphi_2] = r_2 \cos \varphi_2 + i r_2 \sin \varphi_2 \end{array}$$

Then

$$\begin{array}{c} \mathbf{z}_1 \mathbf{z}_2 = \\ r_1 r_2 \left((\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i \left(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2 \right) \right) = \\ r_1 r_2 \left(\cos \left(\varphi_1 + \varphi_2 \right) + i \sin \left(\varphi_1 + \varphi_2 \right) \right). \end{array}$$

We obtain

$$\mathbf{z}_1 \mathbf{z}_2 = \left[r_1 r_2, \varphi_1 + \varphi_2 \right].$$

Thus we have shown that

$$\begin{array}{ll} |\mathbf{z}_1 \mathbf{z}_2| & = |\mathbf{z}_1| \cdot |\mathbf{z}_2| \\ \arg \left(\mathbf{z}_1 \mathbf{z}_2 \right) & = \arg \mathbf{z}_1 + \arg \mathbf{z}_2 \end{array} :$$

an argument of the product of two complex numbers is the sum of their arguments and a modulus of the product of two complex numbers is the product of their moduli.

Let $\mathbf{z}_1 = a_1 + i b_1 = \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}$ and $\mathbf{z}_2 = a_2 + i b_2 = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$. Then

$$\begin{array}{c} z_1 \cdot \ z_2 = \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \\ \begin{pmatrix} a_1 a_2 - b_1 b_2 \\ a_1 b_2 + a_2 b_1 \end{pmatrix} = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1) \, i. \end{array}$$

Let $\mathbf{z} = r \cos \theta + i r \sin \theta = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ be the non-zero complex numbet. The complex number $\mathbf{z}_\varphi = \cos \varphi + i \sin \varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$ has a modulus 1 and an argument φ, so the product

$$\mathbf{z}_\varphi \mathbf{z} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} =$$
$$\begin{pmatrix} r\cos\theta\cos\varphi - r\sin\theta\sin\varphi \\ r\sin\theta\cos\varphi + r\cos\theta\sin\varphi \end{pmatrix} = \begin{pmatrix} r\cos(\theta+\varphi) \\ r\sin(\theta+\varphi) \end{pmatrix}$$

has the same modulus as $\mathbf{z}$ and its argument is $\theta + \varphi$: multiplying $\mathbf{z}$ by $\mathbf{z}_\varphi$ rotates $\mathbf{z}$ counterclockwise by φ.The matrix $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ is called *a rotation by φ matrix.*

Definition 9.3
If $z = a + ib$, then the complex conjugate of z is the complex number defined by $\bar{z} = a - ib$.

We list some of the properties of the complex conjugate in the following proposition. The proof follow from the definiton.

Propsition 9.4
1. $\overline{\mathbf{z}_1+\mathbf{z}_2} = \bar{\mathbf{z}}_1 + \bar{\mathbf{z}}_2$.
2. $\overline{\mathbf{z}_1\cdot\mathbf{z}_2} = \bar{\mathbf{z}}_1 \cdot \bar{\mathbf{z}}_2$.
3. $\overline{(\bar{\mathbf{z}})} = \mathbf{z}$.
4. $\mathbf{z}\cdot\bar{\mathbf{z}} = |\mathbf{z}|^2$.

Let now $\mathbf{z} = [r, \varphi]$ and $\mathbf{z}^{-1} = [q, \psi]$. Then

$$\mathbf{z}\cdot\mathbf{z}^{-1} = [r,\varphi]\cdot[q,\psi] = [r\cdot q, \varphi+\psi] = 1 = [1,0],$$

i.e.

$$r\cdot q = 1, \qquad \varphi + \psi = 0 + 2\pi k, \quad k \in \mathbf{Z},$$

and

$$q = r^{-1}, \qquad \psi = -\varphi + 2\pi k, \quad k \in \mathbf{Z} \text{ for example } \psi = -\varphi.$$

Thus

$$\mathbf{z} = r(\cos\varphi + i\sin\varphi), r \neq 0 \Rightarrow$$
$$\left\{ \begin{array}{c} \mathbf{z}^{-1} = r^{-1}(\cos\varphi - i\sin\varphi) \\ \bar{\mathbf{z}} = r(\cos\varphi - i\sin\varphi) \end{array} \right\} \Rightarrow$$
$$\mathbf{z}\cdot\bar{\mathbf{z}} = r^2 \Rightarrow$$
$$\mathbf{z}^{-1} = \frac{1}{\mathbf{z}} = \frac{\bar{\mathbf{z}}}{\mathbf{z}\cdot\bar{\mathbf{z}}} = \frac{\bar{\mathbf{z}}}{r^2} = \frac{1}{r^2}\bar{\mathbf{z}}.$$

Geometrically, the multiplicative inverse $\mathbf{z}^{-1}$ is obtained by reflecting $\mathbf{z}$ in the real axis and stetching by a factor of $\dfrac{1}{r^2}$.

We show how trigonometric form can be used to give geometric interpretations of division of complex numbers.

Let

$$\begin{aligned}\mathbf{z}_1 &= [r_1, \varphi_1] = r_1 \cos\varphi_1 + ir_1 \sin\varphi_1 \\ \mathbf{z}_2 &= [r_2, \varphi_2] = r_2 \cos\varphi_2 + ir_2 \sin\varphi_2\end{aligned}$$

Then

$$\frac{\mathbf{z}_1}{\mathbf{z}_2} = \mathbf{z}_1 \cdot \mathbf{z}_2^{-1} = \mathbf{z}_1 \cdot \frac{1}{r_2^2}\overline{\mathbf{z}}_2$$

$$\frac{r_1}{r_2}\left(\left(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2\right) + i\left(\cos\varphi_1\sin\varphi_2 - \sin\varphi_1\cos\varphi_2\right)\right) =$$

$$r_1 r_2\left(\cos\left(\varphi_1 - \varphi_2\right) + i\sin\left(\varphi_1 - \varphi_2\right)\right).$$

We obtain

$$\frac{\mathbf{z}_1}{\mathbf{z}_2} = \left[\frac{r_1}{r_2}, \varphi_1 - \varphi_2\right].$$

Thus we have shown that

$$\begin{aligned}\left|\frac{\mathbf{z}_1}{\mathbf{z}_2}\right| &= \frac{|\mathbf{z}_1|}{|\mathbf{z}_2|} \\ \arg\left(\frac{\mathbf{z}_1}{\mathbf{z}_2}\right) &= \arg\mathbf{z}_1 - \arg\mathbf{z}_2\end{aligned} \quad :$$

an argument of the quotient of two complex numbers is the subtract of their arguments and a modulus of the quotient of two complex numbers is the quotient of their moduli.

If n is a positive integer and $\mathbf{z} = r\left(\cos\varphi + i\sin\varphi\right)$ then

$$\mathbf{z}^n = \underbrace{\mathbf{z}\cdot\mathbf{z}\cdots\mathbf{z}}_{n\text{-factors}} = r^n\left(\cos\left(\underbrace{\varphi+\varphi+\cdots+\varphi}_{n\text{-terms}}\right) + i\sin\left(\underbrace{\varphi+\varphi+\cdots+\varphi}_{n\text{-terms}}\right)\right)$$

and

$$\mathbf{z}^n = r^n\left(\cos n\varphi + i\sin n\varphi\right).$$

If n is a positive integer and $\mathbf{z} = r(\cos\varphi + i\sin\varphi)$ then
$\mathbf{z}^{-n} = (\mathbf{z}^{-1})^n = \underbrace{\mathbf{z}^{-1}\cdot\mathbf{z}^{-1}\cdots\mathbf{z}^{-1}}_{n\text{-factors}}$
$= r^{-n}\left(\cos\left(\underbrace{-\varphi-\varphi-\cdots-\varphi}_{n\text{-terms}}\right) + i\sin\left(\underbrace{-\varphi-\varphi-\cdots-\varphi}_{n\text{-terms}}\right)\right)$
and

$$\mathbf{z}^{-n} = r^{-n}\left(\cos\left(-n\varphi\right) + i\sin\left(-n\varphi\right)\right).$$

If $r = 1, n \in \mathbf{Z}$, then we have $\mathbf{z} = \cos\varphi + i\sin\varphi$ and the equality

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi$$

is called *De Moivre's formula*.

Lecture 10

Complex numbers: nth roots of complex numbers. Roots of 1.

nth roots of complex number

Definition 10.1
Let n be an positive integer. An nth root of the complex number z is a complex number w such that $\mathbf{w}^n = \mathbf{z}$.

Theorem 10.2
Let n be an positive integer and $z = r(\cos\varphi + i\sin\varphi)$. Then there are exactly n different nth roots of z:

$$\mathbf{w}_k = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right), \qquad k = 0, 1, 2, ..., n-1.$$

Proof. Let $\mathbf{w} = q(\cos\psi + i\sin\psi)$. Using De Moivre's formula, we get

$$q^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi).$$

The equality of two complex numbers shows (Definition 9.2) that

$$q^n = r \qquad \text{or} \qquad q = \sqrt[n]{r}$$

and

$$n\cdot\psi = \varphi + 2\pi k \qquad \text{or} \qquad \psi = \frac{\varphi+2\pi k}{n}, \qquad k \in \mathbf{Z}.$$

Thus

$$\mathbf{w} = \mathbf{w}_k = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right), \qquad k \in \mathbf{Z}'$$

If $k_1, k_2 \in \mathbf{Z}$ and

$$\frac{\varphi + 2\pi k_1}{n} = \frac{\varphi + 2\pi k_2}{n} + 2\pi \cdot l, \qquad l \in \mathbf{Z}$$

then

$$k_1 \equiv k_2 \,(\mathrm{mod}\, n)$$

and $k = 0, 1, 2, ..., n-1$ produce different values of $\mathbf{w}$.

△

Roots fo unity

An important special case of Theorem 10.2 is the numbers called the roots of unity. By unity we mean the complex number $1 = 1 + 0i$.

Definition 10.3
The roots of equation $w^n = 1$ are called nth roots of unity. The set of nth roots of unity is denoted by $U(n)$.

1 has the trigonometic form

$$1 = [1, 0] = \cos 0 + i \sin 0.$$

Thus, nth roots of unity is given by

$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \qquad k = 0, 1, 2, ..., n-1.$$

The nth roots of unity are located on the unit circle of the complex plane. They form the vertices of a n-sided regular polygon with one vertex on 1.

Definition 10.4
An nth root of unity ε is called primitive if $\varepsilon^m \neq 1$ with $m < n$.

The complex number $\varepsilon_1 = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ is a primitive nth root of unity with all $n > 1$.

Theorem 10.5

A nth root of unity $\varepsilon_k = \cos\dfrac{2\pi k}{n} + i\sin\dfrac{2\pi k}{n}$ *is primitive nth root of unity if and only if two integers k and n are relatively prime, i.e.* $\gcd(k,n) = 1$.

Proof. Let $\varepsilon_k = \cos\dfrac{2\pi k}{n} + i\sin\dfrac{2\pi k}{n}$ is primitive nth root of unity and $d = \gcd(k,n) : n = n_1 d, k = k_1 d$. Then

$$\varepsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n} = \cos\frac{2\pi k_1 d}{n_1 d} + i\sin\frac{2\pi k_1 d}{n_1 d} = \cos\frac{2\pi k_1}{n_1} + i\sin\frac{2\pi k_1}{n_1}$$

and

$$\varepsilon_{k_1}^{n_1} = 1.$$

Since ε_k is primitive, we obtain $n = n_1$ and $d = 1$. Thus, two integers k and n are relatively prime.

Conversely, suppose that two integers k and n are relatively prime and $\varepsilon_k^m = 1$. If $\dfrac{2\pi km}{n} = 2r\pi$ with integer r, then $km = nr$ and the integer km is a multiple of n.Since k and n are relatively prime, we obtain that m is a multiple of n and $m \geq n$. Thus, ε_k is primitive nth root of unity.

△

How many primitive nth roots of unity are there, for given n?

This question is answered be the proposition:

Proposition 10.6

The number of primitive nth roots of unity is $\varphi(n)$, *where* φ *is the Euler function.*

Proof is obvious.

Example 10.7

Primitive nth roots of unity

n	$\varphi(n)$	Primitive nth roots of unity
3	2	$\varepsilon_1, \varepsilon_2$
4	2	$\varepsilon_1, \varepsilon_3$
5	4	$\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$
6	2	$\varepsilon_1, \varepsilon_5$
7	6	$\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6$
8	4	$\varepsilon_1, \varepsilon_3, \varepsilon_5, \varepsilon_7$
9	6	$\varepsilon_1, \varepsilon_2, \varepsilon_4, \varepsilon_5, \varepsilon_7, \varepsilon_8$
10	4	$\varepsilon_1, \varepsilon_3, \varepsilon_7, \varepsilon_9$
11	10	$\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6, \varepsilon_7, \varepsilon_8, \varepsilon_9, \varepsilon_{10}$
12	4	$\varepsilon_1, \varepsilon_5, \varepsilon_7, \varepsilon_{11}$

.

Proposition 10.8

A nth root of unity $\varepsilon_k = \cos\dfrac{2\pi k}{n} + i\sin\dfrac{2\pi k}{n}$ *is primitive mth root of unity, where* $m = \dfrac{n}{d}$ and $d = \gcd(n,k)$.

Proof. Since two integers $n_1 = \dfrac{n}{d}$ and $k_1 = \dfrac{k}{d}$ are are relatively prime then by Theorem 10.5 the integer $\varepsilon_k = \cos\dfrac{2\pi k}{n} + i\sin\dfrac{2\pi k}{n} = \cos\dfrac{2\pi k_1}{n_1} + i\sin\dfrac{2\pi k_1}{n_1} = \varepsilon_{k_1}$ is primitive n_1th root of unity.

△

Proposition 10.9

1. If α *ir* $\beta \in U(n)$, *then* $\alpha \cdot \beta \in U(n)$.

2. If $\alpha \in U(n)$, *tthen* $\alpha^{-1} \in U(n)$.

3. If ε *is primitive nth root of unity and* $\alpha \in U(n)$, *then* $\alpha = \varepsilon^k$ *with some* $k \in N$.

4. If ε *is primitive nth root of unity and* β is *an nth root of the complex number* α, *then the numbers* $\varepsilon^0\beta, \varepsilon^1\beta, \varepsilon^2\beta, ..., \varepsilon^{n-1}\beta$ *are all nth roots of* α.

Proof. 1. If α and $\beta \in U(n)$, then $\alpha^n = \beta^n = 1$. Thus $(\alpha\beta)^n = \alpha^n\beta^n = 1$ and $\alpha \cdot \beta \in U(n)$.

2. If $\alpha \in U(n)$, then $\alpha^n = 1$. Thus $(\alpha^{-1})^n = \alpha^{-n} = (\alpha^n)^{-1} = 1$ and $\alpha^{-1} \in U(n)$.

3. If ε is primitive nth root of unity, then $\left(\varepsilon^k\right)^n = (\varepsilon^n)^k = 1$ and the number $\varepsilon^k \in U(n)$. Suppose now that $\varepsilon^k = \varepsilon^m$, with $0 \le k < m \le n-1 < n$. Then

$\varepsilon^{m-k} = 1$ with $0 < m - k < n$, a contradiction because ε is primitive nth root of unity. So the numbers $1 = \varepsilon^0, \varepsilon^1, \varepsilon^2, ..., \varepsilon^{n-1}$ are all distinct nth roost of unity.

4. Let $\beta^n = \alpha$ and ε be nth root of unity. Then $\left(\varepsilon^k\beta\right)^n = \left(\varepsilon^k\right)^n \beta^n = 1 \cdot \alpha = \alpha$ and $\varepsilon^k\beta$ is an nth root of the complex number α with all $k \in \mathbf{Z}$.Suppose now that $\varepsilon^k\beta = \varepsilon^m\beta$, with $0 \leq k < m \leq n - 1 < n$. Then $\varepsilon^{m-k} = 1$ with $0 < m - k < n$, a contradiction because ε is primitive nth root of unity. So the numbers $\beta = \varepsilon^0\beta, \varepsilon^1\beta, \varepsilon^2\beta, ..., \varepsilon^{n-1}\beta$ are all distinct nth roots of α.

$\triangle$

Lecture 11

Groups. Definitions and examples Basic properties Isomorphism of groups

Definitions and examples

Definition 11.1

A binary operation $*$ on the set G is a function $G \times G \to G$ that assigns to each pair $(g_1, g_2) \in G \times G$ a unique element $g_1 * g_2$ in G. A group $(G, *)$ is the a set G with a binary operation $*$ that satisfies the following axioms:

G1. The binary operation is associative:

$$g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3 \text{ for } g_1, g_2, g_3 \in G.$$

G2. There exist the identity element $e \in G$, such that for any $g \in G$

$$e * g = g * e = g.$$

G3. For each $g \in G$ there exist an inverse element in G , denoted by g^{-1}, such that

$$g * g^{-1} = g^{-1} * g = g.$$

A binary operation $*$ on the set G is *commutative* if and only if $g_1 * g_2 = g_2 * g_1$ for all $g_1, g_2 \in G$. A group $(G, *)$ is *commutative* if its binary operation $*$ is commutative.

A group is *finite* if it contains a finite number of elements. The *order of a finite group* is the number of elements that it contains. If the group G contains n elements, we write $|G| = n$.

Example 11.2

1. The integers $\mathbf{Z}$ with the usual addition $+$ is an infinite commutative group.
2. The rational numbers $\mathbf{Q}$ with the addition is an infinite commutative group.
3. The real numbers $\mathbf{R}$ with the addition is an infinite commutative group.
4. The complex numbers $\mathbf{C}$ with the addition is an infinite commutative group.
5. The set of all congruence classes modulo m ,$\mathbf{Z}_m$, with the addition is a finite commutative group of order m.

6. The nonzero integers $\mathbf{Z}$ with the usual multiplication $\cdot$ is an infinite commutative group.

7. The nonzero rational numbers $\mathbf{Q}$ with the multiplication is an infinite commutative group.

8. The nonzero real numbers $\mathbf{R}$ with the multiplication is an infinite commutative group.

9. The nonzero complex numbers $\mathbf{C}$ with the multiplication is an infinite commutative group.

10. The set of all congruence classes modulo m, witch have multiplicative inverses, U_m with the multiplication is a finite group of order $\varphi(n)$,where φ is the Euler function.

11. The sets of matrices $M_{r\times n}(\mathbf{Z}), M_{r\times n}(\mathbf{Q}), M_{r\times n}(\mathbf{R}), M_{r\times n}(\mathbf{C})$ whith the addition are infinite commutative groups.

12. The set of matrices $M_{r\times n}(\mathbf{Z}_{.m})$ with the addition is a finite commutative group of order $r \cdot n \cdot m$.

13. The sets of the inversible matrices over the rational numbers, over the real numbers, over the complex numbers $GL(n,\mathbf{Q}), GL(n,\mathbf{R}), GL(n,\mathbf{C}), GL(n,\mathbf{Z}_p)$,here p is prime, with the addition are infinite noncommutative groups.

14. The set of the inversible matrices over the rational numbers $GL(n,\mathbf{Z}_p)$,here p is prime, with the multiplication is a finite noncommutative group of order $n^2(p-1)$.

15. The set of nth roots of unity $U(n)$ with the multiplication is a finite commutative group of order n.

16. The set of the permutations S_n with the multiplication of permutations is a finite group of order $n!$.

Basic properties of groups

Proposition 11.3
The identity element in a group G is unique.

Proof. Suppose that e and e' are both identities in $G : e * g = g * e = g$ and $e' * g = g * e' = g$ for all $g \in G$. If e is the identity, then $e = e * e'$; if e' is the identity, then $e * e' = e'$.Thus $e = e'$.

△

Proposition 11.4
Inverses in a group are unique.

Proof. Suppose that g' and g'' are both inverses of g in a group G and e is the identity in G . Then $g' = g' * e = g' * (g * g'') = (g' * g) * g'' = e * g'' = g''$.

$\triangle$

The following proposition is fundamental.

Proposition 11.5(right and left cancellations)
Let G be a group and $a, b, c \in G$. Then $a*c = b*c$ implies $a = b$ and $c*a = c*b$ implies $a = b$.

Proof. Multiplying both sides of $a * c = b * c$ by c^{-1} , we obtain

$$a * c * c^{-1} = b * c * c^{-1}$$
$$a = b.$$

Multiplying both sides of $c * a = c * b$ by c^{-1} , we obtain

$$c^{-1} * c * a = c^{-1} * c * b$$
$$a = b$$

$\triangle$

Proposition 11.6
Let G be a group and $a, b \in G$. Then the equations $a * x = b$ and $x * a = b$ have unique solutions in G.

Proof. Suppose that

$$a * x = b.$$

Multiplying both sides by a^{-1}, we obtain

$$a^{-1} * a * x = a^{-1} * b$$
$$e * x = a^{-1} * b$$
$$x = a^{-1} * b.$$

To show uniqueness, suppose that x_1 and x_2 are both solutions of $a*x = b$.Then $x_1 = (a^{-1} * a) * x = a^{-1} * (a * x_1) = a^{-1} * (b) = a^{-1} * (a * x_2) = (a^{-1} * a) * x_2 = x_2$.

The proof for the existence and uniquess of the solution of $x * a = b$ is similar. $\triangle$

Isomorphism of groups

Definition 11.7

Let $(G, *)$ and $(H, \circ)$ are two groups. A map ϕ of a group G into a group H is a isomorphism if

1. Each element of H has at most one element mapped into it: the equation $\phi(g_1) = \phi(g_2)$ implies $g_1 = g_2$ for all $g_1, g_2 \in G$.

2. Each element h of H has at least one element g of G mapped into it: $\phi(g) = h$.

3. $\phi(g_1 * g_2) = \phi(g_1) \circ \phi(g_2)$ for all $g_1, g_2 \in G$.

Is said to have a isomorphic groups and written $G \approx H$.

Proposition 11.8

$\approx$ is an equivalence relation in any set of groups:

(i) $G \approx G$,

(ii) if $G \approx H$, then $H \approx G$,

(iii) if $G_1 \approx G_2$, and $G_2 \approx G_3$, then $G_1 \approx G_3$.

Proof is obvious.

Example 11.9

1 . $(\mathbf{C},+) \approx \left(\left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in M_2(\mathbf{R}) \right\}, + \right)$. Isomorphism: $a+ib \to \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

2. $(\mathbf{C}\setminus\{0\},\cdot) \approx \left(\left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in M_2(\mathbf{R}), a^2 + b^2 \neq 0 \right\}, \cdot \right)$. Isomorphism: $a + ib \to \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

3. $(\mathbf{R},+) \approx \left(\left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in M_2(\mathbf{R}) \right\}, + \right)$. Isomorphism: $a \to \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

4. $(\mathbf{R}\backslash\{0\},\cdot) \approx \left(\left\{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \in M_2(\mathbf{R}), a \neq 0\right\},\cdot\right)$. Isomorphism: $a \rightarrow \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

5. $(\mathbf{Z}_n,+) \approx (U(n),\cdot)$. Isomrphism: $k \rightarrow \left[1, \frac{2\pi k}{n}\right]$.

6. $(\mathbf{R},+) \approx \left(\left\{\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, a \in \mathbf{R}\right\},\cdot\right)$. Isomorphism: $a \rightarrow \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$.

Lecture 12

Vectors. Dot product and projections. Distance from a point to a line in the plane

Dot product and projections

Definition 12.1
Let $\mathbf{u}$ *and* $\mathbf{v}$ *are two vectors in 2-space or 3-space and* φ *be the angle between* $\mathbf{u}$ *and* $\mathbf{v}$. *The dot product* $\mathbf{u} \cdot \mathbf{v}$ *is defined by*

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos\varphi \qquad \textit{if } \mathbf{u} \neq \mathbf{0} \textit{ and } \mathbf{v} \neq \mathbf{0}$$

and

$$\mathbf{u} \cdot \mathbf{v} = 0 \qquad \textit{if } \mathbf{u} = \mathbf{0} \textit{ or } \mathbf{v} = \mathbf{0}.$$

If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are two vectors in 2-spaces, then the dot product of $\mathbf{u}$ and $\mathbf{v}$ is defined by

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

If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are two vectors in 3-spaces, then the corresponding formula is

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

The length $\|\mathbf{u}\|$ of a vector $\mathbf{u} = (u_1, u_2, u_3)$ is defined by

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$

The angle φ between two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ is then defined by

$$\cos\varphi = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}, \qquad 0 \leq \varphi \leq \pi.$$

Since $\|\mathbf{u}\| > 0$ and $\|\mathbf{v}\| > 0$ it follows that $\cos\varphi$ has the same sign as $\mathbf{u}\cdot\mathbf{v}$. Thus φ is acute if and only if $\mathbf{u}\cdot\mathbf{v} > 0$, φ is obtuse if and only if $\mathbf{u}\cdot\mathbf{v} < 0$, and $\varphi = \dfrac{\pi}{2}$ if and only if $\mathbf{u}\cdot\mathbf{v} = 0$. In other words, two vectors nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal (perpendicular) if and only if $\mathbf{u}\cdot\mathbf{v} = 0$.

Proposition 12.2
Let $\mathcal{L} : ax + by = c$ be the line in 2-space. Then the nonzero vector $\mathbf{n} = (a, b)$ is perpendicular to $\mathcal{L}$.

Proof. Let $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ be distinct points on the line $\mathcal{L}$:

$$\begin{aligned} ax_1 + by_1 &= c \\ ax_2 + by_2 &= c. \end{aligned}$$

Subtracting the equations, we have

$$\begin{aligned} a(x_2 - x_1) + b(y_2 - y_1) &= 0 \\ (a, b)\cdot(x_2 - x_1, y_2 - y_1) &= 0 \\ \mathbf{n}\cdot(x_2 - x_1, y_2 - y_1) &= 0. \end{aligned}$$

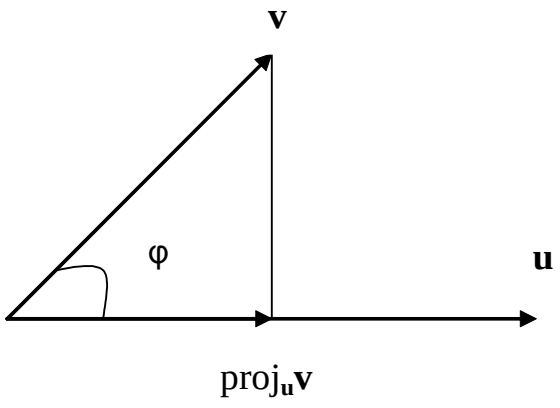
Since the nonzero vector $\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1)$ and the line $\mathcal{L}$ are parallel it follows that $\mathbf{n} = (a, b)$ is perpendicular to $\mathcal{L}$.

△

Proposition 12.3
Let $\mathcal{L}_1 : a_1x + b_1y = c_1$ and $\mathcal{L}_2 : a_2x + b_2y = c_2$ are two lines in 2-space. Then
1. $\mathcal{L}_1$ is parallel to $\mathcal{L}_2$ if and only if $a_1b_2 - b_1a_2 = 0$;
2. $\mathcal{L}_1$ is perpendical to $\mathcal{L}_2$ if and only if $a_1a_2 + b_1b_2 = 0$.

Proof is obvious.

As seen in Figure the vector $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|}\dfrac{\mathbf{u}}{\|\mathbf{u}\|} = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}$ is perpendicular to $\mathbf{u}$. It is called *the orthogonal projection of* $\mathbf{v}$ *on* $\mathbf{u}$ and it is denoted by $\text{proj}_{\mathbf{u}}\mathbf{v} = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\mathbf{u}$.
The lenght of the orthoganal projection of $\mathbf{v}$ on $\mathbf{u}$ is $\|\text{proj}_{\mathbf{u}}\mathbf{v}\| = \dfrac{|\mathbf{u}\cdot\mathbf{v}|}{\|\mathbf{u}\|} = \|\mathbf{u}\|\cos\varphi$.



Let $\mathbf{i}=(1,0,0)\,,\mathbf{j}=(0,1,0)\,,\mathbf{k}=(0,0,1)$ are the standart unit vectors in 3-space and $\mathbf{u}=(u_1,u_2,u_3)$. Then we can write $\mathbf{u}=u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k}$,where

$$\begin{gathered}u_1=\mathbf{u}\cdot\mathbf{i}=\|\mathrm{proj}_{\mathbf{i}}\mathbf{u}\|=\|\mathbf{u}\|\cos\varphi_1,\\ u_2=\mathbf{u}\cdot\mathbf{j}=\left\|\mathrm{proj}_{\mathbf{j}}\mathbf{u}\right\|=\|\mathbf{u}\|\cos\varphi_2,\\ u_3=\mathbf{u}\cdot\mathbf{k}=\|\mathrm{proj}_{\mathbf{k}}\mathbf{u}\|=\|\mathbf{u}\|\cos\varphi_3.\end{gathered}$$

Here the angles $\varphi_1,\varphi_2,\varphi_3$ between $\mathbf{u}$ and the vectors $\mathbf{i},\mathbf{j},\mathbf{k}$ respestively are called *the direction angles of* $\mathbf{u}$ and the numbers $\cos\varphi_1,\cos\varphi_2,\cos\varphi_3$ are called t*he direction cosines of* $\mathbf{u}$.

We have also

$$\begin{gathered}\mathbf{u}\cdot\mathbf{u}=\|\mathbf{u}\|\cdot\|\mathbf{u}\|\cos 0\\ u_1^2+u_2^2+u_3^2=\|\mathbf{u}\|^2\\ \left(\|u\|\cos\varphi_1\right)^2+\left(\|u\|\cos\varphi_2\right)^2+\left(\|u\|\cos\varphi_3\right)^2=\|\mathbf{u}\|^2\\ \|\mathbf{u}\|^2\left(\cos^2\varphi_1+\cos^2\varphi_2+\cos^2\varphi_3\right)=\|\mathbf{u}\|^2\\ \cos^2\varphi_1+\cos^2\varphi_2+\cos^2\varphi_3=1.\end{gathered}$$

Distance from a point to a line in the plane

Proposition 12.3

The distance D between the point $P_0\,(x_0,y_0)$ and the line $\mathcal{L}:ax+by=c$ is defined by the formula

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

Proof. Let $P(x_1, y_1)$ be any point on the line $\mathcal{L}$:

$$ax_1 + by_1 = c.$$

Let P be the initial point of the vector $\mathbf{n} = (a, b)$. By proposition 12.2 the vector $\mathbf{n}$ is perpendicular to the line $\mathcal{L}$ and the distance D is equal to the leght of $\text{proj}_{\mathbf{n}}\overrightarrow{PP_0}$:

$$D = \left\|\text{proj}_{\mathbf{n}}\overrightarrow{PP_0}\right\| = \frac{\left|\overrightarrow{PP_0} \cdot \mathbf{n}\right|}{\|\mathbf{n}\|} =$$
$$\frac{|(x_0 - x_1, y_0 - y_1) \cdot (a, b)|}{\sqrt{a^2 + b^2}} = \frac{|(x_0 - x_1)\, a + (y_0 - y_1)\, b|}{\sqrt{a^2 + b^2}} =$$
$$\frac{|x_0 a + y_0 b - (x_1 a + y_1 b)|}{\sqrt{a^2 + b^2}} = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

△

Lecture 13

Vectors. Cross product. Geometric interpretation. Scalar triple product

Cross product

Definition 13.1

Let $\mathbf{u} = (u_1, u_2, u_3)$ *and* $\mathbf{v} = (v_1, v_2, v_3)$ *are vectors in 3-space. Then* $\mathbf{u} \times \mathbf{v}$, *the cross product of* $\mathbf{u}$ *and* $\mathbf{v}$, *is defined by*

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{k}.$$

Remark 13.2

1. $\mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right) =$
$(u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$
2. The dot product of two vectors a vector and the dot product is a scalar.

The cross product of two vectors has the following properties which follow from properties of determinants.

Proposition 13.3

Let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ *are vectors in 3-space. Then the leght of*

1. $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ *is orthogonal to* $\mathbf{u}$)
2. $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ ($\mathbf{u} \times \mathbf{v}$ *is ortogonal to* $\mathbf{v}$)
3. $\|\mathbf{u} \times \mathbf{v}\|^2 = \|u\|^2 \|v\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (*Lagrange identity*)
4. $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$
5. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$
6. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = (\mathbf{u} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{w})$
7. $k(\mathbf{u} \times \mathbf{v}) = (k\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (k\mathbf{v})$
8. $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$
9. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.

The proof follow from definition 13.1 and properties of determinants

Proposition 13.4

Let $\mathbf{i}, \mathbf{j}, \mathbf{k}$ *are the standart unit vectors in 3-space. Then it is true*

$$\begin{array}{ccc} \mathbf{i}\times\mathbf{i}=\mathbf{0} & \mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{i}\times\mathbf{k}=-\mathbf{j} \\ \mathbf{j}\times\mathbf{i}=-\mathbf{k} & \mathbf{j}\times\mathbf{j}=\mathbf{0} & \mathbf{j}\times\mathbf{k}=\mathbf{i} \\ \mathbf{k}\times\mathbf{i}=\mathbf{j} & \mathbf{k}\times\mathbf{j}=-\mathbf{i} & \mathbf{k}\times\mathbf{k}=\mathbf{0} \end{array}.$$

The proof is left to the reader.

Geometric interpretation

Theorem 13.5
Let $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in 3-space. Then the leght of* $\mathbf{u}\times\mathbf{v}$ *is equal to the area of the parallelogram determinated by* $\mathbf{u}$ *and* $\mathbf{v}$.

Proof. Let φ be the angle between $\mathbf{u}$ and $\mathbf{v}$.Then using Lagrange identity and the definition of dot product of $\mathbf{u}$ and $\mathbf{v}$ we have

$$\begin{aligned} \|\mathbf{u}\times\mathbf{v}\|^2 &= \|u\|^2\|v\|^2-(\mathbf{u}\cdot\mathbf{v})^2 \\ &= \|u\|^2\|v\|^2-\|u\|^2\|v\|^2\cos^2\varphi \\ &= \|u\|^2\|v\|^2(1-\cos^2\varphi) \\ &= \|u\|^2\|v\|^2\sin^2\varphi. \end{aligned}$$

But $\sin\geq 0$ since $0\leq\varphi\leq 0$, so we have

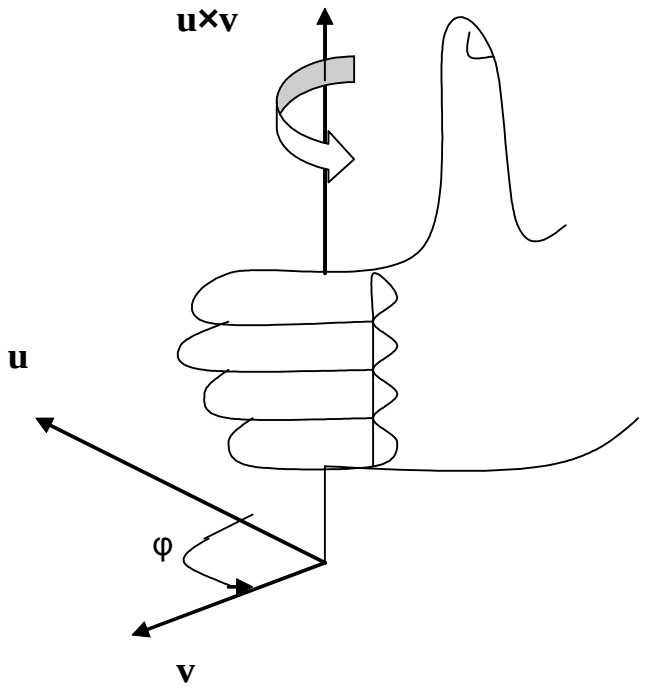
$$\|\mathbf{u}\times\mathbf{v}\|=\|\mathbf{u}\|\,\|\mathbf{v}\|\sin\varphi.$$

We know that the area S of the parallelogram determinated by $\mathbf{u}$ and $\mathbf{v}$ is given by

$$S=(\text{base})\,(\text{altitude})=\|\mathbf{u}\|\,\|\mathbf{v}\|\sin\varphi=\|\mathbf{u}\times\mathbf{v}\|.$$

△

We know that the cross product $\mathbf{u}\times\mathbf{v}$ is ortohogonal to both $\mathbf{u}$ and $\mathbf{v}$.If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, it can be shown that the direction of $\mathbf{u}\times\mathbf{v}$ can be determined using *the right hand rule*: let φ be the angle between $\mathbf{u}$ and $\mathbf{v}$, and suppose $\mathbf{u}$ is rotated through the angle φ until it coincides with $\mathbf{v}$. If the fingers of the right hand are cupped so they point in the direction of rotation, then the thumb indicates the direction of $\mathbf{u}\times\mathbf{v}$:



Scalar triple product

Definition 13.6

Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ *are vectors in 3-space. Then we call* $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ *the scalar triple product of* $\mathbf{u},\mathbf{v}$ *and* $\mathbf{w}$.

If $\mathbf{u}=(u_1,u_2,u_3)$, $\mathbf{v}=(v_1,v_2,v_3)$, $\mathbf{w}=(w_1,w_2,w_3)$ then

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=\mathbf{u}\cdot\left(\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}\mathbf{i}-\begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}\mathbf{j}+\begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}\mathbf{k}\right)$$

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=\begin{vmatrix} v_2 & v_3 \\ w_2 & w_3 \end{vmatrix}u_1-\begin{vmatrix} v_1 & v_3 \\ w_1 & w_3 \end{vmatrix}u_2+\begin{vmatrix} v_1 & v_2 \\ w_1 & w_2 \end{vmatrix}u_3$$

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

The scalar triple product of vectors has the following property which follow from properties of determinants:

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})=\mathbf{v}\cdot(\mathbf{w}\times\mathbf{u})$$

In the next proposition is geometric interpretation of scalar triple product of vectors.

Proposition 13.7

Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ *are vectors in 3-space.*

1. The absolute value of $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ is equal to the volume of the parallelepiped in 3-space *determined by the vectors* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{w}$.

2. The volume of the tetrahedron determined by the vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is $\frac{1}{6}\left|\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\right|$.

3. $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})=0$ *if and only if the vectors* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{w}$ *lie in the same plane.*

Definition 13.8

If scalar triple product $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})\neq 0$, *then the set of three vectors* $\mathbf{u}$ **,** $\mathbf{v}$ *and* $\mathbf{w}$ *is called base in 3-space.*

The following special charakterization of the base in 3-space is fundamental.

Theorem 13.9

Let $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$ *be the base in 3-spase. Then for every vector u there exist unique scalars* $\lambda_1,\lambda_2,\lambda_3$ *such that*

$$\mathbf{u}=\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\lambda_3\mathbf{a}_3..$$

Proof. Let $\mathbf{a}_1=(a_{11},a_{12},a_{13}),\mathbf{a}_2=(a_{21},a_{22},a_{23}),\mathbf{a}_3=(a_{31},a_{32},a_{33})$. If for $\mathbf{u}=(u_1,u_2,u_3)$ there exist scalars $\lambda_1,\lambda_2,\lambda_3$ such that

$$\mathbf{u}=\lambda_1\mathbf{a}_1+\lambda_2\mathbf{a}_2+\lambda_3\mathbf{a}_3$$

then

$$(u_1,u_2,u_3)=\lambda_1(a_{11},a_{12},a_{13})+\lambda_2(a_{21},a_{22},a_{23})+\lambda_3(a_{31},a_{32},a_{33})$$

and

$$\begin{aligned}\lambda_1 a_{11} + \lambda_2 a_{21} + \lambda_3 a_{31} &= u_1\\ \lambda_1 a_{12} + \lambda_2 a_{22} + \lambda_3 a_{32} &= u_2\\ \lambda_1 a_{13} + \lambda_2 a_{23} + \lambda_3 a_{33} &= u_3\end{aligned}\ .$$

The coefficient matrix $A = \begin{pmatrix} a_{11} & a_{21} & a_{31}\\ a_{12} & a_{22} & a_{32}\\ a_{13} & a_{23} & a_{33}\end{pmatrix}$ of the system is inversible because

$$\det A = \det A^T = \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) \neq 0.$$

Using Cramer's formula we get:

$$\lambda_1 = \frac{\det \begin{pmatrix} u_1 & a_{21} & a_{31}\\ u_2 & a_{22} & a_{32}\\ u_3 & a_{23} & a_{33}\end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{21} & a_{31}\\ a_{12} & a_{22} & a_{32}\\ a_{13} & a_{23} & a_{33}\end{pmatrix}} = \frac{\mathbf{u} \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

$$\lambda_2 = \frac{\det \begin{pmatrix} a_{11} & u_1 & a_{31}\\ a_{12} & u_2 & a_{32}\\ a_{13} & u_3 & a_{33}\end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{21} & a_{31}\\ a_{12} & a_{22} & a_{32}\\ a_{13} & a_{23} & a_{33}\end{pmatrix}} = \frac{\mathbf{a}_1 \cdot (\mathbf{u} \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

$$\lambda_3 = \frac{\det \begin{pmatrix} a_{11} & a_{21} & u_1\\ a_{12} & a_{22} & u_2\\ a_{13} & a_{23} & u_3\end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{21} & a_{31}\\ a_{12} & a_{22} & a_{32}\\ a_{13} & a_{23} & a_{33}\end{pmatrix}} = \frac{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{u})}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

and

$$\mathbf{u} = \frac{\mathbf{u} \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\mathbf{a}_1 + \frac{\mathbf{a}_1 \cdot (\mathbf{u} \times \mathbf{a}_3)}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\mathbf{a}_2 + \frac{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{u})}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\mathbf{a}_3.$$

To show *the uniqueness*, suppoese that also

$$\mathbf{u} = \lambda_1' \mathbf{a}_1 + \lambda_2' \mathbf{a}_2 + \lambda_3' \mathbf{a}_3.$$

Then cleary we have

$$(\lambda_1 - \lambda_1') \mathbf{a}_1 + (\lambda_2 - \lambda_2') \mathbf{a}_2 + (\lambda_3 - \lambda_3') \mathbf{a}_3 = \mathbf{0}.$$

Multiplying through by $\mathbf{a}_2 \times \mathbf{a}_3$ in the last equality, we have

$$(\lambda_1 - \lambda_1') (\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)) = \mathbf{0}.$$

But

$$\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) \neq 0,$$

then

$$\lambda_1 - \lambda_1' = 0$$

and

$$\lambda_1 = \lambda_1'.$$

Similarly, we get $\lambda_2 = \lambda_2'$ and $\lambda_3 = \lambda_3'$.

$\triangle$

Lecture 14

Planes in 3-spaces

Planes in 3-spaces

Theorem 14.1

Let $\mathcal{P}$ be a plane in 3-space. Then there exist the constants a, b, c, d and a, b, c are not all zero, $a^2+b^2+c^2 \neq 0$, such that each point $P(x, y, z)$ in $\mathcal{P}$ satisfies the equation $ax + by + cz + d = 0$.

Proof. Let $P_0(x_0; y_{0;} z_0)$ be the point in $\mathcal{P}$ and let $\mathbf{n} = (a; b; c)$ be the nonzero vector that is perpendicular to the plane $\mathcal{P}$ (the vector $\mathbf{n} = (a; b; c)$ is called *the normal* to the plane $\mathcal{P}$) . If $P(x, y, z) \in \mathcal{P}$ then the vector $\overrightarrow{P_0P} = (x - x_0; y - y_0; z - z_0)$ is perpendicular to the vector $\mathbf{n}$:

$$\mathbf{n} \cdot \overrightarrow{A_0A} = 0.$$

This equation can be written as

$$\begin{gathered} a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \\ ax + by + cz + (-ax_0 - by_0 - cz_0) = 0 \\ ax + by + cz + d = 0, \end{gathered}$$

where $d = -ax_0 - by_0 - cz_0$.
The last equation is called *the general form* of the equation of the plane $\mathcal{P}$.

△

Theorem 14.2

If a, b, c, d are the constants and a, b, c are not all zero, $a^2+b^2+c^2 \neq 0$, then there exist a plane $\mathcal{P}$ such that each point $P(x, y, z)$ in $\mathcal{P}$ satisfies the equation $ax + by + cz + d = 0$.

Proof. Is given that $a^2 + b^2 + c^2 \neq 0$. Assume that $a \neq 0$. Let

$$\begin{gathered} (x_1; y_1; z_1) = \left(-\tfrac{d}{a}; 0; 0\right), \\ (x_2, y_2, z_2) = \left(-\tfrac{d}{a} - \tfrac{b}{a}; 1; 0\right), \\ (x_3, y_3, z_3) = \left(-\tfrac{d}{a} - \tfrac{c}{a}; 0; 1\right) \end{gathered}$$

be three distinct solutions of the equation $ax + by + cz + d = 0$:

$$\begin{aligned} ax_1 + by_1 + cz_1 + d &= 0 \\ ax_2 + by_2 + cz_2 + d &= 0 \\ ax_3 + by_3 + cz_3 + d &= 0. \end{aligned}$$

Subtracting the second equation from the first equation and the third equation from the first equation we have

$$\begin{aligned} a(x_2 - x_1) + b(y_2 - y_1) + c(z_2 - z_1) &= 0 \\ a(x_3 - x_1) + b(y_3 - y_1) + c(z_3 - z_1) &= 0. \end{aligned}$$

Let $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), P_3(x_3, y_3, z_3)$ be three points in 3-space then the vector $\mathbf{n} = (a, b, c)$ is perpendicular to $\overrightarrow{P_2P_1}$ and $\overrightarrow{P_3P_1}$:

$$\begin{aligned} \mathbf{n} \cdot \overrightarrow{P_2P_1} &= 0 \\ \mathbf{n} \cdot \overrightarrow{P_3P_1} &= 0 \end{aligned}$$

Since the points P_1, P_2, P_3 are not in line(non-collinear points), then there is only one such plane $\mathcal{P}$, containing the following points. If $P(x, y, z) \in \mathcal{P}$ then the vector $\overrightarrow{PP_1}$ is a linear combination of $\overrightarrow{P_2P_1}$ and $\overrightarrow{P_3P_1}$:

$$\overrightarrow{PP_1} = s\ \overrightarrow{P_2P_1} + t\ \overrightarrow{P_3P_1}\ .$$

Then the vector $\mathbf{n} = (a, b, c)$ is perpendicular to $\overrightarrow{PP_1}$:

$$\overrightarrow{PP_1} \cdot \mathbf{n} = 0$$

or

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0$$

and

$$ax + by + cz + d = 0.$$

△

Corollary 14.3

1. Let P_1, P_2, P_3 be three non-collinear points. Then the plane $\mathcal{P}$ through P_1, P_2, P_3 is given by

$$\overrightarrow{PP_1} = s\ \overrightarrow{P_2P_1} + t\ \overrightarrow{P_3P_1}, \qquad s,t \in \mathbf{R}.$$

This is called *the vector form* of the equation of a plane.

2. Let P_1, P_2, P_3 be three non-collinear points. Then the plane $\mathcal{P}$ through P_1, P_2, P_3 is given by

$$\overrightarrow{PP_1} \cdot \left(\overrightarrow{P_2P_1} \times \overrightarrow{P_3P_1}\right) = 0,$$

where the vector$\left(\overrightarrow{P_2P_1} \times \overrightarrow{P_3P_1}\right) = \mathbf{n}$ is perpendicular to $\overrightarrow{P_2P_1}$ and $\overrightarrow{P_3P_1}$ and therefore to the plane $\mathcal{P}$.

This is called t*he normal form* for a plane.

Now we will discuss some of geometric possibilities of the plane $\mathcal{P} : ax + by + cz + d = 0$.

1. Let $a = b = 0$. Then the vector $\mathbf{n} = (0; 0; c)$ is parallel to the Oz axis and the plane $\mathcal{P}$ is parallel to the (xy)-plane.

2. Let $b = c = 0$. Then the vector $\mathbf{n} = (a; 0; 0)$ is parallel to the Ox axis and the plane $\mathcal{P}$ is parallel to the (yz)-plane.

3. Let $a = c = 0$.Then the vector $\mathbf{n} = (0; b; 0)$ is parallel to the Oy axis and the plane $\mathcal{P}$ is parallel to the (xz)-plane.

4. Let $a = 0, b \neq 0, c \neq 0$.Then the vector $\mathbf{n} = (0; b; c)$ is orthogonal to the Ox axis and the plane $\mathcal{P}$ is parallel to the Ox axis.

5. Let $a \neq 0, b = 0, c \neq 0$.Then the vector $\mathbf{n} = (a; 0; c)$ is orthogonal to the Oy axis and the plane $\mathcal{P}$ is parallel to the Oy axis.

6. Let $a \neq 0, b \neq 0, c = 0$.Then the vector $\mathbf{n} = (a; b; 0)$ is orthogonal to the Oz axis and the plane $\mathcal{P}$ is parallel to the Oz axis.

7. Let $d = 0$. Then the plane $\mathcal{P}$ passing through the origin $(0; 0; 0)$.

8. Let $a \neq 0, b \neq 0, c \neq 0, d \neq 0$. Multiplying through by $\left(-\dfrac{1}{d}\right)$ in the equality $ax + by + cy + d = 0$, we have

$$\frac{x}{-\frac{d}{a}} + \frac{y}{-\frac{d}{b}} + \frac{z}{-\frac{d}{c}} = 1$$

or

$$\tfrac{x}{\alpha} + \tfrac{y}{\beta} + \tfrac{z}{\gamma} = 1,$$

where

$$\alpha = -\frac{d}{a}, \beta = -\frac{d}{b}, \gamma = -\frac{d}{c}.$$

From here we can see that the plane $\mathcal{P}$ intertcepts the coordinate axes are $x = \alpha, y = \beta$ and $z = \gamma$.

Now we find a formula for the distance from a point to a plane.

Proposition 14.4
Let $P'(x'; y'; z')$ be the point in 3-space and $\mathcal{P}$ be the plane .$ax+by+cz+d = 0$. Then the distance D between P' and $\mathcal{P}$ is

$$D = \frac{|ax' + by' + cz' + d|}{\sqrt{a^2 + b^2 + c^2}}$$

Proof. If $P(x; y; z) \in \mathcal{P}$ and A is the initial point of the normal to $\mathcal{P}$, then the distance is equal to the lengh of the orthogonal projection of $\overrightarrow{PP'}$ on **:**

$$D = \left\| \text{proj}_{\mathbf{n}} \overrightarrow{PP'} \right\| = \frac{\left| \overrightarrow{PP'} \cdot \mathbf{n} \right|}{\|\mathbf{n}\|}.$$

But

$$\begin{aligned} \overrightarrow{PP'} &= (x' - x, y' - y, z' - z) \\ \overrightarrow{PP'} \cdot \mathbf{n} &= a(x' - x) + b(y' - y) + c(z' - z) \\ \mathbf{n} &= \sqrt{a^2 + b^2 + c^2}. \end{aligned}$$

Thus

$$D = \frac{|a(x' - x) + b(y' - y) + c(z' - z)|}{\sqrt{a^2 + b^2 + c^2}}.$$

Since $P(x; y; z) \in \mathcal{P}$ its coordinate satisfy the equation of the plane $\mathcal{P}$:

$$ax + by + cz + d = 0$$

or

$$d = -ax_1 - by_1 - cz_1.$$

Thus

$$D = \frac{|ax' + by' + cz' + (-ax - by - cz)|}{\sqrt{a^2 + b^2 + c^2}} = \frac{|ax' + by' + cz' + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

△

Now we will discuss some of geometric possibilities of two planes.

Proposition 14.5

Let $\mathcal{P}_1$ be the plane $a_1x + b_1y + c_1z + d_1 = 0$, and let $\mathcal{P}_2$ be the plane $a_2x + b_2y + c_2z + d_2 = 0$. Then

1. $\mathcal{P}_1$ is parallel to $\mathcal{P}_2$ if and only if

$$\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}.$$

2. $\mathcal{P}_1$ is orthogonal to $\mathcal{P}_2$ if and only if

$$a_1a_2 + b_1b_2 + c_1c_2 = 0.$$

3. The angle φ between $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfies the equation

$$\cos\varphi = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}} \qquad 0 \le \varphi \le \pi.$$

Proof. Let $\mathbf{n}_1=(a_1, b_1, c_1)$ be the normal to $\mathcal{P}_1$, and let $\mathbf{n}_2 = (a_2, b_2, c_2)$ be the normal to $\mathcal{P}_2$. Then

1) $P_1 \parallel P_2 \Longleftrightarrow \mathbf{n}_1 \parallel \mathbf{n}_2 \Longleftrightarrow \dfrac{a_1}{a_2} = \dfrac{b_1}{b_2} = \dfrac{c_1}{c_2}$.

2) $P_1 \perp P_1 \Longleftrightarrow \mathbf{n}_1 \perp \mathbf{n}_2 \Longleftrightarrow \mathbf{n}_1 \cdot \mathbf{n}_2 = 0 \Leftrightarrow a_1a_2 + b_1b_2 + c_1c_2 = 0$.

3) The angle φ between $\mathcal{P}_1$ and $\mathcal{P}_2$ is the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$. Thus

$$\cos\varphi = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{\|\mathbf{n}_1\| \, \|\mathbf{n}_2\|} = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}}.$$

△

Now we will discuss some of geometric possibilities of three planes.

Proposition 14.6

Let $\mathcal{P}_1$ be the plane $a_1x + b_1y + c_1z + d_1 = 0$, let $\mathcal{P}_2$ be the plane $a_2x + b_2y + c_2z + d_2 = 0$ and let $\mathcal{P}_3$ be the plane $a_3x + b_3y + c_3z + d_3 = 0$. Then

1. The planes $\mathcal{P}_1$,$\mathcal{P}_2$,$\mathcal{P}_3$ intersect in one point if and only if $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0.$

2. The planes $\mathcal{P}_1$,$\mathcal{P}_2$,$\mathcal{P}_3$ are parallel to one line if and only if $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0.$

Proof. Let $\mathbf{n}_1 = (a_1, b_1, c_1)$ be the normal to $\mathcal{P}_1$, let $\mathbf{n}_2 = (a_2, b_2, c_2)$ be the normal to $\mathcal{P}_2$ and let $\mathbf{n}_3 = (a_3, b_3, c_3)$ be the normal to $\mathcal{P}_3$. Then

1.

The planes $\mathcal{P}_1$,$\mathcal{P}_2$,$\mathcal{P}_3$ intersect in one point

if and only if

the normals $\mathbf{n}_1, \mathbf{n}_2$ and $\mathbf{n}_3$ are not parallel to one plane

if and only if

$$\mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) \neq 0$$

if and only if

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0.$$

2.

The planes $\mathcal{P}_1$,$\mathcal{P}_2$,$\mathcal{P}_3$ are parallel to one line

if and only if

the normals $\mathbf{n}_1, \mathbf{n}_2$ and $\mathbf{n}_3$ are parallel to one plane

if and only if

$$\mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) = 0$$

if and only if

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0.$$

$\triangle$

Here are more details.

Let $\mathcal{P}_1$ be the plane $a_1x + b_1y + c_1z + d_1 = 0$, let $\mathcal{P}_2$ be the plane $a_2x + b_2y + c_2z + d_2 = 0$ and let $\mathcal{P}_3$ be the plane $a_3x + b_3y + c_3z + d_3 = 0$. Let A be the coefficient matrix and let B be the augmented matrix of the system

$$\begin{aligned} a_1x + b_1y + c_1z + d_1 &= 0 \\ a_2x + b_2y + c_2z + d_2 &= 0 \\ a_3x + b_3y + c_3z + d_3 &= 0. \end{aligned}$$

If $\mathbf{rank}A = 1$, $\mathbf{rank}B = 2$, then
1. *Three parallel planes*:

$$\begin{gathered} \mathcal{P}_1 \| \mathcal{P}_2 \| \mathcal{P}_3 \iff \\ \mathbf{n}_1 \| \mathbf{n}_2 \| \mathbf{n}_3 \iff \\ \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} \neq \frac{d_1}{d_2} \\ \frac{a_1}{a_3} = \frac{b_1}{b_3} = \frac{c_1}{c_3} \neq \frac{d_1}{d_3} \\ \frac{a_2}{a_3} = \frac{b_2}{b_3} = \frac{c_2}{c_3} \neq \frac{d_2}{d_3} \end{gathered}$$

2. *Two coincident planes parallel to a third plane*:

$$\begin{gathered} \mathcal{P}_1 = \mathcal{P}_2 \| \mathcal{P}_3 \iff \\ \mathbf{n}_1 \| \mathbf{n}_2 \| \mathbf{n}_3 \iff \\ \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2} \\ \frac{a_1}{a_3} = \frac{b_1}{b_3} = \frac{c_1}{c_3} \neq \frac{d_1}{d_3}. \end{gathered}$$

If $\mathbf{rank}A = 2$, $\mathbf{rank}B = 3$, then
3. *Two planes parallel*:

$$\begin{gathered}\mathcal{P}_1 \| \mathcal{P}_2 \nparallel \mathcal{P}_3 \iff \\ \mathbf{n}_1 \| \mathbf{n}_2 \nparallel \mathbf{n}_3 \iff \\ \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} \neq \frac{d_1}{d_2}\end{gathered}$$

4. *Three planes with no common intersection*:

$$\begin{gathered}\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3 = \varnothing \iff \\ \mathbf{n}_1 \nparallel \mathbf{n}_2 \nparallel \mathbf{n}_3 \text{ and } \mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) = 0.\end{gathered}$$

If $\mathbf{rank}A = 3$, then
5. *Three planes intersecting at a point*:

$$\begin{gathered}\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3 = \{P\} \iff \\ \mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) \neq 0.\end{gathered}$$

If $\mathbf{rank}A = \mathbf{rank}B = 1$, then
6. *Three coincident planes*:

$$\begin{gathered}\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_3 \iff \\ \mathbf{n}_1 = \mathbf{n}_2 = \mathbf{n}_3 \iff \\ \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2} \\ \frac{a_1}{a_3} = \frac{b_1}{b_3} = \frac{c_1}{c_3} = \frac{d_1}{d_3}.\end{gathered}$$

If $\mathbf{rank}A = \mathbf{rank}B = 2$, then
7. *Two coincident planes intersecting a third plane*:

$$\begin{gathered}\mathcal{P}_1 = \mathcal{P}_2 \nparallel \mathcal{P}_3 \iff \\ \mathbf{n}_1 = \mathbf{n}_2 \nparallel \mathbf{n}_3 \iff \\ \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2}\end{gathered}$$

8. *Three planes intersecting in a line*:
$\mathcal{P}_1 \cap \mathcal{P}_2 \cap \mathcal{P}_3 = \mathcal{L}$ a line $\iff$
$\mathbf{n}_1 \nparallel \mathbf{n}_2, \mathbf{n}_2 \nparallel \mathbf{n}_3, \mathbf{n}_1 \nparallel \mathbf{n}_3$ and $\mathbf{n}_1 \cdot (\mathbf{n}_2 \times \mathbf{n}_3) = 0$.

Lecture 15

Lines in 3-spaces

Lines in 3-spaces

We shall now show how to obtain equations for lines in 3-space.

Definition 15.1
The line $\mathcal{L}$ in 3-space is the intersection of two given planes $\mathcal{P}_1, \mathcal{P}_2 : \mathcal{L} = \mathcal{P}_1 \cap \mathcal{P}_2$. Let $\mathcal{P}_1$ be the plane $a_1x + b_1y + c_1z + d_1 = 0$, let $\mathcal{P}_2$ be the plane $a_2x + b_2y + c_2z + d_2 = 0$. Then the system

$$\begin{array}{l} a_1x + b_1y + c_1z + d_1 = 0 \\ a_2x + b_2y + c_2z + d_2 = 0 \end{array}.$$

is called the general form of the equation of the line $\mathcal{L}$.

Let $P_0\left(x_0; y_0; z_0\right)$ be the point in the line $\mathcal{L}$ and let $\mathbf{v} = (k; l; m)$ be the nonzero vector that is parallel to the line $\mathcal{L}$ (the vector $\mathbf{v} = (k; l; m)$ is called *the direction vector* for the line $\mathcal{L}$) . If $P\left(x, y, z\right) \in \mathcal{L}$ then the vector $\overrightarrow{P_0P} = (x - x_0; y - y_0; z - z_0)$ is parallel to the vector $\mathbf{v}$:

$$\overrightarrow{P_0P} = t\mathbf{v}, \qquad t \in \mathbf{R}. \qquad (1)$$

This equality is called *the vector form of the equation of the line $\mathcal{L}$.*

In termes ao coordinates, the vector form of the equation of $\mathcal{L}$ can be written as

$$(x - x_0; y - y_0; z - z_0) = (tk; tl; tm), \qquad t \in \mathbf{R}$$

or

$$\begin{array}{c} x = x_0 + tk \\ y = y_0 + tl \\ z = z_0 + tm \end{array}$$

where $t \in \mathbf{R}$. This equations are called *parametric equations* for $\mathcal{L}$.

Equivalently, equation (1) becomes

$$\frac{x - x_0}{k} = \frac{y - y_0}{l} = \frac{z - z_0}{m}$$

These are called canonical equations for the line $\mathcal{L}$.

Remark 15.2

Let $\mathbf{n}_1 = (a_1, b_1, c_1)$ be the normal to $\mathcal{P}_1$, let $\mathbf{n}_2 = (a_2, b_2, c_2)$ be the normal to $\mathcal{P}_2$. Then $\mathbf{v} = \mathbf{n}_1 \times \mathbf{n}_2$ is the direction vector for the line $\mathcal{L} = \mathcal{P}_1 \cap \mathcal{P}_2$.

Now we will discuss some of geometric possibilities of the line $\mathcal{L}$, which the direction vector is $\mathbf{v} = (k, l, m)$.

1. $k = 0 \iff \mathbf{v} \cdot \mathbf{i} = 0 \iff \mathcal{L}$ is parallel to the (yz)-plane.
2. $l = 0 \iff \mathbf{v} \cdot \mathbf{j} = 0 \iff \mathcal{L}$ is parallel to the (xz)-plane.
3. $m = 0 \iff \mathbf{v} \cdot \mathbf{k} = 0 \iff \mathcal{L}$ is parallel to the (xy)-plane.
4. $k = l = 0 \iff \mathbf{v} \parallel \mathbf{k} \iff \mathcal{L}$ is parallel to the Oz axis.
5. $l = m = 0 \iff \mathbf{v} \parallel \mathbf{i} \iff \mathcal{L}$ is parallel to the Ox axis.
6. $k = m = 0 \iff \mathbf{v} \parallel \mathbf{j} \iff \mathcal{L}$ is parallel to the Oy axis.

Now we will discuss some of geometric possibilities of two lines.

Proposition 15.3

Let $\mathcal{L}_1$ *be the line* $\dfrac{x - x_1}{k_1} = \dfrac{y - y_1}{l_1} = \dfrac{z - z_1}{m_1}$ *and let* $\mathcal{L}_2$ *be the line* $\dfrac{x - x_2}{k_2} = \dfrac{y - y_2}{l_2} = \dfrac{z - z_2}{m_2}$.

Then

1. $\mathcal{L}_1$ *is parallel to* $\mathcal{L}_2$ *if and only if*

$$\frac{k_1}{k_2} = \frac{l_1}{l_2} = \frac{m_1}{m_2}..$$

2. $\mathcal{L}_1$ *is orthogonal to* $\mathcal{L}_2$ *if and only if*

$$k_1 k_2 + l_1 l_2 + m_1 m_2 = 0.$$

3. The angle φ *between* $\mathcal{L}_1$ *and* $\mathcal{L}_2$ *satisfies the equation*

$$\cos\varphi = \frac{k_1 k_2 + l_1 l_2 + m_1 m_2}{\sqrt{k_1 + l_1 + m_1}\sqrt{k_2 + l_2 + m_2}} \qquad 0 \le \varphi \le \pi.$$

4. $\mathcal{L}_1$ and $\mathcal{L}_2$ is in one plane if and only if

$$\begin{vmatrix} k_1 & l_1 & m_1 \\ k_2 & l_2 & m_2 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{vmatrix} = 0.$$

5. $\mathcal{L}_1$ and $\mathcal{L}_2$ is not in one plane if and only if

$$\begin{vmatrix} k_1 & l_1 & m_1 \\ k_2 & l_2 & m_2 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{vmatrix} \neq 0.$$

Proof.
1.

$$\mathcal{L}_1 \text{ is parallel to } \mathcal{L}_2$$

if and only if

$$\mathbf{v}_1 = (k_1, l_1, m_1) \text{ is parallel to } \mathbf{v}_2 = (k_2, l_2, m_2)$$

if and only if

$$\frac{k_1}{k_2} = \frac{l_1}{l_2} = \frac{m_1}{m_2}.$$

2.

$$\mathcal{L}_1 \text{ is orthogonal to } \mathcal{L}_2$$

if and only if

$$\mathbf{v}_1 = (k_1, l_1, m_1) \text{ is orthogonal to } \mathbf{v}_2 = (k_2, l_2, m_2)$$

if and only if

$$k_1 k_2 + l_1 l_2 + m_1 m_2 = 0.$$

3. The angle φ between $\mathcal{L}_1$ and $\mathcal{L}_2$ is the angle between $\mathbf{v}_1 = (k_1, l_1, m_1)$ and $\mathbf{v}_2 = (k_2, l_2, m_2)$ .Thus

$$\cos\varphi = \frac{k_1k_2 + l_1l_2 + m_1m_2}{\sqrt{k_1 + l_1 + m_1}\sqrt{k_2 + l_2 + m_2}}.$$

4.

$$\mathcal{L}_1 \text{ and } \mathcal{L}_2 \text{ is in one plane}$$

if and only if

$$\mathbf{v}_1 = (k_1, l_1, m_1)\ ,\ \mathbf{v}_2 = (k_2, l_2, m_2) \text{ and } \overrightarrow{P_2P_1} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \text{ is in one plane}$$

if and only if

$$\mathbf{v}_1 \cdot \left(\mathbf{v}_2 \times \overrightarrow{P_2P_1}\right) = 0.$$

5.

$$\mathcal{L}_1 \text{ and } \mathcal{L}_2 \text{ is not in one plane}$$

if and only if

$$\mathbf{v}_1 = (k_1, l_1, m_1)\ ,\ \mathbf{v}_2 = (k_2, l_2, m_2) \text{ and } \overrightarrow{P_2P_1} = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \text{ is not in one plane}$$

if and only if

$$\mathbf{v}_1 \cdot \left(\mathbf{v}_2 \times \overrightarrow{P_2P_1}\right) \neq 0.$$

△

Now we will discuss some of geometric possibilities of a line and a plane.

Proposition 15.4

Let $\mathcal{L}$ be the line $\dfrac{x - x_0}{k} = \dfrac{y - y_0}{l} = \dfrac{z - z_0}{m}$ and let $\mathcal{P}$ be the plane $ax + by + cz + d = 0$. Then

1. $\mathcal{L}$ is parallel to $\mathcal{P}$ if and only if $ak + bl + cm = 0$.

2. $\mathcal{L}$ is orthogonal to $\mathcal{P}$ if and only if $\dfrac{a}{k} = \dfrac{b}{l} = \dfrac{c}{m}$.

Proof. 1.

$\mathcal{L}$ is parallel to $\mathcal{P}$

if and only if

the direct vector $\mathbf{v} = (k, l, m)$ for $\mathcal{L}$ is orthogonal to normal vector $\mathbf{n} = (a, b, c)$ for $\mathcal{P}$

if and only if

$$ak + bl + cm = 0.$$

2.

$\mathcal{L}$ is orthogonal to $\mathcal{P}$

if and only if

the direct vector $\mathbf{v} = (k, l, m)$ for $\mathcal{L}$ is parallel to normal vector $\mathbf{n} = (a, b, c)$ for $\mathcal{P}$

if and only if

$$\frac{a}{k} = \frac{b}{l} = \frac{c}{m}.$$

△

Now we find a formula for the distance from a point to a line.

Proposition 14.4

Let P be the point in 3-space and let Q and R be two distinct points on the line $\mathcal{L}$. Then there exist the unique point S on $\mathcal{L}$ such that the vector $\overrightarrow{PS}$ is orthogonal to vector $\overrightarrow{QR}$, namely

$$\overrightarrow{QS} = t\overrightarrow{QR}, \qquad t = \frac{\overrightarrow{QS} \cdot \overrightarrow{QP}}{\left\| \overrightarrow{QS} \right\|^2}.$$

The distance D between a point P and the line $\mathcal{L}$ is

$$D = \frac{\sqrt{QP^2 \cdot QR^2 - \left(\overrightarrow{QP} \cdot \overrightarrow{QR}\right)^2}}{QR}$$

Proof. The vector $\overrightarrow{PS}$ must be perpendicular to the vector $\overrightarrow{QR}$. Therefore

$$\begin{gathered}
\overrightarrow{PS} \cdot \overrightarrow{QR} = 0 \\
\left(\overrightarrow{QS} - \overrightarrow{QP}\right) \cdot \overrightarrow{QR} = 0 \\
\left(t\overrightarrow{QR} - \overrightarrow{QP}\right) \cdot \overrightarrow{QR} = 0 \\
t\overrightarrow{QR} \cdot \overrightarrow{QR} - \overrightarrow{QP} \cdot \overrightarrow{QR} = 0 \\
t\left\|\overrightarrow{QR}\right\|^2 = \overrightarrow{QP} \cdot \overrightarrow{QR} \\
t = \frac{\overrightarrow{QP} \cdot \overrightarrow{QR}}{\left\|\overrightarrow{QR}\right\|^2}.
\end{gathered}$$

Finally, as $\overrightarrow{PS}$ is perpendicular to $\overrightarrow{QS}$, Phythagoras' theorem gives

$$\begin{gathered}
D^2 = PS^2 = QP^2 - QP^2 = \\
QP^2 - \left\|t\overrightarrow{QR}\right\|^2 = \\
QP^2 - t^2 QR^2 = \\
QP^2 - \left(\frac{\overrightarrow{QR} \cdot \overrightarrow{QP}}{\|QR\|^2}\right)^2 QR^2 = \\
\frac{QP^2 \cdot QR^2 - \left(\overrightarrow{QP} \cdot \overrightarrow{QR}\right)^2}{QR^2}
\end{gathered}$$

△