INTRODUCTION TO MODERN ALGEBRA AND GEOMETRY

HAO SUN

1. The Space: $\mathbb{R}^n$

Notation 1.1. Here are some notations for sets:

- $\mathbb{Z}$: set of integers;
- $\mathbb{Q}$: set of rational numbers;
- $\mathbb{R}$: set of real numbers;
- $\mathbb{C}$: set of complex numbers;

Definition 1.2 (Cartesian product)**.** Given two sets S, T, the *Cartesian product* of S and T is defined as a new set

$$S \times T := \{(s,t) \mid s \in S,\ t \in T\}.$$

Example 1.3. We consider some examples about Cartesian products.

- Let $S = \{1,2,3\}$ and $T = \{1,2\}$. Then, $S \times T$ is
$$S \times T = \{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\}.$$
- Let $S = \mathbb{R}$. Then,
$$\mathbb{R} \times \mathbb{R} = \{(r_1, r_2), r_1, r_2 \in \mathbb{R}\}.$$
This Cartesian product is denoted by $\mathbb{R}^2$.

With the same construction as above, we can define

$$\mathbb{R}^n := \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{n}$$

An element $\mathbf{v} \in \mathbb{R}^n$ is called a *vector*. There is a special vector $\mathbf{0}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, which is called the *zero vector*. If there is no ambiguity, we omit the subscript and use $\mathbf{0}$ for the zero vector.

Now we define some operations on vectors. Given two vectors

$$\mathbf{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

we say $\mathbf{u} = \mathbf{v}$, if $u_i = v_i$ for $1 \le i \le n$. We define their *sum* $\mathbf{u} + \mathbf{v}$ as

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n. \end{pmatrix}$$

Now let $c \in \mathbb{R}$ be a real number, which is called a *constant* or *scalar*. The *scalar multiplication* is defined as

$$c \cdot \mathbf{u} = c \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} cu_1 \\ \vdots \\ cu_n \end{pmatrix}.$$

Remark 1.4. Understand the above operations in $\mathbb{R}^2$ from graphs.

Lemma 1.5 (Algebraic properties for vectors)**.** *Let $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ be vectors in $\mathbb{R}^n$, and let c, d be two scalars in $\mathbb{R}$. The following rules hold*

(1) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$;
(2) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$;
(3) $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$;
(4) $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$;
(5) $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$;
(6) $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$;
(7) $c(d\boldsymbol{u}) = (cd)\boldsymbol{v}$;
(8) $1\boldsymbol{u} = \boldsymbol{u}$.

Proof. The proof of these formula is not hard. We only give the proof of the first one as an example.

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix} = \begin{pmatrix} v_1 + u_1 \\ \vdots \\ v_n + u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \mathbf{v} + \mathbf{u}.$$

□

Given vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$ and scalars $c_1, \dots c_p$, the vector

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \sum_{i=1}^{p} c_i\mathbf{v}_i$$

is called a *linear combination* of $\mathbf{v}_1, \dots, \mathbf{v}_p$ with weights (or coefficients) $c_1, \dots c_p$.

Remark 1.6 (Summation). Let $(a_1, \dots, a_i, \dots)$ be a sequence of numbers (or vectors). We define the *sum* of numbers from a_j to a_n as follows

$$\sum_{i=j}^{n} a_i := a_j + a_{j+1} + \cdots + a_n.$$

The notation $\sum$ is for *summation*, the subscript $i = j$ means that the starting number is a_j, the upperscript n means that we stop at a_n.

Similarly, let $f(x)$ be a continuous function on $\mathbb{R}$. We define

$$\sum_{i=j}^{n} f(i) := f(j) + f(j+1) + \cdots + f(n).$$

Definition 1.7. Given vectors $\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^n$, the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted by $\mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called *the subset of $\mathbb{R}^n$ spanned (or generated) by $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$.*

Remark 1.8. geometric interpretation of span in $\mathbb{R}^2$.

Definition 1.9 (Linearly independence)**.** Let $\mathbf{v}_1, \dots, \mathbf{v}_p$ be vectors in $\mathbb{R}^n$. They are *linearly independent* if the equation

$$x_1\mathbf{v}_1 + \cdots + x_p\mathbf{v}_p = 0$$

has only trivial solution. Otherwise, they are called *linearly dependent*.

Example 1.10. Consider the following three vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

It is easy to check that

$$2\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 = 0.$$

Thus, the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent.

Definition 1.11. A *basis* of $\mathbb{R}^n$ is a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ such that

(1) the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are linearly independent;
(2) given any other vector $\mathbf{v} \in \mathbb{R}^n$, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p, \mathbf{v}$ are linearly dependent.

Here is a basis of $\mathbb{R}^n$:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

This basis is called the *standard basis* of $\mathbb{R}^n$.

Lemma 1.12 (Substitution Lemma)**.** *Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ be a basis of $\mathbb{R}^n$. Let $\boldsymbol{v}$ be a nonzero vector in $\mathbb{R}^n$. Suppose that*

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_p\boldsymbol{v}_p$$

and $c_1 \neq 0$, then $\{\boldsymbol{v}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ is a basis of $\mathbb{R}^n$.

Proof. First, we prove the set of vectors $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is a basis of $\mathbb{R}^n$. Suppose that there exists real numbers $d_1, \ldots, d_p$ such that

$$d_1\mathbf{v} + \cdots + d_p\mathbf{v}_p = 0.$$

If $d_1 = 0$, then

$$d_1\mathbf{v}_2 + \cdots + d_p\mathbf{v}_p = 0.$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ are linearly independent, then $d_2 = \cdots = d_p = 0$. Thus, $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ are linearly independent. If $d_1 \neq 0$, then we have

$$d_1(\sum_{i=1}^{n} c_i\mathbf{v}_i) + d_2\mathbf{v}_2 + \cdots + d_p\mathbf{v}_p = 0.$$

Rewrite the above formula

$$d_1c_1\mathbf{v}_1 + (d_1c_2 + d_2)\mathbf{v}_2 + \cdots + (d_1c_p + d_p)\mathbf{v}_p = 0.$$

We find that the coefficient of $\mathbf{v}_1$ is nonzero, and this contracts to the assumption that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is a basis. Therefore, this case cannot happen. The above discussion shows that $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ are linearly independent.

Now we will show that any vector $w \in \mathbb{R}^n$ can be written as a linear combination of $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. Since

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p,$$

we have

$$\mathbf{v}_1 = -\frac{1}{c_1}\mathbf{v}_1 + \cdots + \frac{c_p}{c_1}\mathbf{v}_p.$$

Given any vector $w \in \mathbb{R}^n$, w can be written as a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$, and suppose that

$$w = f_1\mathbf{v}_1 + \cdots + f_n\mathbf{v}_n.$$

Then,

$$w = f_1(-\frac{1}{c_1}\mathbf{v}_1 + \cdots + \frac{c_p}{c_1}\mathbf{v}_p) + f_2\mathbf{v}_2 + \cdots + f_n\mathbf{v}_n.$$

Therefore, w is written as a linear combination of $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. This finishes the proof of this lemma. □

Theorem 1.13. *If $\{v_1, \ldots, v_p\}$ is a basis of $\mathbb{R}^n$, then $p = n$. This implies that the number of vectors in any basis of $\mathbb{R}^n$ is n.*

Proof. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{R}^n$. Given another basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of $\mathbb{R}^n$, suppose that $p < n$. By the Substitution Lemma, $\{\mathbf{e}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is also a basis. Repeating this process, we find that $\{\mathbf{e}_1, \ldots, \mathbf{e}_p\}$ is also a basis of $\mathbb{R}^n$. However, this is false. Therefore, $p \geq n$. The case of $p > n$ can be proved similarly. Thus, $p = n$. □

Definition 1.14. A *linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ is a map such that for any $a, b \in \mathbb{R}$ and $v, w \in \mathbb{R}^n$, we have

$$T(av_1 + bv_2) = aT(v_1) + bT(v_2).$$

Consider the zero vector $\mathbf{0}$, we have

$$T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) = 2T(\mathbf{0}).$$

Therefore,

$$T(\mathbf{0}) = \mathbf{0}.$$

More precisely,

$$T(\mathbf{0}_n) = \mathbf{0}_m.$$

Now let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of $\mathbb{R}^n$. Given any vector $\mathbf{v} \in \mathbb{R}^n$, we write $\mathbf{v}$ as a linear combination of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Then,

$$\begin{aligned} T(\mathbf{v}) &= T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n). \end{aligned}$$

Therefore, for any vector $v \in \mathbb{R}^n$, the value of $T(v)$ is uniquely determines by the values $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$. In fact, a linear transformation can be expressed by a *matrix*.

Definition 1.15. An $m \times n$ *matrix* A with coefficients in $\mathbb{R}$ is an array of numbers with m rows and n columns

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

where a_{ij} is a real number at the i-th row and j-th column. We say that *the entry at (i, j) is a_{ij}*. Sometimes, we also use the notation $A = (a_{ij})_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} = (a_{ij})$ for the matrix. If a matrix A has m rows and n columns, then we say that the *size* of A is $m \times n$.

Example 1.16. If a matrix is of size $n \times 1$, it is called a *column vector*, while a matrix is called a *row vector* if it is of size $1 \times n$. If a matrix is of size $n \times n$, it is called a *square matrix*. The *diagonal* is defined as the entries at (i, i), where $1 \leq i \leq n$.

There are tow special matrices. The first one is the *zero matrix* 0, of which the entries are all zero. The second one is the *identity matrix* I_n, of which the size is $n \times n$ and the entry on the diagonal is one, while the others are zero.

Let A, B be two matrices. If A and B are of the same size, the *sum* is defined as

$$A + B := (a_{ij} + b_{ij}).$$

Let c be a real number. The *scalar multiplication* is

$$c \cdot A := (ca_{ij}).$$

Now we are going to define the *multiplication* of matrices. We first consider the baby case. Let A be a row matrix of size $1 \times n$ and B be a column matrix of size $n \times 1$. Their multiplication is defined as

$$A \times B = \begin{pmatrix} a_{11} \dots a_{1n} \end{pmatrix} \times \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} := a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} = \sum_{i=1}^{n} a_{1i}b_{i1}.$$

Now let A be a matrix of size $p \times n$ and let B be a matrix of size $n \times q$. The multiplication is defined as

$$A \times B := C = (c_{ij})_{\substack{1 \leq i \leq p, \\ 1 \leq j \leq q}},$$

where $c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}$.

Lemma 1.17. *We have the following properties of matrices.*

(1) $A + B = B + A$;
(2) $(A + B) + C = A + (B + C)$;
(3) $A + 0 = A$;
(4) $r(A + B) = rA + rB$;
(5) $(r + s)A = rA + sA$;
(6) $r(sA) = (rs)A$;
(7) $A(BC) = (AB)C$;
(8) $A(B + C) = AB + AC$;
(9) $(B + C)A = BA + CA$;
(10) $r(AB) = (rA)B = A(rB)$.

Proof. The proof is left to the readers. □

Remark 1.18.
(1) The multiplication of matrix only makes sense when the number of columns of the first matrix equal to the number of rows of the second matrix.
(2) $AB \neq BA$ in general.
(3) If $AB = AC$, it is not true in general that $B = C$.

Lemma 1.19. *Let A be an $m \times n$ matrix. Then, the map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined as*

$$T_A(\boldsymbol{v}) = A\boldsymbol{v}$$

is a linear transformation.

Conversely, let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then, there exists a unique $m \times n$ matrix A (with respect to the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$) such that $T(\boldsymbol{v}) = A\boldsymbol{v}$.

Proof. By Lemma 1.17, for any two vectors $\mathbf{v}, \mathbf{w}$ and real numbers a, b, we have

$$T_A(a\mathbf{v} + b\mathbf{w}) = A(a\mathbf{v} + b\mathbf{w}) = aA\mathbf{v} + bA\mathbf{w} = aT_A(\mathbf{v}) + bT_A(\mathbf{w}).$$

Therefore, T_A is a linear transformation.

As we discussed above, we know that given any vector $\mathbf{v} \in \mathbb{R}^n$, the value $T(\mathbf{v})$ is uniquely determined by $T(\mathbf{e}_1), \dots T(\mathbf{e}_n)$, where $\mathbf{e}_1, \dots, \mathbf{e}_n$ is the standard basis. Since $T(\mathbf{e}_j)$ is a vector in $\mathbb{R}^m$, suppose that

$$T(\mathbf{e}_j) = a_{1j}\mathbf{f}_1 + \cdots + a_{mj}\mathbf{f}_m,$$

where a_{ij} are real numbers and $\mathbf{f}_1, \dots, \mathbf{f}_m$ is the standard basis of $\mathbb{R}^m$. We define the matrix A as

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

It is easy to check that

$$T(\mathbf{e}_i) = A\mathbf{e}_i.$$

Therefore, for any vector $v \in \mathbb{R}^n$, we have

$$T(\mathbf{v}) = A\mathbf{v}.$$

□

Add determinants, will be use as an example of Groups

2. Example: Grassmannian

Definition 2.1 (Subspaces of $\mathbb{R}^n$)**.** A *subspace* V of $\mathbb{R}^n$ is a set such that

(1) the zero vector is in V;
(2) if $\mathbf{u}, \mathbf{v} \in V$, then $c\mathbf{u} + d\mathbf{v} \in V$ for arbitrary $c, d \in \mathbb{R}$.

The second property actually shows that V is closed under addition and scalar multiplication.

Lemma 2.2. *Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$ be vectors in $\mathbb{R}^n$ (not necessarily to be linearly independent). Then,* $\mathrm{Span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$ *is a subspace of* $\mathbb{R}^n$*.*

Proof. By definition, the set $\mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ includes all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$, i.e.

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p \in \mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$$

for arbitrary $c_i \in \mathbb{R}$. Now taking $c_i = 0$, clearly $\mathbf{0} \in \mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

For the second condition, given $\mathbf{u}, \mathbf{v} \in \mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$. As linear combinations, we have

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$
$$\mathbf{u} = d_1\mathbf{v}_1 + \cdots + d_p\mathbf{v}_p.$$

Then,

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + \cdots + (c_p + d_p)\mathbf{v}_p,$$

which is also a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_p$. □

Any subspace of $\mathbb{R}^n$ can be given in this way.

Proposition 2.3. *Any subspace of $\mathbb{R}^n$ can be written as* $\mathrm{Span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$ *for some vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_p$. Furthermore, the vectors can be chosen to be a basis of the subspace.*

Example 2.4. A subspace V of dimension one in $\mathbb{R}^n$ is a line passing through the origin. Suppose that $V = \mathrm{Span}\{\mathbf{v}\}$. Then, the vector corresponds to the direction of the line. In $\mathbb{R}^2$, this vector is related to the slope of the line.

Now we consider the following set

$$\mathrm{Gr}_k(\mathbb{R}^n) := \{\text{all dimension } k \text{ subspaces in } \mathbb{R}^n\ \},$$

and this set (or space) is called the *Grassmannian*.

Now we consider the example $k = 1$ and $n = 2$. In this case, $\mathrm{Gr}_1(\mathbb{R}^2)$ parametrizes all dimension one subspaces in $\mathbb{R}^2$. Note that a vector space of dimension one is generated by a single (nonzero) vector $\mathbf{v} \in \mathbb{R}^2$. Let $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, then at least one of v_1 and v_2 is nonzero. Suppose that $v_1 \neq 0$. Without loss of generality, we assume that $v_1 = 1$. We define

$$U_1 = \{\text{all dimension subspaces of } \mathbb{R}^2 \text{ generated by } (v_1, v_2), \text{ where } v_1 = 1\}.$$

We can defined U_2 in a similar way. It is easy to find that

$$\mathrm{Gr}_1(\mathbb{R}^2) = U_1 \cup U_2.$$

Note that the intersection $U_1 \cap U_2$ is non-empty, which includes all subspaces generated by $v = (v_1, v_2)$ where $v_1, v_2 \neq 0$. Now let us understand U_i from the geometry: U_1 includes all lines (passing through the origin) except the y-axis and U_2 includes all lines (passing through the origin) except the x-axis. Therefore, their intersection includes all lines except the x-axis and y-axis. Now we go back to the definition of U_i. It is easy to check that U_i can be understood as $\mathbb{R}$. Therefore, we say that $\mathrm{Gr}_1(\mathbb{R}^2)$ is of *dimension one*.

Remark 2.5. Note that this is not the precise definition of $\mathrm{Gr}_1(\mathbb{R}^2)$. The Grassmannian is actually a manifold, of which the dimension is defined by the local charts. add reference

Homework (2 pts). Using the idea above, find a decomposition of $\mathrm{Gr}_1(\mathbb{R}^n)$ and guess the dimension of $\mathrm{Gr}_1(\mathbb{R}^n)$. How about $\mathrm{Gr}_k(\mathbb{R}^n)$ (extra 2 pts)? (If the explanation is given in the language of manifolds, there will be extra 4 pts.)

3. Vector Space

Definition 3.1. A *vector space over* $\mathbb{R}$ is a nonempty set V equipped with two operations: the *addition* $+$ and *scalar multiplication* $\cdot$, subject to the axioms as follows, where $u, v, w \in V$ and $c, d \in \mathbb{R}$,

(1) $u + v \in V$ (closed under addition);
(2) $u + v = v + u$ (commutativity);
(3) $(u + v) + w = u + (v + w)$ (associativity);
(4) there is a vector $0 \in V$ such that $0 + v = v = v + 0$ for arbitrary $v \in V$, and the vector 0 is called *zero vector*;
(5) $cu \in V$ (closed under scalar multiplication);
(6) $c(u + v) = cu + cv$ (distribution law);
(7) $(c + d)u = cu + du$ (distribution law);
(8) $c(du) = (cd)u$ (associativity);
(9) $1u = u$.

An element $v \in V$ is called a *vector*. Similarly, we can define vector spaces over $\mathbb{Q}$ and $\mathbb{C}$.

In this section, we use the notation v for vectors rather than $\mathbf{v}$ as we did in the last section.

Example 3.2. The first example is $\mathbb{R}^n$. Equipped with vector addition and scalar multiplication, clearly, $\mathbb{R}^n$ is a vector space.

The second example is the line $y = x$ in $\mathbb{R}^2$. More precisely, we consider the subset

$$R = \{(x, y) \mid y = x\} \in \mathbb{R}^2.$$

Equipped with the addition of vectors and scalar multiplication of vectors, it is easy to check that R is a vector space.

The third example is the line $y = x + 1$ in $\mathbb{R}^2$. More precisely, we consider the subset

$$S = \{(x, y) \mid y = x + 1\} \in \mathbb{R}^2.$$

As a subset of $\mathbb{R}^2$, there is a natural candidate for addition on S, which is induced from that of $\mathbb{R}^2$. However, it is easy to check that this operation is not closed, which contradicts to axiom (1). More precisely, although $(-1, 0), (0, 1) \in S$, their sum $(-1, 0) + (0, 1) = (-1, 1) \in S$. Therefore, S is a vector space with respect to the addition and scalar multiplication induced from that of $\mathbb{R}^2$. Furthermore, since the line does not pass through the origin, the zero vector is not included in S. Hence, it also does not satisfy Axiom (4).

Now we compare the second and the third example. By drawing the graph of these two lines, we find that they are parallel to each other. It is natural to imagine that the second line $y = x + 1$ should also be a vector space by intuition. The problem is that we have to figure out an appropriate definition for addition and scalar multiplication on S to make it to be a vector space. We only give the definition of addition and scalar multiplication and leave it for the reader to check that it gives a vector space structure on S.

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2 - 1)$$
$$r \cdot (x, y) := (rx, ry - r + 1).$$

This example actually tells us that vector space is a concept of *geometry*. However, we have to use *algebraic approach* to realize it as a vector space.

Definition 3.3. Let $v_1, \ldots, v_p$ be vectors in V. They are *linearly independent* if the equation

$$c_1 v_1 + \cdots + c_p v_p = 0$$

has only trivial solution. Otherwise, they are called *linearly dependent*.

Definition 3.4. A *basis* of V is a set of vectors $\{v_1, \ldots, v_n\}$ such that

(1) the vectors $v_1, \ldots, v_n$ are linearly independent;
(2) given any other vector $v \in V$, the vectors $v_1, \ldots, v_n, v$ are linearly dependent.

Lemma 3.5. *Let $\{v_1, \ldots, v_n\}$ be a basis of V. Then, any vector v can be uniquely expressed as a linear combination of $\{v_1, \ldots, v_n\}$.*

Proof. Let v be a vector in V. By definition of basis, v can be expressed as a linear combination of $\{v_1, \ldots, v_n\}$, i.e. there exist real numbers $a_1, \ldots, a_n$ such that

$$v = a_1 v_1 + \cdots + a_n v_n.$$

Suppose that we have another expression

$$v = b_1 v_1 + \cdots + b_n v_n.$$

If $a_k \neq b_k$ for some k, then

$$\sum_{i=1}^{n} (a_i - b_i) v_i = 0$$

implies that $v_1, \ldots, v_n$ are linearly dependent, which contradicts to the fact that $\{v_1, \ldots, v_n\}$ is a basis. □

Definition 3.6. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Given any vector $v \in V$, the *$\mathscr{B}$-coordinate* of v is $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, if

$$v = a_1 v_1 + \cdots + a_n v_n.$$

The $\mathscr{B}$-coordinate of v is denoted by $[v]_{\mathscr{B}}$.

Theorem 3.7. *If $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_m\}$ are two sets of basis of V, then we have $n = m$.*

Proof. The proof is the same as Theorem 1.13. □

Definition 3.8. Let V be a vector space. We say that the *dimension* of V is n, if the number of vectors in a basis of V is n,

Example 3.9. We consider the set $\mathbb{Q}$. By considering the usual addition and multiplication of numbers, $\mathbb{Q}$ is a vector space (over $\mathbb{Q}$) of dimension one. In this example, we will give the idea of defining new addition and scalar multiplication to make the set $\mathbb{Q}$ as a vector space (over $\mathbb{Q}$) of dimension two, The idea is based on the concept of cardinality of infinite sets.

Definition 3.10. Let V and W be two vector spaces. A *linear transformation* $T : V \to W$ is a map such that the equality holds

$$T(av_1 + bv_2) = aT(v_1) + bT(v_2)$$

for any $v_1, v_2 \in V$ and $a, b \in \mathbb{R}$.

Fixing a basis $\mathscr{B}$ of V (of dimension n) and a basis $\mathscr{C}$ of W (of dimension m), the linear transformation T can be expressed as an $m \times n$ matrix uniquely, denoted by ${}_{\mathscr{B}}[T]_{\mathscr{C}}$. The idea is given as follows. Given any vector $v \in V$, suppose that $v = \sum_{i=1}^{n} a_i v_i$. Then,

$$T(v) = T(\sum_{i=1}^{n} a_i v_i) = a_i \sum_{i=1}^{n} T(v_i).$$

Thus, we only have to find the values of $T(v_i)$. Note that $T(v_i)$ is a vector in W. Since $\{w_1, \dots, w_m\}$ is a basis of W, the vector $T(v_i)$ can be expressed as a linear combination of $\{w_1, \dots, w_m\}$, i.e.

$$T(v_i) = \sum_{j=1}^{m} c_{ji} w_j.$$

Then, the matrix ${}_{\mathscr{B}}[T]_{\mathscr{C}}$ is exactly given by (c_{ij}), i.e.

$$\begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mn} \end{pmatrix}.$$

Example 3.11. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the map

$$T(x, y) = (x, -y).$$

Clearly, T can be written as the matrix

$${}_{\mathscr{E}}[T]_{\mathscr{E}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with respect to the standard basis of $\mathbb{R}^2$.

Now we consider the set $\mathscr{B} = \{v_1 = (1, 1), v_2 = (1, -1)\}$. Clearly, $\mathscr{B}$ is a basis of $\mathbb{R}^2$. Also, it is easy to check that

$$T(v_1) = v_2, \quad T(v_2) = v_1.$$

Therefore, the matrix ${}_{\mathscr{B}}[T]_{\mathscr{B}}$ is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Definition 3.12. Let V be a vector space. A *subspace* W of V is a set such that

(1) the zero vector is in W;
(2) for any $u, v \in W$, any linear combination of u and v is also included in W.

4. Example: Rational Numbers $\mathbb{Q}$

We have learned that a vector space is a set equipped with two operations: addition and scalar multiplication. In this section, we give some examples to the point that a set can be equipped with distinct structures of vector space.

Definition 4.1 (Bijectivity)**.** Let S, T be two sets. A map $f : S \to T$ is *injective*, if $s_1 \neq s_2$, then $T(s_1) \neq T(s_2)$. It is called *surjective*, if for any $t \in T$, there exists $s \in S$ such that $f(s) = t$. A map f is *bijective* if it is both injective and surjective.

Definition 4.2 (Cardinality)**.** Let S be a finite set. The *cardinality* of S is the number of elements in it, and is denoted by $\mathrm{Card}(S)$. Let S and T be two sets (not necessarily to be finite). These two sets are of the same *cardinality* if there exists a bijective map $f : S \to T$.

Let S and T be two finite sets. They are of the same cardinality if and only if $\mathrm{Card}(S) = \mathrm{Card}(T)$. For the infinite case, proving that S and T are of the same cardinality is equivalent to show that there exist injective map $S \to T$ and surjective map $S \to T$.

Definition 4.3. Let S be an infinite set. It is *countable* if the cardinality of S is the same as $\mathbb{Z}$.

Lemma 4.4. *The set of positive integers $\mathbb{Z}_+$ is countable.*

Proof. It is enough to construct a surjective map $f : \mathbb{Z}_+ \to \mathbb{Z}$. The map is given as follows

$$f(n) = \begin{cases} -k, & n = 2k-1 \\ k-1, & n = 2k \end{cases}$$

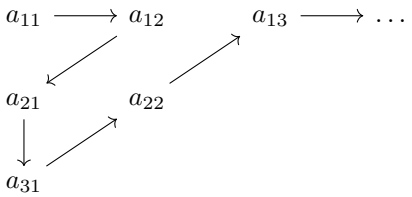
□

Theorem 4.5. $\mathbb{Q}$ *is countable,*

Proof. It is enough to prove that the set of positive rational numbers $\mathbb{Q}_+$ is countable. More precisely, we will construct a surjective map

$$f : \mathbb{Z}_+ \to \mathbb{Q}_+.$$

Define $a_{ij} := \frac{i}{j}$, where i, j are positive integers.



Then, the map f is defined as

$$f(1) = a_{11},\ f(2) =_{12},\ f(3) = a_{21}, \ldots\ .$$

In this way, we construct a surjective map $f : \mathbb{Z}_+ \to \mathbb{Q}_+$. □

Corollary 4.6. *The set $\mathbb{Z}^2$ is countable, and therefore, $\mathbb{Q}^2$ is countable.*

Proof. To prove $\mathbb{Z}^2$ is countable, it is enough to construct a surjective map

$$\mathbb{Z}_+ \to \mathbb{Z}_+^2.$$

The idea of constructing this map is exactly the same as the proof of Theorem 4.5.

Now we consider the set $\mathbb{Q}^2$. By Theorem 4.5, there exists a surjective map $f : \mathbb{Z} \to \mathbb{Q}$. Therefore, we have a surjective map $f \times f : \mathbb{Z}^2 \to \mathbb{Q}^2$. This implies that the cardinality of $\mathbb{Q}^2$ is the same as $\mathbb{Z}^2$. Since $\mathbb{Z}^2$ is countable, the set $\mathbb{Q}^2$ is countable. □

Homework (4 pts)**.** This homework is a continuation of Example 3.9. By Corollary 4.6, we know that there exists a bijective map $\mathbb{Q} \to \mathbb{Q}^2$. Using this property, try to define addition and scalar multiplication on $\mathbb{Q}$ to make it to be a vector space of dimension 2 over $\mathbb{Q}$. (It is not necessary to give a general formula for the addition. It is enough to give some examples about the operation you want to define on $\mathbb{Q}$.) Is it possible to make $\mathbb{Q}$ to be a vector space of dimension n over $\mathbb{Q}$? Is it possible to make it to be a vector of infinite dimension over $\mathbb{Q}$ (extra 2 pts)?

Theorem 4.7. $\mathbb{R}$ *is not countable.*

Proof. It is enough to prove that $\mathbb{R} \cap [0,1]$ is not countable. Suppose that $\mathbb{R} \cap [0,1]$ is countable. Then, there exists a surjective map $f : \mathbb{Z}_+ \to \mathbb{R} \cap [0,1]$. For each $n \in \mathbb{Z}_+$, denote by a_n the n-th digit of $f(n)$. For example, if $f(1) = 3.2314\ldots$, then $a_1 = 2$, and if $f(3) = \pi$, then $a_3 = 1$. Now we can define a new real number as follows

$$b = 0.b_1 b_2 \ldots,$$

where $b_i = a_i + 1 \bmod 10$. Since the n-th digit of b is not the same as $f(n)$, we have $b \neq f(n)$. Therefore, $b \in f(\mathbb{Z})$ and this contradicts to the assumption that f is surjective. □

5. Example: Polynomials

Definition 5.1. A *polynomial* is an infinite sequence of (real) numbers

$$(a_0, a_1, \ldots, a_i, \ldots)$$

such that there exists a nonnegative integer n, and when $i > n$, we have $a_i = 0$. If n is the smallest integer satisfying this property, then we say that the *degree of the polynomial* is n. Denote by $\mathbb{P}$ the set of all polynomials (with coefficients in $\mathbb{R}$). Denote by $\mathbb{P}_n$ the set of all polynomials of degree $\leq n$.

Definition 5.2. The *addition* of two polynomials $(a_0, a_1, \ldots, a_i, \ldots)$ and $(b_0, b_1, \ldots, b_i, \ldots)$ is defined as

$$(a_0, a_1, \ldots, a_i, \ldots) + (b_0, b_1, \ldots, b_i, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots, a_i + b_i, \ldots).$$

The *scalar multiplication* is defined as

$$r \cdot (a_0, a_1, \ldots, a_i, \ldots) = (ra_0, ra_1, \ldots, ra_i, \ldots).$$

where r is a real number.

The *multiplication* is defined as

$$(a_0, a_1, \ldots, a_i, \ldots) \times (b_0, b_1, \ldots, b_i, \ldots) = (c_0, c_1, \ldots, c_i, \ldots),$$

where $c_i = \sum_{j=1}^{i} a_j b_{i-j}$.

Lemma 5.3. *The set of polynomials $\mathbb{P}$ is a vector space (over $\mathbb{R}$), and this vector space is of infinite dimension. Similarly, $\mathbb{P}_n$ is a vector space of dimension $n+1$.*

Proof. The proof is left to the reader. Note that although we give the definition of multiplication, it does not play a role in the proof of vector spaces. □

Notation 5.4. If we introduce a variable x and consider a_i as the coefficient of the term x^i, the above definition coincides with the definition we familiar with

$$(a_0, a_1, \ldots, a_i) \to \sum_{i=0}^{\infty} a_i x^i.$$

Although the precise definition of a polynomial is given as a sequence of numbers, we use the classical notation $f(x)$ to do calculations for convenience.

Definition 5.5 (Differentiation)**.** Let $f(x)$ be a continuous function on its domain $D \subseteq \mathbb{R}$. Let $a \in D$. We say that $f(x)$ is *differentiable* at $x = a$, if the limit

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists. Denote the limit by $f'(a)$. If $f(x)$ is differentiable at every point in D, we say that the function $f(x)$ is *differentiable*, and denote by $f'(x)$ its differentiation.

Lemma 5.6. *Let $f(x)$ and $g(x)$ be two differentiable functions on D. Then, $f(x) + g(x)$ and $rf(x)$ are also differentiable functions on D, where r is a real number.*

Proof. □

Definition 5.7 (Integration)**.** It is very hard to give the precise definition of integration, which comes from *Riemann sum*. Therefore, in this note, the integration is regarded as the inverse operation of differentiation.

We say that $F(x)$ is an *indefinite integration* of $f(x)$ if $F'(x) = f(x)$. We would like to use the following notation

$$F(x) = \int f(x)dx.$$

Example 5.8. Let $f(x) = x^n$. We will calculate that $f'(x) = nx^{n-1}$ in this example. We first review a formula

$$x^n - a^n = (x - a)(x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}).$$

Next, by definition,

$$\begin{aligned} f'(a) &= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= \lim_{x \to a} \frac{x^n - a^n}{x - a} \\ &= \lim_{x \to a} (x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}) \\ &na^{n-1}. \end{aligned}$$

Clearly, $f(x) = x^n$ is differential at any real number $x = a$. Thus, we have $f'(x) = nx^{n-1}$ by substituting a by x.

With respect to the above calcution, the integration of $f(x)$ is $\frac{1}{n+1}x^{n+1}$.

Now we consider two maps:

$$\begin{aligned} T_x &: \mathbb{P} \to \mathbb{P} \\ T_\partial &: \mathbb{P} \to \mathbb{P} \end{aligned}$$

such that $T_x(f(x)) = xf(x)$ and $T_\partial(f(x)) = \frac{\partial f(x)}{\partial x}$. The map T_x can be regarded as the map for integration, while the map T_∂ is regarded as the map for differentiation.

Lemma 5.9. *The map T_x, T_∂ are linear transformations.*

Proof. We only give the proof for T_x. The proof for T_∂ follows from Lemma 5.6. Let $f(x)$ and $g(x)$ be two polynomials, and let a, b be two real numbers. We have

$$\begin{aligned} T_x(af(x) + bg(x)) &= x(af(x) + bg(x)) \\ &= axf(x) + bxg(x) \\ &= aT_x(f(x)) + bT_x(g(x)). \end{aligned}$$

Therefore, T_x is a linear transformation. □

In the rest of this section, we want to give an example about a difference between finite and infinite spaces. Let A and B be two square matrix of size $n \times n$. The *commutator* $[A, B]$ of A and B is defined as

$$[A, B] := AB - BA.$$

We have a very interesting property:

Lemma 5.10. *Given any square matrices A and B, their commutator cannot be the identity matrix I.*

Proof. Let A be a square matrix. We define the *trace* of A as

$$\mathrm{tr}(A) := \sum_{i=1}^{n} a_{ii}.$$

It is easy to check that Add the proof

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Therefore, Add the proof

$$\mathrm{tr}(AB - BA) = \mathrm{tr}(AB) - \mathrm{tr}(BA) = 0.$$

However, the trace of the identity matrix is n. This finishes the proof of this lemma. □

Corollary 5.11. *Let V be a finite dimensional vector space with linear transformations $S, T : V \to V$. Then, we have*

$$T \circ S - S \circ T \neq I,$$

where I is the identity map $V \to V$.

Proof. This is a direct result of Lemma 5.10. □

Now we go back to the linear transformations T_x and T_∂. Note that these two linear transformations are about infinite dimensional vector space.

Lemma 5.12. *We have*

$$T_\partial \circ T_x - T_x \circ T_\partial = I,$$

where I is the identity map of $\mathbb{P}$.

Proof. We calculate the commutator of T_x and T_∂:

$$\begin{aligned}
(T_\partial \circ T_x - T_x \circ T_\partial)(f(x)) &= T_\partial(T_x(f(x)) - T_x(T_\partial(f(x))) \\
&= T_\partial(x(f(x)) - T_x(\frac{\partial f(x)}{\partial x}) \\
&= f(x) + x\frac{\partial f(x)}{\partial x} - x\frac{\partial f(x)}{\partial x} \\
&= f(x).
\end{aligned}$$

The above calculation holds for an arbitrary polynomial $f(x)$. Therefore,

$$T_\partial \circ T_x - T_x \circ T_\partial = I.$$

□

In this section, we only consider polynomials with a single variable x. In later sections, we prefer to use the notation $\mathbb{R}[x]$ for the set of polynomials with coefficients $\mathbb{R}$. In many cases, the variable is not unique and we probably have more than one variable. For example,

$$f(x, y) = x^2 - y^3$$

is a polynomial with variable x, y. We can define the set of polynomials with two variables x, y and denote it by $\mathbb{R}[x, y]$. This is also a vector space over $\mathbb{R}$. With the same idea, we can define the set of polynomials with multi-variables.

Homework (2 pts)**.** Find general formulas of roots of cubic polynomials and quadratic polynomials. It is a little bit embarrased to leave this problem. However, it seems that I do not have enough time to introduce the Galois theory.

6. Algebraic Structures

6.1. Operations.

Definition 6.1 (Unitary operations and Binary operations)**.** Let S be a set. A *unitary operation* on S is a map

$$\Phi : S \to S.$$

A *binary operation* on S is a map

$$\Psi : S \times S \to S.$$

A binary operation Ψ is *commutative*, if for arbitrary elements s_1, s_2, we have

$$\Psi(s_1, s_2) = \Psi(s_2, s_1).$$

A binary operation Ψ is *associative*, if for arbitrary elements $s_1, s_2, s_3 \in S$, we have

$$\Psi(s_1, \Psi(s_2, s_3)) = \Psi(\Psi(s_1, s_2), s_3).$$

Remark 6.2. The associativity can be understood as the equality of the following two compositions

$$S \times S \times S \xrightarrow{I\times\Psi} S \times S \xrightarrow{\Psi} S$$

$$S \times S \times S \xrightarrow{\Psi\times I} S \times S \xrightarrow{\Psi} S,$$

where $I : S \to S$ is the identity map. This property can be understood as the *commutativity* of the following diagram

$$\begin{array}{ccc} S \times S \times S & \xrightarrow{I\times\Psi} & S \times S \\ \downarrow{\scriptstyle \Psi\times I} & & \downarrow{\scriptstyle \Psi} \\ S \times S & \xrightarrow{\Psi} & S \end{array}$$

Add reference

Example 6.3. Let $S = \mathbb{R}$. The map

$$\begin{aligned} - :& \mathbb{R} \to \mathbb{R} \\ & a \to -a \end{aligned}$$

is a unitary operation. Furthermore, the addition and multiplication

$$+, \times : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

are binary operations.

Definition 6.4. Let S be a set with a binary operation Ψ (resp. unitary operation Φ). A subset T is *closed under* Ψ (resp. Φ) if $\Psi(T, T) \subseteq T$ (resp. $\Phi(T) \subseteq T$). We also say that Ψ is well-defined on T.

Example 6.5. Consider the pair $(\mathbb{Z}_+, +)$, positive integers with addition. The sum of any two positive integers is still a positive integer. Therefore, $\mathbb{Z}_+$ is closed under $+$.

Now consider $(\mathbb{Z}_-, \times)$, negative integers with multiplication. The multiplication of two negative integers is positive. Thus, $\mathbb{Z}_-$ is not closed under $-$.

For the third example, we consider the pair $(\mathbb{R}\backslash\{0\}, \times)$. It is easy to see that the multiplication of any two nonzero real numbers is a real number. Therefore, $\mathbb{R}\backslash\{0\}$ is closed under $\times$.

Definition 6.6 (Identity element)**.** Let S be a set with a binary operation Ψ. An element e is called an *identity element* (w.r.t Ψ), if for any $s \in S$, we have

$$\Psi(e, s) = \Psi(s, e) = s.$$

The identity element of $(\mathbb{R}, +)$ is 0, and the identity element of $(\mathbb{R}/\{0\}, \times)$ is 1.

Lemma 6.7. *Let S be a set with a binary operation Ψ. If the identity element exists, then it is unique.*

Proof. Suppose that e_1 and e_2 are identity elements. Then, $\Psi(e_1, e_2) = e_2$ because e_1 is an identity element. Also, $\Psi(e_1, e_2) = e_1$ because e_2 is an identity element. Thus, $e_2 = e_1$. □

Definition 6.8 (Inverse element)**.** Let S be a set with a binary operation Ψ. Suppose that there exists an identity element e for Ψ. Let $s \in S$. An *inverse element* of s (w.r.t e) is an element $t \in S$ such that

$$\Psi(t, s) = \Psi(s, t) = e.$$

Example 6.9. We take $\mathbb{R}$ as an example and consider some special operations on it. The first binary operation is the addition: $+$. Clearly, $\mathbb{R}$ is closed under the operation $+$. The identity element of $+$ is 0, since

$$a + 0 = 0 + a = a.$$

Thus, the inverse element of an element $r \in \mathbb{R}$ is easy to see to be $-r$.

The second binary operation is the multiplication: $\times$, which is also well-defined on $\mathbb{R}$. The identity element in this case is 1. However, not every element has an inverse element. An element $r \in \mathbb{R}$ has an inverse if and only if $r \neq 0$. Note that if we take off 0, and consider the set $\mathbb{R}\backslash\{0\}$, then every element has an inverse.

Lemma 6.10. *Let S be a set with an associative binary operation Ψ. Suppose that there exists an identity element e. Given $s \in S$, if its inverse element exists, then it is unique.*

Proof. Suppose that t_1 and t_2 are inverse elements of s. Then,

$$t_1 = \Psi(t_1, e) = \Psi(t_1, \Psi(s, t_2)) = \Psi(\Psi(t_1, s), t_2) = \Psi(e, t_2) = t_2.$$

□

6.2. **Groups.**

Definition 6.11 (Group)**.** Let G be a set with a well-defined binary operation Ψ. We say that (G, Ψ) is a group if it satisfies the following properties:

(1) the binary operation Ψ is associative;
(2) there exists an identity element e;
(3) for any element $g \in G$, it has an inverse.

If Ψ is commutative, then (G, Ψ) is called an *abelian group*. For simplicity, we use the notation G for the group and omit the notation for the operation if there is no ambiguity.

Example 6.12. The real number equipped with the addition is a group, i.e. $(\mathbb{R}, +)$ is a group. However, $(\mathbb{R}, \times)$ is not a group because 0 does not have an inverse. As we discussed above, $(\mathbb{R}\backslash\{0\}, \times)$ is a group. There are examples of non-abelian groups. One of the basic examples comes from matrices. More precisely, the multiplication of matrices is not commutative. Since we did not define the determinant of matrices. We will not discuss this example in detail.

Definition 6.13. Let (G, Ψ) be a group. A *subgroup* of G is a subset H such that

(1) H is closed under Ψ;
(2) $e \in H$;
(3) if $h \in H$, then $h^{-1} \in H$.

Construction 6.14. In this example, we discuss a popular way to give subgroups of G. Given elements $g_1, \dots, g_n \in G$, we define the subset

$$H := \langle g_1, \dots, g_n \rangle := \{g_{j_1}^{i_1} \dots g_{j_m}^{i_m}, m \in \mathbb{Z}_+, i_1, \dots, i_m \in \mathbb{Z}, j_1, \dots, j_m \in \mathbb{Z} \cap [1, 10]\}$$

of G. Note that the above set includes all distinct elements in G which can be written in the form $g_1^{i_1} \dots g_n^{i_n}$.

We claim that the set $\langle g_1, \dots, g_n \rangle$ is a subgroup of G. First, given two elements in this set $\langle g_1, \dots, g_n \rangle$, clearly their multiplications is also included in it. Then, let $m = 1$ and $i_1 = j_1 = 0$. We get the identity element e. Finally, let $g_{j_1}^{i_1} \dots g_{j_m}^{i_m}$ be an element in the set. Clearly, $g_{j_m}^{-i_m} \dots g_{j_1}^{-i_1}$ is the inverse element. In this case, we say that H is the subgroup of G *generated* by $g_1, \dots, g_n$.

The *generation* is a very useful way to generate *sub-algebraic structure*. Recall that when we discuss vector space, we use a similar idea to construct subspaces (see Lemma 2.2). Since there are two operations in the definition of vector spaces (addition and scalar multiplication), given vectors $v_1, \dots, v_p$, we define a subset *generated* by those vectors and closed under the given operation. In this way, we obtain a subspace. For subgroups, the idea is exactly the same. We still give a class of elements $g_1, \dots, g_n$, and use them to generate a subset of G which is large enough to includes the identity element, inverse elements and closed under multiplication.

Definition 6.15 (Homomorphisms of Groups)**.** Let G, H be two groups. A *homomorphism* $f : G \to H$ is a map of set such that

$$f(g_1 g_2) = f(g_1) f(g_2)$$

for arbitrary $g_1, g_2 \in G$.

Remark 6.16. We compare *homomorphisms of groups* with *linear transformations of vector spaces* we defined in Definition 3.10, As we already know, linear transformation *preserves* addition and scalar multiplication, which means that the result does not change whenever we do the operations. The idea is similar for homomorphisms of groups. Since there is only one operation for groups, a homomorphism is a map which *preserves* the operation. Generally speaking, given any algebraic structure (we will see for rings and modules), a homomorphism always means a map *preserves* the operations which defines the algebraic structure.

We consider a special homomorphism of groups:

Definition 6.17 (Conjugation)**.** Let $a \in G$ be an element in a group G. We define

$$\gamma_a : G \to G, \quad g \to aga^{-1},$$

which is a map of sets. This map is called a *conjugation*.

Lemma 6.18. *The conjugation map is a homomorphism of groups.*

Proof. Let g_1, g_2 be elements in G. Let γ_a be the conjugation map. We have

$$\gamma_a(g_1 g_2) = a(g_1 g_2)a^{-1} = ag_1(a^{-1}a)g_2 a^{-1} = \gamma_a(g_1)\gamma_a(g_2).$$

Thus, γ_a is a homomorphism. □

Example 6.19 (Cyclic Group)**.** Let S be a set. An *equivalence relation* $\sim$ (sometimes we also use the notation $\sim_{\mathcal{R}}$) is a relation on S, which satisfies the following properties:

(1) for any $a \in S$, we have $a \sim a$;
(2) if $a \sim b$, then $b \sim a$;
(3) if $a \sim b$ and $b \sim c$, then $a \sim c$.

If $a \sim b$, then we say that a *is equivalent to* b. Let $a \in S$. We define $[a]_{\mathcal{R}} \subseteq S$ to be the subset including all elements equivalent to a. The subset $[a]_{\mathcal{R}}$ is called an *equivalence class* of S, and a is called an

representative of this set. Clearly, the set $[a]_{\mathcal{R}}$ could have distinct representatives. For example, if $a \sim b$, then $[a]_{\mathcal{R}} = [b]_{\mathcal{R}}$, which means that a and b are both representatives of this set. Now we define

$$S_{\mathcal{R}} := \{ \text{ all equivalence classes in } S \text{ with respect to the equivalence relation } \mathcal{R}.\}$$

The obvious equivalence relation on $\mathbb{Z}$ is given by the equality " $=$ ". In this case, each equivalence class contains only one element.

Now we consider another equivalence relation on $\mathbb{Z}$. We fix a positive integer n. Given two integers a, b, we say $a \sim_n b$, if $a - b \equiv 0 \bmod n$, i.e. $n|(a-b)$. Then, there are exactly n equivalence classes, which are classified by the remainders modulo n:

$$[0]_n, [1]_n, \ldots, [n-1]_n.$$

Thus, the set of all equivalence classes $\mathbb{Z}_n$ has n elements.

Now we consider how to define *addition* on this set. Let $a \in [i]_n$ and $b \in [j]_n$, i.e.

$$a \equiv i \bmod n, \quad b \equiv j \bmod m.$$

Then, we have

$$a + b \equiv i + j \bmod n.$$

This property holds for arbitrary elements a and b. Therefore, the addition

$$[i]_n + [j]_n = [i+j]_n$$

is well-defined. It is easy to check that the identity element is $[0]_n$, and the inverse element of $[i]_n$ is $[-i]_n$ (or $[n-i]_n$). Therefore, $(\mathbb{Z}_n, +)$ is a group, which is called a *cyclic group*.

The idea of equivalence classes is very important. It also gives a very important class of groups called *quotient groups*.

Construction 6.20 (Quotient Groups)**.** Let G be a group. Given a subgroup H of G, we say that H is *normal* if for arbitrary $g \in G$, we have $gH = Hg$ (as sets). Now we use H to give a definition of equivalence relation. We say $g_1 \sim g_2$ if $g_1 H = g_2 H$. It is easy to check that this is an equivalence relation. We use the notation G/H for the set of equivalence classes, and an element in G/H is written as gH, where $g \in G$. The *multiplication* on G/H is given as

$$g_1 H \cdot g_2 H = g_1 (g_2 H g_2^{-1}) g_2 H = (g_1 g_2) H.$$

It is easy to check that the multiplication is well-defined and gives a group structure on the set G/H.

6.3. **Rings.**

Definition 6.21 (Ring)**.** Let R be a set with two well-defined binary operations Ψ, Φ. We say that (R, Ψ, Φ) is a *ring*, if it satisfies the following properties:

(1) (R, Ψ) is an abelian group, and the identity element is denoted by 0;
(2) the operation Φ is associative;
(3) there exists an identity element for Φ, which is denoted by 1;
(4) the operations satisfy the *distributive law*

$$\Phi(a, \Psi(b, c)) = \Psi(\Phi(a, b), \Phi(a, c)).$$

If Φ is commutative, then (R, Ψ, Φ) is called a *commutative ring*. If there is no ambiguity, we omit the operations and say R is a ring. The operation Ψ is usually called the *addition*, while Φ is called the *multiplication*.

Example 6.22 (Polynomial Ring)**.** Let $\mathbb{R}[x]$ be the set of polynomials with coefficients in $\mathbb{R}$. The first operation is given as the addition of polynomials, which is denoted by $+$, and the second operation is defined as the multiplication of polynomials and denoted by $\times$. It is easy to check that with these two operations, $\mathbb{R}[x]$ is a ring, which is called the *polynomial ring*. The coefficient can be taken to be $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$. Furthermore, both the operations are commutative.

Example 6.23 (Matrices)**.** Let $M_n(\mathbb{R})$ be the set of $n \times n$-matrices with coefficients in $\mathbb{R}$. Under the addition of matrices, $(M_n(\mathbb{R}), +)$ is an abelian group, of which the identity element 1 is the zero matrix. The second operation we take is the multiplication of matrices, of which the identity element is the identity matrix. Therefore, $(M_n(\mathbb{R}), +, \times)$ is a ring.

Definition 6.24 (Subrings)**.** Let R be a ring. A *subring* of R is a subset S such that S is closed under addition and multiplication and contains the identity element A.

Definition 6.25 (Ideals)**.** Let R be a ring. An *ideal* of R is a subset I such that

(1) $0 \in I$;
(2) if $a, b \in I$, then $a + b \in I$;
(3) if $r \in R$ and $x \in I$, then $rx \in I$.

Construction 6.26 (Quotient Ring)**.** We will construct quotient ring in this example. Let R be a ring, and let I be an ideal. We define an equivalence relation on R based on I. Let $a, b \in R$. We say $a \sim b$ if $a - b \in I$. Check that this is actually an equivalence relation. Denote by R/I the set of equivalence classes, and we use the notation $a+I$ for an element in R/I. If $a-b \in I$, then $a+I = b+I$. The addition on R/I is defined as

$$(a + I) + (b + I) = (a + b) + I.$$

The multiplication is given as

$$(a + I)(b + I) = ab + I.$$

We claim that the above definition gives a ring structure $(R/I, +, \times)$, and this ring is called the *quotient ring*.

Homework (3 pts)**.** Let $R = \mathbb{C}[x]$ be the polynomial ring. Given a polynomial $f(x) \in \mathbb{C}[x]$, denote by $I = (f(x))$ the ideal generated by the single element $f(x)$. Answer the following questions:

(1) Explain how a single element $f(x)$ generates an ideal I.
(2) Prove that R/I is a ring (identity elements in R/I, define addition and multiplication, and check all of the conditions).
(3) If $f(x)$ is of degree n, prove that R/I is isomorphic to $\mathbb{C}^n$ as vector spaces over $\mathbb{C}$.

Then, R/I is a quotient ring.

Definition 6.27 (Homomorphisms of Rings)**.** Let R, S be two rings. A *homomorphism* $f : R \to S$ is a map such that

$$f(r_1 + r_2) = f(r_1) + f(r_2), \quad f(r_1 r_2) = f(r_1) f(r_2).$$

Example 6.28. Consider the set of complex numbers $\mathbb{C}$. Recall that complex numbers can be written in the form $a + bi$, where i is the imaginary unit and a, b are real numbers, and the multiplication is defined as

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Now we consider a set of matrices

$$M = \{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \}.$$

We claim that both $\mathbb{C}$ and M are rings. Now we define a map

$$f : \mathbb{C} \to M$$

as

$$a + bi \to \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

For addition,

$$\begin{aligned} f((a+bi)+(c+di)) &= f((a+c)+(b+d)i) \\ &= \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \\ &= f(a+bi)+f(c+di). \end{aligned}$$

For multiplication,

$$\begin{aligned} f((a+bi)(c+di)) &= f((ac-bd)+(ad+bc)i) \\ &= \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \\ &= f(a+bi)f(c+di). \end{aligned}$$

Therefore, this map f is a homomorphism of rings.

Recall that the cancellation law says that if $ab = ac$, then we have $b = c$. However, for a general ring R, the cancellation law does not hold, for instance the ring of matrices. This property induces the following definition:

Definition 6.29. A ring R is a *domain* if $1 \neq 0$ and the cancellation law holds. In some textbooks, it is called an *integral domain*.

Lemma 6.30. *Let R be a domain. If $a, b \in R$ be two nonzero elements, then $ab \neq 0$.*

Proof. Suppose that if there exists a, b such that $ab = 0$, then we have $a \cdot b = a \cdot 0 = b \cdot 0$. Since the cancellation law holds, then at least one of a, b is zero. Therefore, if $a, b \neq 0$, then $ab \neq 0$. □

Definition 6.31 (Field)**.** A ring R is a *field* if for any nonzero element $r \in R$, there exists an element s such that $s \cdot r = 1 = r \cdot s$. The element s is called the *inverse element* of r under the multiplication.

Construction 6.32. Let R be a ring. In this example, we construct a field associated to R. Consider the set R^2. We define a relation on R^2 in the following ways. If $r_1 s_2 = r_2 s_1$, then we say that

$$(r_1, s_1) \sim (r_2, s_2).$$

It is easy to check that this is an equivalence relation. If (r, s) is an element in R^2, we denote by $[r, s]$ the corresponding equivalence class. Let F be the set of equivalence classes under the equivalence relation defined above. Now we want to give a ring structure for F, and the most important part is to figure out the addition and multiplication. Let $[r_1, s_1]$ and $[r_2, s_2]$ be two elements in F. The two operations are defined as follows:

$$\begin{aligned} [r_1, s_1] + [r_2, s_2] &:= [r_1 s_2 + r_2 s_1, s_1 s_1] \\ [r_1, s_1] \cdot [r_2, s_2] &:= [r_1 r_2, s_1 s_2]. \end{aligned}$$

It is easy to check that the zero element is $[0, 1]$, the identity element is $[1, 1]$ and the operations satisfy all of the conditions in the definition of rings. Actually it is a field. This field is usually denoted by $\mathrm{Frac}(R)$, and called the *fraction field of R*.

For example, the fraction field of $\mathbb{Z}$ is exactly $\mathbb{Q}$. In this case, the notation $[r, s]$ corresponds to the usual fraction $\frac{r}{s}$, where r is regarded the numerator and $s$is regarded as the denominator. Furthermore, if R is a domain, then $\mathrm{Frac}(R)$ can always be understood in this way, and there is a natural injection $R \to \mathrm{Frac}(R)$.

However, when R is not a domain, this property does not hold.

Homework (3 pts)**.** Consider the set $A = \{(a, b) \mid a, b \in \mathbb{R}\}$ and answer the following questions.

(1) We define the addition $+$ on A the same as for vectors. Prove that $(A, +)$ is an abelian group.

(2) We define the multiplication $\times$ as follows

$$(a_1, b_1) \times (a_2, b_2) := a_1a_2 + (a_1b_2 + a_2b_1)x.$$

Show that $(A, +, \times)$ is a ring.
(3) Based on the above definition, show that A is not a domain.
(4) Finally, calculate $\mathrm{Frac}(A)$ and check that whether the natural map $A \to \mathrm{Frac}(A)$ is an injection.

Lemma 6.33. *A domain with finitely many elements is a field.*

Proof. Take an element $r \in R$. We define a map

$$T_r : R \to R, \quad s \to rs.$$

Note this is only a map and it is not a homomorphism of rings in general. Suppose that there exists two elements s_1, s_2 such that $T_r(s_1) = T_r(s_2)$. Therefore, $rs_1 = rS_2$. Then, by cancellation law, we have $s_1 = s_2$. Therefore, the map T_r is injective. Since R is a finite set, T_r must be surjective. Therefore, there exists an element $s \in \mathbb{R}$ such that $rs = 1$. This gives an inverse element for r. The above argument holds for arbitrary element $r \in R$. Therefore, R is a field. □

Definition 6.34 (Module)**.**

Remark 6.35 (Algebraic Structures)*.*

7. Groups: Cyclic Groups and Permutation Groups

Let $[n] = \{1, \ldots, n\}$ be the finite set with integers from 1 to n.

Definition 7.1. Denote by S_n the set of bijective maps of $[n]$. An element in S_n is called a *permutation*.

Lemma 7.2. *The cardinality of S_n is $n!$.*

Proof. This is a combinatorial problem. If we want to construct a bijective map of $[n]$, we have to figure out the number of choices for each $f(i)$. Without loss of generality, we start from $i = 1$. For $f(1)$, we have n choices because we just have to choose one number from $[n]$ to be its value. After we fix the value of $f(1)$, we consider $f(2)$. Since the map we want should be bijective, the number of choices for $f(2)$ is $n - 1$, which does not include $f(1)$. Inductively, the number of chocies for $f(3)$ is $n - 2$ and so on. Thus, the total number we get is $n! = n \cdot (n - 1) \cdot \ldots \cdot 1$. □

Lemma 7.3. *Let f and g be two elements in S_n. Then, the composite $f \circ g$ is also an element in S_n. Then, the composite can be regarded as a binary operation on the set S_n.*

Proof. Note that f and g are bijective maps of the set $[n]$. We have to show that $f \circ g$ is also a bijective map of $[n]$.

For the injectivity, if $(f \circ g)(i) = (f \circ g)(j)$, then $f(g(i)) = f(g(j))$. Since f is injective, then $g(i) = g(j)$. Since g is injective, we have $i = j$. Thus $f \circ g$ is injective.

For the surjectivity, given an element i, since f is surjective, there exists j such that $f(j) = i$. Since g is also surjective, there exists k such that $g(k) = j$. Thus,

$$(f \circ g)(k) = f(g(k)) = f(j) = i,$$

and this proofs the surjectivity of $f \circ g$. □

Lemma 7.4. *Let I be the identity map of $[n]$, which is clearly a bijective map. Then, for any element $f \in S_n$, we have*

$$f \circ I = I \circ f = f.$$

Proof. Note that for any element $j \in [n]$, we have $I(j) = j$. Then,

$$(f \circ I)(j) = f(I(j)) = f(j), \quad (I \circ f)(j) = I(f(j)) = f(j).$$

□

Lemma 7.5. *Given a bijective map f of $[n]$, there exists a bijective map g of $[n]$ such that*

$$f \circ g = I = g \circ f.$$

Proof. We define g as follows: $g(b) = a$ if $f(a) = b$. Clearly, g is a bijective map. For any element $j \in [n]$, we have

$$(f \circ g)(j) = f(g(j)) = j.$$

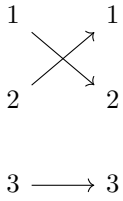
Thus, $f \circ g = I$. The same argument holds for $g \circ f$. □

The above lemmas give us the following theorem.

Theorem 7.6. *The set S_n is a group, of which the binary operation is given by the composite of maps.*

Proof. Clearly, the composition of maps is associative. By Lemma 7.4, the identity map I is the identity element. Lemma 7.5, any bijective map has an inverse. Thus, S_n is a group. □

Notation 7.7. We introduce another notation for permutations. We first consider an example. Let $\sigma : [3] \to [3]$ be the following map



The diagram can be regarded as

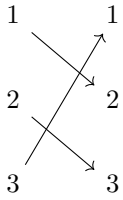
$$1 \to 2 \to 1, 3 \to 3.$$

This diagram is understood as σ sends 1 to 2, then sends 2 to 1. Finally, σ sends 3 to itself. We move a step further

$$(12)(3),$$

where the first bracket means that $1 \to 2 \to 1$ and the second bracket means that $3 \to 3$. When the number go back to a existing one, we omit the information from the bracket. Sometimes, we also omit the bracket which includes only one number, i.e. $(12)(3) = (12)$. Here is another explanation of this notation. Since $\sigma(3) = 3$, we say that the element 3 is *fixed* by σ. Then, the notation $\sigma = (12)$ only includes numbers which are not fixed by σ.

For the second example, consider $\tau : [3] \to [3]$



Then, $\tau = (123)$. Note that $(123) \neq (132)$.

Using new notation to give multiplication

Definition 7.8 (r-cycles)**.** A bijective map includes only one bracket with r numbers is called a *r-cycle*. A *transposition* is a 2-cycle.

Definition 7.9 (Disjointness)**.** Given two permutations σ, τ, we say that σ and τ are *disjoint* if whenever $\sigma(i) \neq i$, then $\tau(i) = i$, and whenever $\tau(j) \neq j$, then $\sigma(i) = i$.

Theorem 7.10. *Every permutation can be written as a product of disjoint r-cycles. This presentation is unique up to permutation of cycles. More precisely, let σ be a permutation. It can be written as $\sigma = \alpha_1 \dots \alpha_l$, where α_i is an r_i-cycle for some positive integer r_i.*

Proof. The proof is given by induction and based on the construction of the notation. Let σ be a permutation in S_n. If for any element k, $\sigma(k) = k$, then $\sigma = (1)$ is the identity element. Therefore, we suppose that there exists an element k such that $\sigma(k) \neq k$. Then, we obtain the following sequence of infinite numbers

$$k, \sigma(k), \dots, \sigma^i(k), \dots.$$

We claim that there exists a positive integer $1 < j \leq n$ such that $k = \sigma^j(k)$ and $k \neq \sigma^{j-1}(k)$. This claim is easily proved by *pigeon hole principle*. Furthermore, $\sigma^{j+i}(k) = \sigma^i(k)$, which implies that the sequence is periodic. These numbers form a j-cycle

$$\alpha_1 = (k\ \sigma(k) \dots \sigma^j(k)).$$

Let $Y := [n] \backslash \{k, \sigma(k), \dots, \sigma^j(k)\}$. Define a permutation σ' in S_n as follows: if $s \in Y$, then $\sigma'(s) := \sigma(s)$; if $s \neq Y$, then $\sigma'(s) = s$. It is easy to check that σ' and α_1 are disjoint, and furthermore, $\sigma = \alpha_1 \sigma'$. Then, by induction, σ' can be written as a product of disjoint cycles $\sigma' = \alpha_2 \dots \alpha_l$. Thus, $\sigma = \alpha_1 \dots \alpha_l$ can be written as a product of disjoint cycles. □

Theorem 7.11. *Any permutation can be written as a product of transpositions.*

Proof. By Theorem 7.10, we only have to show that any r-cycles can be written as a product of transpositions. The proof is based on induction. Let $\sigma = (i_1 \dots i_r)$ be an r-cycle. We have

$$\sigma = (i_1 \dots i_r) = (i_1 i_2)(i_2 \dots i_r) = \tau \sigma',$$

where $\tau = (i_1 i_2)$ is a transposition and $\sigma' = (i_2 \dots i_r)$ is a $(r-1)$-cycle. By induction, we get the result. □

Definition 7.12. A *partition* of n is a finite nonincreasing sequence of positive integers, of which the sum is n. More precisely, let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_l)$ be a sequence of positive integers such that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l, \text{ and } \sum_{i=1}^{l} \lambda_i = n,$$

then $\boldsymbol{\lambda}$ is a partition of n. A permutation σ is of *type* $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_l)$, if $\sigma = \alpha_1 \dots \alpha_l$ and α_i is an λ_i-cycle.

Recall that a conjugation γ_τ of S_n is defined as $\gamma_\tau(\sigma) = \tau \sigma \tau^{-1}$, where $\tau, \sigma \in S_n$. We say that σ_1 is *conjugate* to σ_2 if there exists $\tau \in S_n$ such that $\sigma_1 = \tau \sigma_2 \tau^{-1}$.

Lemma 7.13. *Let $\tau = (ij)$ be a transposition. Given a permutation $\sigma \in S_n$, we obtain a new permutation σ' by switching the position of i, j in σ, and this permutation σ' is exactly $\tau \sigma \tau^{-1}$.*

Proof. We only given an example. Let $\sigma = (12345)$ and $\tau = (14)$. Then, it is easy to check

$$(14)(12345)(14) = (42315).$$

□

Theorem 7.14. *Let σ_1 and σ_2 be two permutations in S_n. They are of the same type if and only if σ_1 is conjugate to σ_2.*

Proof. We first prove that if σ_1 and σ_2 are conjugate, then they are of the same type. Suppose that $\sigma_2 = \tau\sigma_1\tau^{-1}$, where $\tau \in S_n$. By Theorem 7.11, τ can be written as a product of transpositions. Then, it is enough to show that $\tau\sigma_1\tau$ is of the same type as σ_1, where $\tau = (ij)$ is a transposition. Note that τ only switches the position of i, j in σ_1 by Lemma 7.13, and does not change the type of σ_1. Thus, σ_1 and $\tau\sigma_1\tau^{-1}$ are of the same type.

For the other direction, suppose that σ_1 and σ_2 are of the same type. By Theorem 7.10, we write

$$\sigma_1 = \alpha_1 \dots \alpha_l, \quad \sigma_2 = \beta_1 \dots \beta_l$$

as products of disjoint cycles such that α_i and β_i are r_i-cycles. We first find a permutation τ_1 such that $\alpha_1 = \tau_1\beta_1\tau_1^{-1}$. to be added □

Example 7.15. Given a transposition $(ij) \in S_n$, we calculate the multiplication $(ij)\alpha$ in this example, where $\alpha \in S_n$ is an arbitrary element. By Theorem 7.10, we write α as a product of disjoint r-cycles $\alpha_1 \dots \alpha_l$. Since α_i are disjoint, there are two cases: (1) i, j are included in the same r-cycle; (2) i, j are included in distinct cycles.

For the first case, without loss of generality, suppose that $i, j \in \alpha_1$. In this case, it is enough to calculate

$$(ij)\alpha_1 = (ij)(i\dots j\dots).$$

By calculation, we have

$$(ij)\alpha_1 = (i\dots)(j\dots).$$

We find that under the multiplication, the cycle α_1 breaks into two cycles $\beta_1 = (i\dots)$ and $\beta = (j\dots)$. For example,

$$(23) \cdot (12435) = (24)(351).$$

In conclusion, we have

$$(ij)\alpha = (ij)\alpha_1 \dots \alpha_l = \beta_1\beta_2\alpha_2 \dots \alpha_l.$$

For the second case, suppose that $i \in \alpha_1 = (i\dots)$ and $j \in \alpha_2 = (j\dots)$. Then,

$$(ij)\alpha_1\alpha_2 = (ij)(i\dots)(j\dots) = (i\dots j\dots).$$

Under the multiplication, these two cycles α_1 and α_2 are connected and become one cycle $\beta = (i\dots j\dots)$. For example,

$$(23) \cdot (24)(135) = (24351).$$

In conclusion, we have

$$(ij)\alpha = (ij)\alpha_1\alpha_2 \dots \alpha_l = \beta\alpha_3 \dots \alpha_l.$$

In the first case, the transposition (ij) is called a *cut operator*, while is called a *join operator* in the second case. In the literature, this operation by transpositions is called the *cut-and-join operator*.

Homework (2 pts)**.** Discuss the multiplication by 3-cycles with same idea above. How about r-cycles (extra 2 pts)?

Theorem 7.16. *Any finite group G is a subgroup of some permutation group.*

8. Open Intervals and Closed Intervals

Let S be a set, and let X, Y be two subsets of S. For sets, there are three operations: union, intersection and complement.

$$S\backslash(X \cap Y) = S\backslash X \cup S\backslash Y, \quad S\backslash(X \cup Y) = S\backslash X \cap S\backslash Y.$$

Definition 8.1. Let S be a set in $\mathbb{R}$. Let e be a point (number) in $\mathbb{R}$. We say that e is an *interior point* of S if there exists $\epsilon > 0$ such that $(e-\epsilon, e+\epsilon) \subseteq S$. We say that e is an *exterior point* of S, if there exists $\epsilon > 0$ such that $(e-\epsilon, e+\epsilon) \nsubseteq S$. e is called an *endpoint* of S if for any $\epsilon > 0$, we have $(e-\epsilon, e+\epsilon) \nsubseteq S$ and $(e-\epsilon, e+\epsilon) \cap S \neq \emptyset$.

Note that given a set S, any point in $\mathbb{R}$ lies in one of the above three cases.

Definition 8.2. An *open interval* is a set of $\mathbb{R}$ which does not include its endpoints. A *closed interval* is defined as the complement of an open interval.

It is easy to see that any point in an open interval is an interior point. Now let a, b be two real numbers. The interval (a, b) is open, while $[a, b]$ is closed. Since we work on real numbers, we use the notation $(-\infty, \infty) = \mathbb{R}$ for the interval including all real numbers. Furthermore, $(-\infty, a)$ is defined as open and $(-\infty, a]$ is defined as closed. The interval $(-\infty, \infty)$ and the empty interval is regarded as both open and closed.

Lemma 8.3. *Let $\{U_i\}_{i\in I}$ be a set of open intervals, where I is the index set (probably not a finite set). Then, their union $\bigcup_{i\in I} U_i$ is an open interval. Thus, the intersection of closed intervals is also closed.*

Proof. Let $x \in \bigcup_{i\in I} U_i$. Then, it is included in some U_i. Since U_i is open, x is an interior point in it. There exists $\epsilon > 0$ such that $(e-\epsilon, e+\epsilon) \subseteq U_i \subseteq \bigcup_{i\in I} U_i$. Therefore, any point in $\bigcup_{i\in I} U_i$ is an interior point, which means that it is an open interval. □

Lemma 8.4. *Let F_1 and F_2 be closed intervals. Then, $F_1 \cup F_2$ is a closed interval. Thus, the intersection of finitely many open intervals is still open.*

Proof. It is enough to show that $\mathbb{R}\backslash(F_1 \cup F_2)$ is an open interval. Note that

$$\mathbb{R}\backslash(F_1 \cup F_2) = \mathbb{R}\backslash F_1 \cap \mathbb{R}\backslash F_2.$$

Since $\mathbb{R}\backslash F_1$ and $\mathbb{R}\backslash F_2$ are open, then $\mathbb{R}\backslash(F_1 \cup F_2)$ is open. □

Example 8.5. Let $U_i = (a - \frac{1}{i}, b + \frac{1}{i})$. Then,

$$\bigcap_{i=1}^{n} U_i = [a, b].$$

Here is a proof of the equality. Since

$$U_i = (a - \frac{1}{i}, b + \frac{1}{i}) \supseteq [a, b],$$

then the intersection $\bigcap_{i=1}^{n} U_i$ also includes $[a, b]$. Now suppose that there exists x such that $x \in \bigcap_{i=1}^{n} U_i$ and $x \notin [a, b]$. Without loss of generality, we suppose that $x < a$. Set $\epsilon = |x - a|$. Then, $x \notin U_i$ for $i > \frac{1}{\epsilon}$. This is a contradiction. Therefore, $\bigcap_{i=1}^{n} U_i = [a, b]$. This example tells us that the intersection of infinitely many open intervals may be a closed interval.

Let $F_i = [a + \frac{1}{i}, b - \frac{1}{i}]$. Then,

$$\bigcup_{i=1}^{n} F_i = (a, b).$$

The proof is similar to the above example. Notice that $F_i = [a + \frac{1}{i}, b - \frac{1}{i}] \subseteq (a, b)$, then the union $\bigcup_{i=1}^{n} F_i$ is included in (a, b). Now suppose that there exists x such that $x \in (a, b)$ and $x \notin \bigcup_{i=1}^{n} F_i$.

Suppose that $|x-a| < |x-b|$. Let $\epsilon = |x-a|$. Then, $x \in F_i$ for $i > \frac{1}{\epsilon}$. Thus, $x \in \bigcup_{i=1}^{n} F_i$, and this is a contradiction. Therefore, we have $\bigcup_{i=1}^{n} F_i = (a,b)$. This example shows that the union of infinitely many closed intervals may be an open interval.

Probably explain logic notations: and, or, not, all, exists.

Definition 8.6 (Limits)**.** Given a function $f : \mathbb{R} \to \mathbb{R}$, we say that when x approaches x_0, the *limit* of $f(x)$ is L, if it satisfies the following conditions. For any $\epsilon > 0$, there exists $\delta > 0$, whenever $|x - x_0| < \delta$, we have $|f(x) - L| < \epsilon$. We use that notation $\lim_{x \to x_0} f(x) = L$ for the limit.

Definition 8.7 (Continuity)**.** Let U be an open subset of $\mathbb{R}$, and let $f : U \to \mathbb{R}$ be a function. We say that $f(x)$ is *continuous* if it satisfies the following conditions

(1) for any $u \in U$, when x approaches u, the limit of $f(x)$ exists;
(2) for any $u \in U$, the value $f(u)$ exists;
(3) $\lim_{x \to u} f(x) = f(u)$.

Lemma 8.8. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, for any open interval $V \subseteq \mathbb{R}$, we have $f^{-1}(V)$ is an open interval.*

Proof. It is equivalent to prove that for each point $y \in V$, $f^{-1}(y)$ is included in an open interval $I_y \subseteq f^{-1}(V)$. Let y be a point in V. Since V is open, we choose ϵ such that $(y-\epsilon, y+\epsilon) \subseteq V$. Let $x := f^{-1}(y)$ be the preimage of y. By the continuity of f, there exists δ such that $f((x-\delta, x+\delta)) \subseteq (y-\epsilon, y+\epsilon)$. Thus, $(x-\delta, x+\delta) \subseteq f^{-1}((y-\epsilon, y+\epsilon)) \subseteq f^{-1}(V)$. This finishes the proof. □

9. Topology

Add ideas of subsets, probably in the cardinality

9.1. Basic Definition.

Definition 9.1 (Topology)**.** Let X be a set. A *topology* $\mathcal{T}$ on X is a collection of subsets of X satisfying the following properties:

(1) $\emptyset$ and X are in $\mathcal{T}$;
(2) let $U_i \in \mathcal{T}$, then their union $\cup_{i \in I} U_i$ is still in $\mathcal{T}$, where I could be an infinite index set;
(3) let $U_i \in \mathcal{T}$, then their intersection $\cap_{i \in I} U_i$ is still in $\mathcal{T}$, where I is a finite index set.

A set X equipped with a topology $\mathcal{T}$ is called a *topological space*. We use either $(X, \mathcal{T})$ or X for the topological space. An element U in $\mathcal{T}$ is called an *open set* of X (or U is *open* in X). On the other hand, its complement $X \backslash U$ is called an *closed set* of X (or *closed* in X).

Example 9.2. Let $X = \{1,2,3\}$. Give some examples of topology on X. discrete topology

Example 9.3 (Co-finite Topology)**.** Let X be a set. Define $\mathcal{T}_f$ as follows

$$\mathcal{T}_f = \{U \mid U \subseteq X, \text{ and } U = \emptyset \text{ or } X \backslash U \text{ is a finite set.}\}$$

$\mathcal{T}_f$ is a collection of subsets of X, and furthermore, it is a topology on X. Clearly, the first condition is satisfied. Now given a collection of elements $U_i \in \mathcal{T}_f$, we have

$$X \backslash (\cup_{i \in I} U_i) \subseteq X \backslash U_i.$$

Since $X \backslash U_i$ is a finite set, $X \backslash (\cup_{i \in I} U_i)$ is also a finite set. Thus, $\cup_{i \in I} U_i \in \mathcal{T}_f$. Finally, for the third condition, given a finite collections of elements U_i in $\mathcal{T}_f$, we have

$$X \backslash (\cap_{i \in I} U_i) \subseteq \cup_{i \in I} X \backslash U_i.$$

Since $X \backslash U_i$ is a finite set and the index set I is finite, then $X \backslash (\cap_{i \in I} U_i)$ is also a finite set. Thus, $\cap_{i \in I} U_i \in \mathcal{T}_f$. In conclusion, $\mathcal{T}_f$ is a topology on X, and this topology is called *co-finite topology* or *finite complement topology*.

Definition 9.4 (Basis)**.** Let X be a set. Let $\mathcal{B}$ be a collection of subsets of X such that

(1) for each $x \in X$, there is at least one element $B \in \mathcal{B}$ containing x;
(2) if $x \in B_1 \cap B_2$, where $B_1, B_2 \in \mathcal{B}$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

The *topology* $\mathcal{T}$ *generated by* $\mathcal{B}$ is defined as follows: a subset U of X is said to be open, if for each $x \in U$, there exists an element $B \in \mathcal{B}$ such that $x \in B \subseteq U$. We say that $\mathcal{B}$ is a basis of $\mathcal{T}$, and elements in $\mathcal{B}$ are called *basis elements*.

Generally speaking, it is very hard to give a precise description of all elements included in a topology. This is one of the reasons why we introduce the concept of *basis*. We only have to describe some key elements and use them to generate a topology.

Example 9.5. Let $\mathcal{B} = \{(a, b), a, b \in \mathbb{R}\}$ be the collection of all open intervals in $\mathbb{R}$. The topology generated by $\mathcal{B}$ is called the *standard topology*.

Definition 9.6 (Product Topology)**.** Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be two topological spaces. The *product topology* on $X \times Y$ is generated by the basis

$$\mathcal{B} := \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_Y\}.$$

Example 9.7. Example that we have to use it generates the topology

Definition 9.8 (Subspace Topology)**.** Let $(X, \mathcal{T})$ be a topological space. Let $Y \subseteq X$ be a subset. The collection $\mathcal{T}_Y := \{U \cap Y \mid U \in \mathcal{T}\}$ is a topology on Y. This topology $\mathcal{T}_Y$ is called the *subspace topology*.

Definition 9.9. Let X be a set. Suppose that $\mathcal{T}_1$ and $\mathcal{T}_2$ are two topologies on X. If $\mathcal{T}_1 \subseteq \mathcal{T}_2$, we say that $\mathcal{T}_1$ is *coarser* than $\mathcal{T}_2$ or $\mathcal{T}_2$ is *finer* than $\mathcal{T}_1$.

Example 9.10. In this subsection, we introduce two topologies on $\mathbb{R}$. One is the stanrdard topology (denote by $\mathcal{T}_1$), which is given by open intervals, and the other one is co-finite topology (denote by $\mathcal{T}_2$), which is given by complement of finite set. It is easy to check that the complement of any finite set is an open interval. Therefore, $\mathcal{T}_2 \subseteq \mathcal{T}_1$. *Thus, the co-finite topology coarser than the standard topology or the standard topology is finer than the co-finite topology.*

Lemma 9.11. *Let $\mathcal{B}_i$ be a base for the topology $\mathcal{T}_i$ on X for $i = 1, 2$. Then, the following statements are equivalent:*

- *$\mathcal{T}_1$ is finer than $\mathcal{T}_2$.*
- *For each $x \in X$ and each element $B \in \mathcal{B}_2$ containing x, there is an element $B' \in \mathcal{B}_1$ such that $x \in B' \subseteq B$.*

Proof. Given $x \in B \subseteq X$ and $B \in \mathcal{B}_2$, since $B \in \mathcal{T}_2$ and $\mathcal{T}_2 \subseteq \mathcal{T}_1$ by definition, we have $B \in \mathcal{T}_1$. Since $\mathcal{T}_1$ is generated by $\mathcal{B}_1$, there is an element $B' \in \mathcal{B}_1$ such that $x \in B' \subseteq B$.

For the other direction, we have to show that given an element $T \subseteq \mathcal{T}_2$, T is also included in $\mathcal{T}_1$. Since $\mathcal{B}_2$ generates $\mathcal{T}_2$, for each $x \in T$, there is an element $B_x \in \mathcal{B}_2$ containing x. By the second condition, we can find $B'_x \in \mathcal{B}_1$ such that $x \in B'_x \subseteq B_x$. Thus, T is also an open subset in $\mathcal{T}_1$. This finishes the proof. □

9.2. **Hausdorff Space.**

Definition 9.12. Let X be a topological space, and let A be a subset of X. The *interior* of A is defined as the union of all open sets contained in A. Denote it by $\mathrm{Int}A$. The *closure* of A is defined as the intersection of all closed sets containing A. Denote it by $\overline{A}$.

Given an arbitrary subset $A \subseteq X$, we have

$$\mathrm{Int}A \subseteq A \subseteq \overline{A}.$$

Lemma 9.13. *Let A be a subset of a topological space X. A point $x \in \overline{A}$ if and only if every open set U containing x intersects A.*

Proof. Suppose that $x \in \overline{A}$. If $x \in A$, clearly $U \cap A \neq \emptyset$ for any U containing x. If $x \notin A$, suppose that there exists an open set U such that $x \in U$ and $U \cap A = \emptyset$. Consider the closed set $Z := X \backslash U$. Clearly, $X \subseteq Z \neq \emptyset$. By definition of $\bar{A}$, we have $x \in Z$. However, this contradicts with the fact that $x \in U$.

For the other direction, suppose that $x \notin \bar{A}$. Then, there exists a closed set Z such that $A \subseteq Z$ and $x \notin Z$. Consider the open set $U := X \backslash Z$. Clearly, $x \in U$. By the condition, we know that $U \cap A \neq \emptyset$, which contracts with the fact $A \subseteq Z$. This finishes the proof. □

Definition 9.14 (Limit Points)**.** Let A be a subset of a topological space X. Let $x \in X$ be a point. We say that x is a *limit point* of A if $x \in \overline{A - \{x\}}$.

Let A be a subset of a topological space X. Let A' be the set of all limit points of A. Then, $\bar{A} = A \cup A'$.

Lemma 9.15. *A subset contains all limit points is closed.*

Proof. □

Definition 9.16 (Hausdorff Topology)**.** A topological space X is a *Hausdorff space* if for each pair x_1, x_2 of distinct points of X, there exist disjoint neighborhoods U_1 of x_1 and U_2 of x_2.

Lemma 9.17. *Every finite point set in a Hausdorff space is closed.*

Proof. We only have to show that a single point set $\{x\}$ is closed. Choose any other point y. By definition of Hausdorff space, we can find $x \in U$ and $y \in V$ such that U, V are open and $U \cap V = \emptyset$. Since $x \notin V$, then $y \notin \overline{\{x\}}$. Thus, we have $\overline{\{x\}} = \{x\}$ and this implies that $\{x\}$ is closed. □

Definition 9.18. We say that a sequence $\{x_i\}_{i \in \mathbb{Z}}$ of points *converges* to x if for any open set U containing x, there exists $N \geq 0$ such that $x_n \in U$ for all $n > N$.

Theorem 9.19. *Let X be a Hausdorff space. A sequence of points in X converges to at most one point of X.*

Proof. Suppose that there exists a sequence of points $\{x_n\}$ converges to two distinct points x and y. By definition of Hausdorff space, we can find open set U and V such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$. Since $\{x_n\}$ converges to both x and y, we can find $N \geq 0$ such that $x_n \in U$ and $x_n \in V$ for $n > N$. This contradicts with the fact that $U \cap V = \emptyset$. □

Proposition 9.20. *Let X be a topological space. Denote by $\Delta : X \to X \times X$ the diagonal map. Prove that X is a Hausdorff space if and only if Δ is closed.*

Corollary 9.21. *Let X be a set with infinitely many elements. The cofinite topology T_f on X is not Hausdorff.*

Definition 9.22 (Continuous Functions)**.** Let X and Y be topological spaces. A function $f : X \to Y$ is *continuous* if for any open subset V of Y, the preimage $f^{-1}(V)$ is open in X.

Lemma 9.23. *Let X and Y be topological spaces. Let $f : X \to Y$ be a function. The following statements are equivalent:*

(1) f is continuous;
(2) for any subset $A \subseteq X$, we have $f(\bar{A}) \subseteq \overline{f(A)}$;
(3) for any closed set B of Y, we have $f^{-1}(B)$ is closed in X;
(4) for any $x \in X$ and any open subset V containing $f(x)$, there is an open subset U of x in X such that $f(U) \subseteq V$.

9.3. **Metric Topology.**

Definition 9.24. Let X be a set. A *metric* on X is a function

$$d : X \times X \to \mathbb{R}$$

such that

(1) $d(x, y) \geq 0$ for $x, y \in X$;
(2) $d(x, y) = 0$ if and only if $x = y$;
(3) $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$.

Let $x, y \in X$. The number $d(x, y)$ is called the *distance* between x and y.

Let d be a metric on X. Given a number $\epsilon > 0$, define the *ϵ-ball centered at x* as

$$B_d(x, \epsilon) = \{y \mid d(x, y) < \epsilon\}.$$

Lemma 9.25. *Let $\mathcal{B} = \{B_d(x, \epsilon) \mid x \in X, \epsilon > 0\}$. Then, $\mathcal{B}$ is a basis of X*

Proof. Given any $\epsilon > 0$, clearly, $x \in B_d(x, \epsilon)$ for any $x \in X$. Now we consider the second condition. Given $x \in B_d(y, \epsilon_y) \cap B_d(z, \epsilon_z)$, consider the following inequalities

$$d(x, y) + \epsilon_x < \epsilon_y, \quad d(x, z) + \epsilon_x < \epsilon_z.$$

Let $\epsilon_x := \min\{\epsilon_y - d(x, y), \epsilon_z - d(x, z)\}$. Then, we have

$$B_d(x, \epsilon_x) \subseteq B_d(y, \epsilon_y) \cap B_d(z, \epsilon_z).$$

□

Definition 9.26. Let d be a metric on X. The *metric topology* on X (induced by d) is defined as the topology generated by the collection of all ϵ-balls. Furthermore, let X be a topological space. X is *metrizable* if there exists a metric d that induces this topology.

The standard topology on $\mathbb{R}^n$ is a metric topology.

Lemma 9.27. *Let d_1 and d_2 be two metrics on X. Denote by $\mathcal{T}_1$ and $\mathcal{T}_2$ the corresponding topology. Then, $\mathcal{T}_1$ is finer than $\mathcal{T}_2$ if and only if for each $x \in X$ and every $\epsilon > 0$, there exists $\delta > 0$ such that*

$$B_{d_1}(x, \delta) \subseteq B_d(x, \epsilon).$$

Proof. □

Theorem 9.28. *The topologies on $\mathbb{R}^n$ induced by the euclidean metric d and the square metric ρ are the same as the product topology on $\mathbb{R}^n$.*

Proof. □

10. Basic idea of Algebraic Geometry

We first consider an example: curves. Curves are one of the most basic objects in geometry. To give a precise description of curves, we prefer to establish a coordinate system and regard the curve as the solution of some equations. For example, $Y^2 - X^3 = 0$. This gives an algebraic way to describe geometric objects, and this is also one of the approaches we already know.

Based on this concept, we first consider the examples $\mathbb{C}^n$. The vector space $\mathbb{C}^n$ is called an *affine space* (over $\mathbb{C}$). The coefficient can be an arbitrary field k. As an affine space, we prefer to use the notation $\mathbb{A}^n$. An element in $\mathbb{A}^n$ is called a *point*. A polynomial in $\mathbb{C}[x_1, \ldots, x_n]$ can be regarded as a function

$$f(x_1, \ldots, x_n) : \mathbb{A}^n \to \mathbb{A}^1.$$

The polynomial ring is called *the ring of regular functions on* $\mathbb{A}^n$, and is denoted by $\mathbb{C}[\mathbb{A}^n]$. An element in $k[\mathbb{A}^n]$ is called a *regular function*.

Let $f \in k[\mathbb{A}^n]$ be a regular function. Denote by $Z(f)$ the zero set of the regular function f. Furthermore, given a set of regular functions $T \subseteq \mathbb{C}[\mathbb{A}^n]$, define $Z(T)$ to be the set of common zeros of elements of T. Let $I = (f)$ be the ideal generated by a single regular function f. Clearly, $Z(f) = Z(I)$.

11. Symmetric Polynomials