Grad Student Prep Course: Discrete Mathematics

Prepared by Nicholas Camacho
University of Iowa

Contents

Introduction

This 5-day prep course will, of course, not be able to handle the material presented in a complete fashion. In many places, some concepts are used without being previously introduced. Instead, the purpose of this brief course is to provide a basic overview of material to help prepare first-year graduate students with their qualifying exam courses. These notes contain a baseline for what the expectations are from the student's undergraduate curriculum.

Many of the stated results, definitions, theorems, examples, etcetera, from these notes come directly from [**1**],[**2**],[**3**], [**4**], [**5**], and [**6**].

If you find any typos, or if you have general suggestions for improving these notes, please don't hesitate to let me know.

Nicholas Camacho
August 2020
nicholas-camacho@uiowa.edu

DAY 1

Sets, Functions, and Proof-Writing

1.1. Sets

1.1.1. Basic Definitions and Notation.

Definition 1.1.1. A *set* is a collection of things called *elements* whose membership is unambiguous.

- To denote membership of an element x in a set A, we write
$$x \in A.$$
- If an object x is not a member of a set A, we write
$$x \notin A.$$
- There is a set with no elements, called the *empty set*, and is denoted $\varnothing$.
- Given two sets A and B, we write

  $A = B$ if A and B have the same elements.

  $A \neq B$ if A and B do not share at least one element in common.

  $A \subset B$ if all the elements of A are also members of B.

  $A \subseteq B$ if $A \subset B$ or $A = B$.

  $A \subsetneq B$ if $A \subset B$ and $A \neq B$.

  $A \supset B$ if all the elements of B are also members of A.

  $A \supseteq B$ if $A \supset B$ or $A = B$.

  $A \supsetneq B$ if $A \supset B$ and $A \neq B$.

Example 1.1.2.

(a) If a set A has only a few elements, say a, b, and c, one can simple write a statement
$$A = \{a, b, c\}.$$

(b) Usually, the way to specify a set A is to specify a property that each member of A has. If A consists of all real numbers which are even integers, we can write
$$A = \{x \in \mathbb{R} : x \text{ is an even integer}\}.$$
The colon ":" can also be a vertical bar "|" and stands for the words "such that." So, the above can be read "A is the set of elements x in $\mathbb{R}$ such that x is an even integer." Notice that we can immediately recognize that A above is a subset of $\mathbb{R}$, that is $A \subseteq \mathbb{R}$.

1.1.2. Set Operations. Given two sets A and B, we have the operations

(Union of sets) $A \cup B = \{x : x \in A \text{ or } x \in B\}$,

(Intersection of sets) $A \cap B = \{x : x \in A \text{ and } x \in B\}$,

(Difference of sets) $A - B = \{x : x \in A \text{ and } x \notin B\}$,

(Cartesian Product) $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.

Remark 1.1.3. If A is a subset of some "universal set" U in which we are working, we may also write A^c to denote $U - A$. In this case, if B is also a subset of U, then $A - B$ is also called the *complement of B relative to A*.

Unions, intersections, and products can, using the Axiom of Choice[1], be extended to an arbitrary number of sets: If J is a set of an arbitrary number of elements (finite or (un)countably infinite), and if for each $\alpha \in J$ there is a set A_α, we have[2]

$$\bigcup_{\alpha \in J} A_\alpha = \{x : \exists \alpha \in J, x \in A_\alpha\},$$
$$\bigcap_{\alpha \in J} A_\alpha = \{x : x \in A_\alpha \forall \alpha \in J\}, \text{ and}$$
$$\prod_{\alpha \in J} A_\alpha = \{(x_\alpha)_{\alpha \in J} : x_\alpha \in A_\alpha \forall \alpha \in J\}.$$

Using these operations, some important set-theoretic rules are the following:

(a) The Distributive Laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ and}$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

(b) DeMorgan's Laws:

Complement of union is intersection of complements:

$$A - (B \cup C) = (A - B) \cap (A - C),$$

Complement of intersection is union of complements:

$$A - (B \cap C) = (A - B) \cup (A - C).$$

Moreover, DeMorgan's Laws also hold for an arbitrary number of sets: If A is a set and $\{A_\alpha\}_{\alpha \in J}$ is a collection of sets, then:

Complement of union is intersection of complements:

$$A - \bigcup_{\alpha \in J} A_\alpha = \bigcap_{\alpha \in J} (A - A_\alpha)$$

Complement of intersection is union of complements:

$$A - \bigcap_{\alpha \in J} A_\alpha = \bigcup_{\alpha \in J} (A - A_\alpha)$$

[1]See section 2.4 of Day 2.

[2]Instead of relying on an indexing set J, we could also write this by letting $\mathcal{J}$ be a set containing an arbitrary number of sets, and index using the elements of $\mathcal{J}$ themselves: $\bigcup_{A \in \mathcal{J}} A$.

Example 1.1.4 (The Finite Complement Topology)**.**

Definition. A *topology* on a set X is a collection of subsets $\mathcal{T}$ of X with the following properties:

(a) $\varnothing$ and X are in $\mathcal{T}$.
(b) the union of the elements of any subcollection of elements of $\mathcal{T}$ is also in $\mathcal{T}$.
(c) The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

Let X be a set, and let $\mathcal{T}$ be the collection of all subsets U of X such that $X - U$ is either finite or all of X. Prove that $\mathcal{T}$ is a topology on X.

Proof.

□

1.2. Functions

1.2.1. Basic Definitions and Notation.

Definition 1.2.1.

(a) A *function* is a subset f of the Cartesian product $A \times B$ of two sets A and B, such that each element of A appears as the first coordinate of at most one ordered pair. The set A is called the *domain* of the function f, and B is called the *range* of f. The subset of B which consists of all the second coordinates of elements of f is called the *image* of f:

$$\text{image of } f = \{b \in B : \exists a \in A, (a, b) \in f\}.$$

If f is a function with domain A and range B, we express this fact by writing

$$f : A \longrightarrow B,$$

which reads "f is a function from A to B" or "f is a mapping from A into B." If $a \in A$, we write $f(a)$ to denote the unique element $b \in B$ such that $(a, b) \in f$, and $f(a)$ is referred to as the *image of a under f*.

(b) A function $f : A \to B$ is said to be *injective* if for each pair of distinct points of A, their images under f are distinct. In other words, f is injective if for every $a, a' \in A$, we have

$$a \neq a' \implies f(a) \neq f(a'),$$

or equivalently,

$$f(a) = f(a') \implies a = a'.$$

(c) A function $f : A \to B$ is said to be *surjective* if for each element $b \in B$, there is at least one element $a \in A$ whose image under f equals b, i.e. $f(a) = b$.

(d) A function that is both injective and surjective is called *bijective*. If $f : A \to B$ is a bijective function, there exists a unique function $f^{-1} : B \to A$ defined by the rule that for each $b \in B$, $f^{-1}(b)$ is the unique element $a \in A$ such that $f(a) = b$. Moreover, if f is bijective, then f^{-1} is bijective also.

(e) If $f : A \to B$ and if A_0 is a subset of A, we define the *restriction of f to A_0* to be the function mapping from A_0 into B defined by

$$\{(a, f(a)) : a \in A_0\}.$$

We denote the restriction of f to A_0 by $f|_{A_0}$, which is read "f restricted to A_0."

(f) If $f : A \to B$ and $A_0 \subseteq A$, we denote $f(A_0)$ the set of all images of points of A_0 under the function f; this set is called the *image of A_0 under f*.

$$f(A_0) = \{b \in B : \exists a \in A_0, b = f(a)\}.$$

If $B_0 \subseteq B$, we denote by $f^{-1}(B_0)$ the set of all elements of A whose images under f lie in B_0, and $f^{-1}(B_0)$ is called the *preimage of B_0 under f*.

$$f^{-1}(B_0) = \{a \in A : f(a) \in B_0\}.$$

Note that if $f : A \to B$ is bijective, we now have two meanings for the notation $f^{-1}(B_0)$. However, these two meanings yield the exact same subset of A.

(g) Given functions $f : A \to B$ and $g : B \to C$, we define the composite $g \circ f$ of f and g as the function $g \circ f : A \to C$ defined by the equation

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

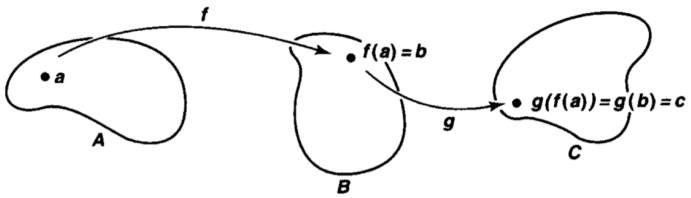


FIGURE 1. From James Munkres' *Topology* 2nd Ed., page 18

Example 1.2.2 (Properties of f and f^{-1})**.**

(1) Let $f : A \to B$, and let $A_0, A_1 \subseteq A$ and $B_0, B_1 \subseteq B$. When we apply f^{-1} to inclusions, unions, intersections, and differences, the operations are preserved. In other words, we have:
   (a) $B_0 \subseteq B \implies f^{-1}(B_0) \subseteq f^{-1}(B)$,
   (b) $f^{-1}(B_0 \cup B_1) = f^{-1}(B_0) \cup f^{-1}(B_1)$,
   (c) $f^{-1}(B_0 \cap B_1) = f^{-1}(B_0) \cap f^{-1}(B_1)$, and
   (d) $f^{-1}(B_0 - B_1) = f^{-1}(B_0) - f^{-1}(B_1)$.

   When we apply f to these operations, only inclusions and unions are preserved:
   (a) $A_0 \subseteq A \implies f(A_0) \subseteq f(A)$, and
   (b) $f(A_0 \cup A_1) = f(A_0) \cup f(A_1)$.

   Prove that:
   (c) $f(A_0 \cap A_1) \subseteq f(A_0) \cap f(A_1)$; show that equality holds if f is injective.
   (d) $f(A_0 - A_1) \supseteq f(A_0) - f(A_1)$; show that equality holds if f is injective.

(2) Prove the following equivalence, which gives another definition of bijectivity:

   A function $f : A \to B$ is bijective if and only if there exists a function $g : B \to A$ such that for all $a \in A$ and for all $b \in B$

$$g(f(a)) = a \quad \text{and} \quad f(g(b)) = b.$$

(3) Using the same notation as above, show that
   (a) $A_0 \subseteq f^{-1}(f(A_0))$, and that equality holds if f is injective.
   (b) $f(f^{-1}(B_0)) \subseteq B_0$, and that equality holds if f is surjective.

1.3. Other structures on sets

1.3.1. Equivalence Relations and Partitions.

Definition 1.3.1. A *relation* on a set A is a subset C of the cartesian product $A \times A$. We use the notation xCy to mean the same thing as $(x, y) \in C \subseteq A \times A$.

A function $f : A \to A$ (from a set to itself) is a relation on A, but note that such a function is a particular kind of relation: every element of A appears as the first coordinate of the subset $f \subseteq A \times A$ exactly once.

Example 1.3.2. Let P denote the set of all people in the world, and define $D \subseteq P \times P$ by the equation

$$D = \{(x, y) : x \text{ is a descendant of } y\}.$$

Consider another relation $S \subseteq P \times P$ defined by

$$S = \{(x, y) : \text{ the parents of } x \text{ are the parents of } y\}.$$

Notice that the relation S is "symmetric" in the sense that if the parents of x are the parents of y, then the parents of y are the parents of x. However, the relation D does not have this property: if x is a descendant of y, then y is not a descendant of x.

Definition 1.3.3. An *equivalence relation* on a set A is a relation $E \subseteq A \times A$ with the following properties:

(a) (Reflexivity) For every $x \in A$: xEx.
(b) (Symmetry) For every $x, y \in A$: If xEy, then yEx.
(c) (Transitivity) For every $x, y, z \in A$: If xEy and yEz, then xEz.

Notation. Another common way to denote a relation (and in particular, an equivalence relation) is the "tilde" symbol: $\sim$. If E is an equivalence relation and xEy, we will instead write $x \sim y$.

Definition 1.3.4. Given an equivalence relation on a set A and an element $x \in A$, we define a certain subset $C \subseteq A$, called the *equivalence class* determined by x, by the equation

$$C := [x] := \{y \in A : y \sim x\}.$$

Lemma 1.3.5. *Two equivalence classes $[x]$ and $[x']$ are either disjoint or equal.*

PROOF.

□

Definition 1.3.6. A partition of a set A is a collection of pairwise disjoint nonempty subsets of A whose union is all of A.

Using the Lemma, prove: Given a partition of A there is exactly one equivalence relation on A from which it is derived.

Example 1.3.7.

(1) (Projective Space) Define two points in the plane to be equivalent if they lie on the same line through the origin. This defines an equivalence relation on the plane, and the equivalence classes form a space called *projective space*, and is denoted by $\mathbb{P}^1$. Moreover, an equivalence class for a point $(x, y) \in \mathbb{R}^2$ is denoted by $(x : y)$.

We can extend this to $\mathbb{R}^{n+1}$: Two points $a = (a_0, \ldots, a_n)$ and $b = (b_0, \ldots, b)$ in $\mathbb{R}^n - \{0\}$ are on the same line through the origin if and only if there exists $\lambda \in \mathbb{R} - \{0\}$ such that

$$\lambda(a_0, \ldots, a_n) = (\lambda a_0, \ldots, \lambda a_n) = (b_0, \ldots, b_n).$$

This defines an equivalence relation on $\mathbb{R}^{n+1} - \{0\}$, and the set of equivalence classes is called *projective n-space over* $\mathbb{R}$, denoted $\mathbb{P}^n(\mathbb{R})$ or $\mathbb{RP}^n$.

(2) Define two points in the plane to equivalent if they have the same y-coordinate. This defines an equivalence relation, whose equivalence classes are the set of horizontal lines.

1.3.2. Orders and Partial Orders.

Definition 1.3.8.

(1) A *partial order* on a set A is relation $E \subseteq A \times A$ with the following properties:
   (a) (Reflexivity) For every $x \in A$: xEx.
   (b) (Antisymmetry) For every $x, y \in A$: If xEy and yEx, then $x = y$.
   (c) (Transitivity) For every $x, y, z \in A$: If xEy and yEz, then xEz.

   A set together with a partial order is called a *poset*.

(2) A *total order* on a set A is a relation $E \subseteq A \times A$ that is a partial order with the additional property of Comparability (or the Trichotomy Law):
   (d) For all $x, y \in A$, either xEy or yEx.

Example 1.3.9.

(a) The $\leq$ on the set of integers is a partial order, making the pair $(\mathbb{Z}, \leq)$ a poset. Moreover, $(\mathbb{Z}, \leq)$ is also a total order.

(b) Given any set A, the *power set* of A is the set of all subsets of A, including both A and $\varnothing$, and is denoted $\mathcal{P}(A)$. The subset relation "$\subseteq$" on $\mathcal{P}(A)$ makes the pair $(\mathcal{P}(A), \subseteq)$ a poset.

(c) The strict "less than" relation $<$ on $\mathbb{Z}$ is not a partial order, since it is not reflexive. However, this kind of relation does have a name:

Definition 1.3.10. A *linear order* on a set A is relation $E \subseteq A \times A$ with the following properties:

(a) (Nonreflexivity) For every $x \in A$, the relation xEx does not hold.
(b) (Comparability) For every $x, y \in A$ with $x \neq y$: Either $x \sim y$ or $y \sim x$.
(c) (Transitivity) For every $x, y, z \in A$: If xEy and yEz, then xEz.

Example 1.3.11 (Dictionary Ordering)**.** Suppose that A and B are sets with linear orders $<_A$ and $<_B$. Define an order relation $<$ on $A \times B$ by defining

$$a_1 \times b_1 < a_2 \times b_2$$

if $a_1 <_A a_2$, or if $a_1 = a_2$ and $b_1 <_B b_2$. Then $<$ is a linear order, and is called the dictionary order on $A \times B$.

DAY 2

Sets and Integers

2.1. Defining the Integers

We assume that there exists a set $\mathbb{R}$, called the set of *real numbers* with two operations $+$ and $\cdot$, with the usual algebraic properties (associativity, commutativity, multiplicative identity 1, additive identity 0, distributivity of $\cdot$ over $+$, etcetera) and order properties (least upper bound property, $<$ is "dense" in $\mathbb{R}$, etcetera).

Definition 2.1.1. A subset A of the real numbers is said to be *inductive* if it contains the number 1, and if for every $x \in A$, the number $x+1$ is also in A. Let $\mathcal{A}$ be the collection of all inductive subsets of $\mathbb{R}$. Then the set $\mathbb{Z}_+$ of *positive integers* is defined by the equation

$$\mathbb{Z}_+ = \bigcap_{A \in \mathcal{A}} A.$$

The set $\mathbb{Z}$ of *integers* is the set consisting of the positive integers $\mathbb{Z}_+$, the number 0, and the negatives of the elements of $\mathbb{Z}_+$.

Remark 2.1.2.

(1) It follows from this definition that the positive integers are given by

$$\mathbb{Z}_+ = \{1, 2, 3, 4, 5, \dots\}.$$

(2) There are a three properties of $\mathbb{Z}_+$ that are used quite frequently:
   (a) (Well-Ordering Property) Every nonempty subset of $\mathbb{Z}_+$ has a smallest element.
   From the well-ordering property, the following two principles can be proven:
   (b) (Strong Mathematical Induction Principle) Let A be a set of integers, and let $k \in \mathbb{Z}_+ \cup \{0\}$ be fixed. Suppose that $k \in A$ and that for each positive integer $n \geq k$, the following holds:

$$\{k, \dots, n\} \subset A \implies n+1 \in A.$$

   Then $A = \{k, k+1, k+2, \dots\}$.
   (c) (Weak Mathematical Induction Principle) Let A be a set of integers, and let $k \in \mathbb{Z}_+ \cup \{0\}$ be fixed. Suppose that $k \in A$ and that for each positive integer $n \geq k$, the following holds:

$$n \in A \implies n+1 \in A.$$

   Then $A = \{k, k+1, k+2, \dots\}$.

We usually want the set A to comprise of the set of all positive integers for which a particular statement is true, and, after having shown that $0 \in A$, we proceed to show that $A = \mathbb{Z}_+ \cup \{0\}$. Moreover, Weak and Strong Induction are actually equivalent statements.

Example 2.1.3. Consider the Fibonacci numbers $\{F_n\}_{n\in\mathbb{Z}_+}$:

$$F_1 = 1, F_2 = 1,$$
$$F_{n+1} = F_n + F_{n-1} \text{ for } n = 2, 3, 4, \ldots$$

Show that for $k \geq 1$, $\sum_{i=1}^{k} F_i = F_{k+2} - 1$.

2.2. Finite Sets and the Pigeonhole Principle

Definition 2.2.1. A set is said to be *finite* if there is a bijective correspondence of A with some section of the positive integers. That is, A is finite if it is empty or if there is a bijection

$$f : A \longrightarrow \{1, \ldots, n\}$$

for some positive integer n. In the former case, we say that A has *cardinality* 0; in the latter case, we say that A has *cardinality* n

Theorem 2.2.2 (Generalized pigeon-hole principle)**.** *If each of N items is put into exactly one of n pigeon-holes and $N > mn$, then there is one pigeon-hole with at least $m + 1$ items.*

Proof.

□

Theorem 2.2.3. *Let B be a nonempty set. Then the following are equivalent:*

(1) B is finite.
(2) There is a surjective function from a section of the positive integers onto B.
(3) There is an injective function from B into a section of the positive integers.

Example 2.2.4. Let A be a nonempty finite set that is linearly ordered.

(1) Show that A has a largest element.
(2) Show that A has the order type of a section of the positive integers.

Proof.

□

2.3. Countable and Uncountable Sets; Diagonalization Arguments

Definition 2.3.1.

(1) A set A is said to be *infinite* if it is not finite. It is said to be countably infinite if there is a bijective correspondence

$$f : A \longrightarrow \mathbb{Z}_+.$$

(2) A set is said to be *countable* if it is either finite or countably infinite. A set that is not countable is said to be *uncountable*.

Theorem 2.3.2. *Let B be a nonempty set. Then the following are equivalent:*

(1) B is countable.
(2) There is a surjective function $f : \mathbb{Z}_+ \to B$.
(3) There is an injective function $g : B \to \mathbb{Z}_+$.

Example 2.3.3.

(1) The set $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countably infinite, via the map

$$\begin{aligned} f : \mathbb{Z}_+ \times \mathbb{Z}_+ &\longrightarrow \mathbb{Z}_+ \\ (n, m) &\longmapsto 2^n 3^m. \end{aligned}$$

(2) The set $\mathbb{Q}_+$ of positive rationals is countably infinite. Define

$$\begin{aligned} g : \mathbb{Z}_+ \times \mathbb{Z}_+ &\longrightarrow \mathbb{Q}_+ \\ (n, m) &\longmapsto m/n. \end{aligned}$$

Since $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countable, there exists a surjection $f : \mathbb{Z}_+ \to \mathbb{Z}_+ \times \mathbb{Z}_+$, and the composite $g \circ f : \mathbb{Z}_+ \to \mathbb{Q}_+$ is a surjection.

Next, we move on to diagonalization arguments.

> "Diagonal proofs make use of self-reference to create objects that have certain properties, usually of a negative kind. These objects are sometimes called *self-defeating objects*" [**5**, Page 143]

Theorem 2.3.4 (Cantor's Theorem)**.** *If A is a set, the set of subsets of A, denoted $\mathcal{P}(A)$ cannot be put into one-to-one correspondence with A. In fact, there is no onto map $A \to \mathcal{P}(A)$.*

Proof.

□

Theorem 2.3.5. *The set $\mathbb{R}$ of real numbers is uncountable.*

Proof.

□

2.4. Axiom of Choice

Axiom of Choice: Give a collection $\mathcal{A}$ of disjoint nonempty sets, there exists a set C consisting of exactly one element from each element of $\mathcal{A}$; that is, a set C such that C is contained in the union of the elements of $\mathcal{A}$, and for each $A \in \mathcal{A}$, the set $C \cap A$ contains a single element.

From the Axiom of Choice, one can prove:

Lemma 2.4.1 (Existence of a choice function)**.** *Given a collection $\mathcal{B}$ of nonempty sets (not necessarily disjoint), there exists a function*

$$c : \mathcal{B} \longrightarrow \bigcup_{B \in \mathcal{B}} B$$

such that $c(B)$ is an element of B, for each $B \in \mathcal{B}$.

Theorem 2.4.2. *Let A be a set. The following statements about A are equivalent:*

(1) There exists an injective function $f : \mathbb{Z}_+ \to A$.
(2) There exists a bijection of A with a proper subset of itself.
(3) A is infinite.

PROOF. (Only proving (3) $\Longrightarrow$ (1))

□

DAY 3

Matrices

3.1. Vector Spaces

Note to the student: For the purpose of generality, we will use F to denote a field. If you are not familiar with fields, just let $F = \mathbb{R}$ or $F = \mathbb{C}$ for now.

Definition 3.1.1. A vector space over any field F is a set V that has the following properties:

(1) (Assoc. & Comm. Addition) There is a binary operation, $+ : V \times V \to V$, on V that is associative and commutative.
(2) (Identity) There is an element $0 \in V$ such that
$$v + 0 = 0 + v = v$$
for all $v \in V$.
(3) (Inverses) For all $v \in V$, there exists $-v \in V$ such that
$$v + (-v) = 0 = (-v) + v.$$
(4) (Scaling) There is a "scaling" operation (or action), $. : F \times V \to V$, such that:
   (a) (Unitary) $1.v = v$ for all $v \in V$.
   (b) (Associative) $a.(b.v) = (ab).v$ for all $a, b \in F$ and for all $v \in V$.
   (c) (Distributive) For all $a, b \in F$ and for all $v, w \in V$,
$$a.(v + w) = a.v + a.w, \text{ and}$$
$$(a + b).v = a.v + b.v.$$

Example 3.1.2.

(1) A very common example is the set $V = \mathbb{R}^2$ over $\mathbb{R}$, with the action
$$a.(b, c) = (ab, ac).$$
The elements of $\mathbb{R}^2$ can be interpreted as arrows with their tail at the origin. The action of $\mathbb{R}$ on V is interpreted geometrically as extending, shortening, and flipping the vector arrows.
(2) The set of functions from a field F to itself forms a vector space over F: $V = \{f : F \to F\}$, with addition defined by
$$(f + g)(x) = f(x) + g(x)$$
and scaling given by
$$(c.f)(x) = cf(x).$$
(3) The set
$$V = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is differentiable.}\}$$
is a vector space over $\mathbb{R}$.

(4) The set $\mathrm{Seq}(F)$ of all infinite sequences with members from a field F is a vector space with component-wise operations:

$$(s_n) + (t_n) = (s_n + t_n)$$
$$a(s_n) = (as_n).$$

(5) The set ℓ^∞ of all bounded sequences with members in $\mathbb{C}$ is a vector space, with the same component-wise operations above.

3.2. Basic Operations: Subspaces, Sums, Quotients, and Tensor Products

(IMPPORTANT!) Note to the student: For those who are already familiar with rings and modules, note that in this "Basic Operations" section, every definition and theorem is written in such a way that "vector space V over F" can be replaced with "left R-module M," and "subspace" can be replaced with "submodule," and so on, unless noted otherwise. In the next section, however, the statements (generally) only apply to vector spaces.

Definition 3.2.1. Let V be a vector space over F. A *subspace* of V is a nonempty subset $W \subseteq V$ that is also a vector space under the same addition operation of V and the same scaling action of F.

Lemma 3.2.2 (Subspace Criterion)**.** *A subset $W \subseteq V$ is a subspace of V if and only if:*

(1) $W \neq \varnothing$, and
(2) $aw + bw' \in W$ for all $a, b \in F$ and for all $w, w' \in W$.

Proof.

□

Definition 3.2.3 (Forming New Spaces from Old)**.**

(1) The *direct product* of a family of vector spaces $\mathcal{F} = \{V_i\}_{i \in I}$ is

$$\prod_{i \in I} V_i = \{(v_i)_{i \in I} : v_i \in V_i \text{ for all } i \in I\}.$$

The direct product is a vector space with pointwise addition and pointwise scalar multiplication. When $I = \{1, \dots, n\}$ is finite, the direct product is also written

$$V_1 \times \cdots \times V_n.$$

(2) The *(external) direct sum* of a family of vector spaces $\{V_i\}_{i \in I}$ is

$$\bigoplus_{i \in I} V_i = \left\{ (v_i)_{i \in I} \in \prod_{i \in I} V_i : v_i \in V_i \text{ for all } i \in I \textbf{ and} \text{ almost all } v_i = 0. \right\}.$$

The (external) direct sum is a vector space with pointwise addition and pointwise scalar multiplication.

Definition 3.2.4 (Forming New Spaces from Subspaces)**.**

(1) Let S and T be subspaces of a vector space V. The *sum* of S and T is

$$S + T := \{s + t : s \in S, t \in T\},$$

and is also a vector space.

More generally, the *sum* of any family of subspaces $\{W_i : i \in I\}$ of a vector space V is the set of all finite sums of vectors from the union $\bigcup_i W_i$:

$$\sum_{i \in I} W_i = \left\{ w_1 + \cdots + w_n : w_i \in \bigcup_{i \in I} W_i \right\},$$

again forming a vector space.

(2) A vector space V is the *(internal) direct sum* of a family of subspaces $\{W_i : i \in I\}$ of a vector space V if the following hold:

(a) V is the sum of the family:

$$V = \sum_{i \in I} W_i.$$

(b) For each $i \in I$,

$$W_i \cap \left(\sum_{j \in I - \{i\}} W_j \right) = \{0\}.$$

In this case, each W_i is called a direct summand of V, and we write

$$V = \bigoplus_{i \in I} W_i.$$

If $I = \{1, \dots, n\}$ is finite, we may also write

$$V = W_1 \oplus \cdots \oplus W_n.$$

Theorem 3.2.5. [1]. *Let $\{V_i\}_{i \in I}$ be vector spaces. Let $\boxplus$ denote the external direct sum, and let $\oplus$ denote the internal direct sum. For a vector space V, the following are equivalent:*

(1) $V \cong \boxplus_{i \in I} V_i$.

(2) V contains subspaces $\{W_i\}_{i \in I}$ such that $W_i \cong V_i$ for all i, and every element of V can be written uniquely as a sum $\sum_{i \in I} w_i$ where $w_i \in W_i$ for all i and $w_i = 0$ for almost all i.

[1] For a proof, see [**2**, Ch.VIII,Prop.3.5]

(3) *V contains subspaces $\{W_i\}_{i\in I}$ such that $W_i \cong V_i$ for all i and $V = \bigoplus_{i\in I} W_i$.*

Remark 3.2.6.

(1) When I is finite, the external direct sum is equal to the direct product.
(2) The external direct sum considers arbitrary vector spaces (not necessarily all subspaces of some universal space) and forms a new vector space by essentially concatenating vectors into tuples.
(3) The internal direct sum indicates when a vector space V is the direct product of some of its subspaces.
(4) Notice that we use the same notation for the external direct sum and internal direct sum in their definitions. This is because the two are isomorphic. More precisely, we have by the previous theorem: If V is an external direct sum of vector spaces $\{V_i\}_{i\in I}$, then V is an internal direct sum of subspaces $\{W_i\}_{i\in I}$ where $W_i \cong V_i$ for all i, (this is (1) $\implies$ (3) from the theorem). Conversely, if V is an internal direct sum of subspaces $\{W_i\}_{i\in I}$, then V is isomorphic to the external direct sum $\bigoplus_{i\in I} W_i$, (this is (3) $\implies$ (1) from the theorem).

Definition 3.2.7 (Quotient Spaces)**.** Let V be a vector space and let $W \subseteq V$ be a subspace. The relation

$$v \sim u \iff v - u \in W$$

is an equivalence relation. When $v \sim u$, we say that v and u are *congruent modulo* W, and we write

$$v \equiv u \mod W.$$

Notice that

$$\begin{aligned}
[v] &= \{u \in V : u \equiv v\} \\
&= \{u \in V : u - v \in W\} \\
&= \{u \in V : u = v + w \text{ for some } w \in W\} \\
&= \{v + w : w \in W\} \\
&= v + W.
\end{aligned}$$

The set

$$[v] = v + W = \{v + w : w \in W\}$$

is called a *coset* of W in V, and v is called a *coset representative* for the coset $v + W$. The set of all cosets of W in V is denoted

$$V/W = \{v + W : v \in V\},$$

and is called the *quotient space of V modulo W*, which is a vector space by the operations

$$(v + W) + (u + W) = (v + u) + W \quad \text{and} \quad a(v + W) = av + W$$

for all $v, u \in V$ and for all $a \in F$.

Definition 3.2.8 (Tensor Products)**.** [2] *The General Construction:* Let V and W be vector spaces, and let Free$(V \times W)$ be the free abelian group on $V \times W$. In

[2]Note here that for the tensor product of two R-modules M and N, we would replace V for a right R-module M, and replace W for a left R-module N. After making these substitutions, the construction of the tensor product of R-modules is the same.

other words,

$$\text{Free}(V \times W) = \bigoplus_{i=1}^{\infty} V \times W.$$

Hence, each nonzero element $x \in \text{Free}(V \times W) - \{0\}$ can be (uniquely!) written as a finite sum

$$x = \sum_{i=1}^{n} a_i(v_i, w_i)$$

for a unique $n \in \mathbb{Z}_+$, unique $a_i \in \mathbb{Z} - \{0\}$, and unique $(v_i, w_i) \in V \times W$.

The *tensor product* of V and W over the field F is the quotient group

$$V \otimes_F W = \text{Free}(V \times W)/U$$

where U is the group **generated by** all expressions of the form

- $(v_1, w) - (v_2, w) - (v_1 + v_2, w)$,
- $(v, w_1) - (v, w_2) - (v, w_1 + w_2)$, and
- $(va, w) - (v, aw)$

for all $v, v_1, v_2 \in V$, for all $w, w_1, w_2 \in W$, and for all $a \in F$.

Then , $V \otimes_F W$ is an abelian group. The elements of $V \otimes_F W$ are called *tensors*. For $(v, w) \in F \times W$, we define

$$v \otimes w := (v, w) + U,$$

the coset of (v, w) in $V \otimes_F W$, called a *simple tensor*. Note that every element of $V \otimes_F W$ can be written as a finite sum of simple tensors, and this presentation is usually not unique.

Remark 3.2.9.

(1) We have
- $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$,
- $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$, and
- $va \otimes w = v \otimes aw$

for all $v, v_1, v_2 \in V$, for all $w, w_1, w_2 \in W$, and for all $a \in F$.

(2) We emphasize again that not every tensor is a simple tensor, but a (possibly non-unique) sum of simple tensors. However, many calculations/proofs for tensor products can be restricted to simple tensors since the simple tensors *generate* $V \otimes_F W$ as an abelian group.

(3) Note again that so far, we have only said $V \otimes_F W$ is an abelian group. For vector spaces V and W, it is true that $\text{Free}(V \times W)$ and U are also vector spaces over F, and that the tensor product is also itself a vector space over F, via the action [3]

$$a\left(\sum_{i=1}^{n} v_i \otimes m_i\right) = \sum_{i=1}^{n} av_i \otimes m_i.$$

[3]For the tensor product of modules, $M \otimes_R N$ is an abelian group, just as with the construction given for vector spaces. However, unlike vectors spaces, in order for $M \otimes_R N$ to also be a (left) R-module, we need M to also be a left R-module, (which is always the case when R is commutative, for example).

3.3. Span, Linear Independence, and Bases

Definition 3.3.1. Let V be a vector space over a field F.

(1) A subset S of V is said to *span* V if every element $v \in V$ can be written as a linear combination of the elements of S:
$$v = a_1 s_1 + \cdots a_n s_n$$
for $a_i \in F$ and $s_n \in S$.
(2) A subset $S \subseteq V$ is said to be a set of *linearly independent vectors* if for all $n \in \mathbb{Z}_+$ and for all $v_1, \ldots, v_n \in S$, whenever we have an equation
$$a_1 v_1 + \cdots + a_n v_n = 0_V,$$
for $a_1, \ldots, a_n \in F$, then $a_i = 0_F$ for all $1 \leq i \leq n$.
(3) A *basis* of V is an **ordered** set $S \subseteq V$ of linearly independent vectors that span V.

Example 3.3.2 (Polynomials over F)**.** The set $V = F[x]$ of polynomials in the variable x with coefficients from the field F is a vector space over F. The elements $1, x, x^2, \ldots$ are linearly independent by definition, these elements span V, and hence they are a basis for V.

Theorem 3.3.3. *Let V be a vector space.*

(1) V has a basis.
(2) If $S \subseteq V$ spans V, then S contains a basis of V.
(3) (Replacement Theorem) Suppose V has a finite basis. Every (finite) linearly independent subset of V is contained in a basis for V. More precisely:
Suppose $\{b_1, \ldots, b_m\}$ is a finite basis for V, and suppose $\{a_1, \ldots, a_n\} \subseteq V$ is a linearly independent set. Then there is an ordering of $a_1, \ldots, a_n$ such that for all $0 \leq k \leq m$, the set
$$\{b_1, \ldots, b_k, a_{k+1}, \ldots, a_n\}$$
is a basis for V. In particular, $m \leq n$.
(4) Suppose V has a finite basis of cardinality n. Every linearly independent subset of V has $\leq n$ elements, and every spanning set of V has $\geq n$ elements. Moreover, every basis of V has n elements, and n is called the dimension of V*, written $\dim_F V = n$.*
(5) Suppose $\dim_F V < \infty$. Every subspace of V has a complement. More precisely, if W is a subspace of V, then there exists a subspace U of V such that $V = W \oplus U$.

3.4. Linear Transformations and Dual Spaces

In essence, a linear transformation is a function between vector spaces that preserves the vector space operations. More formally,

Definition 3.4.1. Let V, W be vector spaces over a field F.

(1) A function $\varphi : V \to W$ is a *linear transformation* if
$$\varphi(av_1 + bv_2) = a\varphi(v_1) + b\varphi(v_2)$$
for all $v_1, v_2 \in V$ and for all $a, b \in F$.

(2) A bijective linear transformation $\varphi : V \to W$ is called an *isomorphism*, and we denote this fact by writing $V \cong W$.
(3) A linear transformation $\varphi : V \to V$, from the vector space V into itself is called a *linear operator on* V.
(4) The set of all linear functionals on V,
$$V^* = \{\varphi : V \to F : \varphi \text{ is a linear transformation.}\}$$
is called the *dual space of* V.
(5) The *kernel* of φ, denoted $\mathrm{Ker}(\varphi)$, is the subspace of V given by
$$\mathrm{Ker}(\varphi) = \{v \in V : \varphi(v) = 0\}.$$
The dimension of $\mathrm{Ker}(\varphi)$ is called the *nullity* of φ.
(6) The *rank* of φ, denoted $\mathrm{Rk}(\varphi)$, is the dimension of the image of φ.

Remark 3.4.2.

(1) A linear transformation $\varphi : V \to W$ is completely determined by its action on a basis of V. In particular, suppose V is finite dimensional (though this is also true if $\dim_F V = \infty$), with basis $\{x_1, \ldots, x_n\}$. If $v \in V$, then we can write
$$v = \sum_{i=1}^{n} a_i x_i$$
for $a_i \in F$. Therefore, since φ is a linear transformation,
$$\varphi(v) = \varphi\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i \varphi(x_i).$$
In other words, in order to understand φ, one only needs to know $\varphi(x_i)$ for each $1 \leq i \leq n$.
(2) If $\dim_F V = n$, then $V \cong F^n := \underbrace{F \times F \times \cdots \times F}_{n \text{ factors}}$.
(3) The kernel of a linear transformation is equal to the zero subspace if and only if the transformation is injective.

Theorem 3.4.3. *Let V be a finite dimensional vector space over a field F. Then V^* is a finite dimensional vector space with $\dim_F V^* = \dim_F V$. More precisely, if $\{x_1, \ldots, x_n\}$ is a basis of V, then $\{x_1^*, \ldots, x_n^*\}$ is a basis of V^*, where*
$$x_i^*(x_j) = \delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Proof.

□

3.5. Matrix of a Linear Transformation and Change of Basis

Definition 3.5.1. Let V (resp. W) be a finite dimensional vector space over a field F with basis $B = \{v_1, \ldots, v_n\}$ (resp. $E = \{w_1, \ldots, w_m\}$). Let $\varphi : V \to W$ be a linear transformation. Then for all $j \in \{1, \ldots, n\}$, there exists unique $a_{ij} \in F$ $(i = 1, \ldots, m)$ such that

$$\varphi(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

The $m \times n$ matrix (a_{ij}) is called the *matrix of φ with respect to the bases B, E*, denoted by $M_B^E(\varphi)$.

Remark 3.5.2.

(1) Let V and W be finite dimensional vector spaces of dimensions n and m, respectively. The collection of all linear transformations from V to W is itself a vector space, denoted $\mathrm{Hom}_F(V, W)$. The set of all $m \times n$ matrices with entries in F is also a vector space, denoted $M_{m\times n}(F)$. One can show that there is an isomorphism

$$\mathrm{Hom}_F(V, W) \xrightarrow{\cong} M_{m\times n}(F).$$

In other words, matrices and linear maps are the "same thing" in the sense that, given a matrix , there is a unique linear transformation associated to it, and vice versa.

(2) Let U, V, W be finite dimensional vector spaces over F with bases B_U, B_V, B_W, respectively. Let $\varphi : U \to V$ and $\psi : V \to W$ be linear transformation. Then matrix multiplication shows:

$$M_{B_U}^{B_W}(\psi \circ \varphi) = M_{B_V}^{B_W}(\psi) \cdot M_{B_U}^{B_V}(\varphi).$$

Example 3.5.3.

(1) Let $V = W = \mathbb{Q}^3 = \{(x, y, z) : x, y, z \in \mathbb{Q}\}$ be the usual 3-dimensional vector space of ordered 3-tuples with entries from the field $F = \mathbb{Q}$ of rational numbers and suppose $\varphi : V \to V$ is the linear transformation

$$\varphi(x, y, z) = (9x + 4y + 5z, -4x - 3z, -6x - 4y - 2z).$$

In the standard basis for V, $B = \{e_1, e_2, e_3\}$, where

$$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1),$$

the matrix A representing this linear transformation with respect to these bases is

$$A = \begin{pmatrix} 9 & 4 & 5 \\ -4 & 0 & -3 \\ -6 & -4 & -2 \end{pmatrix}.$$

(2) Let $V = W$ be the 2-dimensional space of solutions of the differential equation

$$y'' - 4y' + 2y = 0$$

over $\mathbb{C}$ and let $B = E = \{v_1 = e^t, v_2 = e^{2t}\}$. Since the coefficients of this equation are constant it is easy to check that if y is a solution then

its derivative y' is also a solution. It follows that the differentiation map (with respect to t),

$$\varphi = \frac{d}{dt} : V \longrightarrow V,$$

is a linear transformation from V to itself. Since

$$\varphi(v_1) = \frac{d(e^t)}{dt} = e^t = v_1 \quad \text{and} \quad \varphi(v_2) = \frac{d(e^{2t})}{dt} = 2e^{2t} = 2v_2,$$

it follows that the matrix corresponding to φ is the diagonal matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Definition 3.5.4. Let V be vector space of finite dimension n, and let B and E be two bases of V. Let $P = M_E^B(\mathrm{id}_V)$ (i.e. express the j^{th} element of E in terms of B and put it into the j^{th} column of P).

Then for all $\varphi \in \mathrm{Hom}_F(V, V) =: \mathrm{End}_F(V)$, we have

$$M_E^E(\varphi) = P^{-1} M_B^B(\varphi) P.$$

This is because

$$\begin{aligned} M_B^B(\varphi) \cdot P &= M_B^B(\varphi) \cdot M_E^B(\mathrm{id}_V) \\ &= M_E^B(\varphi) \\ &= M_E^B(\mathrm{id}_V) \cdot M_E^E(\varphi) \\ &= P \cdot M_E^E(\varphi). \end{aligned}$$

We call $P = M_E^B(\mathrm{id}_V)$ the *change of basis matrix from B to E.*

Example 3.5.5. Let $V = \mathbb{Q}^3$ and $\varphi : V \to V$ be the linear transformation from Example 3.5.3. If $B = \{b_1, b_2, b_3\}$ is the standard basis, we found that

$$A = M_B^B(\varphi) = \begin{pmatrix} 9 & 4 & 5 \\ -4 & 0 & -3 \\ -6 & -4 & -2 \end{pmatrix}.$$

Consider the basis $E = e_1 = (2, -1, -1), e_2 = (1, 0, -1), e_3 = (3, -2, -2)$ of V. Since

$$\begin{aligned} \varphi(e_1) &= \varphi((2, -1, -2)) = (4, -2, -4) = 2e_1 \\ \varphi(e_2) &= \varphi((1, 0, -1)) = (4, -1, -4) = e_1 + 2e_2 \\ \varphi(e_3) &= \varphi((3, -2, -2)) = (9, -6, -6) = 3e_3 \end{aligned}$$

we have

$$C := M_E^E = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Since the change of basis $P = M_E^B(\mathrm{id}_V)$, we need to write each element of E in terms of the elements of B:

$$\begin{aligned} e_1 &= 2b_1 - b_2 - 2b_3 \\ e_2 &= b_1 - b_3 \\ e_3 &= 3b_1 - 2b_2 - 2b_3. \end{aligned}$$

So we have

$$P = M_E^B(\mathrm{id}_V) = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & -2 \\ -2 & -1 & -2 \end{pmatrix} \quad \text{and} \quad P^{-1} = \begin{pmatrix} -2 & -1 & -2 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Moreover, $P^{-1}AP = C$, as you can check.

3.6. Row Reduction

Consider a system of m linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n &= c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n &= c_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n &= c_m \end{aligned}$$

in the n unknowns $x_1, x_2, \ldots, x_n$ where a_{ij}, c_i for $1 \leq i \leq m$ and $1 \leq j \leq n$ are elements of a field F. Associated to this system is the *coefficient matrix*:

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

and the *augmented matrix*

$$(A|C) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{array}\right).$$

The set of solutions in F of this system of equations is not altered if we perform any of the following three operations:

(1) interchange any two equations
(2) add a multiple of one equation to another
(3) multiply any equation by a nonzero element from F,

which correspond to the following three *elementary row operations* on the augmented matrix:

(1) interchange any two rows
(2) add a multiple of one row to another
(3) multiply any row by a nonzero element from F.

Proposition 3.6.1.

(1) The relation "$A \sim C$ if and only if A can be row reduced to C" is an equivalence relation. If $A \sim C$, these two matrices are said to be row equivalent*.*
(2) The rank of two row equivalent matrices is the same.

An $m \times n$ matrix is said to be in *reduced row echelon form* if:

(1) the first nonzero entry a_{ij_i} in row i is 1 and all other entries in the corresponding j_i^{th} column are zero, and
(2) $j_1 < j_2 < \cdots < j_r$ where r is the number of nonzero rows, i.e., the number of initial zeros in each row is strictly increasing (hence the term *echelon*).

The point of reduced row echelon form is that the corresponding system of linear equations is in a particularly simple form, from which solutions can be easily determined. This is particularly useful when trying to determine the kernel and image of a matrix. See [**1**, Exercise 11.2.31].

Example 3.6.2.

(1) The following two matrices are in reduced row echelon form:

$$\left(\begin{array}{cccccc|c} 1 & 0 & 5 & 7 & 0 & 3 & 0 \\ 0 & 1 & -1 & 1 & 0 & -4 & -1 \\ 0 & 0 & 0 & 0 & 1 & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \text{and} \quad \left(\begin{array}{cccc|c} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -3 \end{array}\right).$$

(2) Find the reduced row echelon form of the matrix

$$A = \begin{pmatrix} -7 & -1 & -4 \\ 7 & 1 & 3 \\ 1 & 0 & 0 \end{pmatrix}.$$

3.7. Rank-Nullity Theorem

Theorem 3.7.1 (Rank-Nullity Theorem)**.** *Let $\varphi : V \to W$ be a linear transformation. Then*

$$\dim_F(\mathrm{Ker}(\varphi)) + \dim_F(\mathrm{Im}(\varphi)) = \dim_F(V).$$

In other words, the rank plus nullity of φ equals the dimension of V.

Proof.

□

Corollary 3.7.2. *Let $\varphi : V \to W$ be a linear transformation where $\dim_F(V) = \dim_F(W) < \infty$. Then φ is injective if and only if it is surjective.*

Remark 3.7.3. Let A be a $m \times n$ matrix that is in reduced row echelon form, and moreover suppose that A has r nonzero rows. One can show that $\mathrm{Rk}(A) = r$. Combining this fact with the Rank-Nullity theorem, we see that

$$\dim(\mathrm{Ker}(A)) = n - r.$$

Example 3.7.4. Let $\varphi : \mathbb{R}^4 \to \mathbb{R}^2$ such that

$$\varphi((1,0,0,0)) = (1,-1) \qquad \varphi((1,-1,0,0)) = (0,0)$$
$$\varphi((1,-1,1,0)) = (1,-1) \qquad \varphi((1,-1,1,-1)) = (0,0)$$

Determine a basis for the image and for the kernel of φ.

DAY 4

Matrices (continued)

4.1. Trace and Determinant

Definition 4.1.1.

(1) The *trace* of a matrix A, denoted $\mathrm{Tr}(A)$, is the sum of the elements on the main diagonal of A. In particular (see Theorem 4.1.2), for each $n \in \mathbb{Z}_+$, $\mathrm{Tr} : M_{n\times n}(F) \to F$ is a linear functional on the vector space $M_{n\times n}(F)$.

(2) Let V and W be vector spaces over a field F. For $n \in \mathbb{Z}_+$, a map $\varphi : V^n \to W$ is called an *n-multilinear function on V* if for all $i \in \{1, \ldots, n\}$ and for all $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in V$, the map

$$\begin{aligned} \varphi_i : V &\longrightarrow W \\ x &\longmapsto \varphi(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) \end{aligned}$$

is a linear transformation. If $V = F$, then φ is called an *n-multilinear form on V*.

(3) The *determinant* of a $n \times n$ matrix $A = (a_{ij})$, denoted by $\mathrm{Det}(A)$, is given by

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

where S_n is the permutation group on n symbols, and ϵ is the sign operator. For each $n \in \mathbb{Z}_+$, $\mathrm{Det} : M_{n\times n}(F) \to F$ is an n-multilinear linear functional on the columns of matrices.

Theorem 4.1.2 (Properties of the Trace)**.** *Let A, B, C be square matrices with entries from a field F.*

(1) $\mathrm{Tr}(rA) = r\,\mathrm{Tr}(A)$ *for* $r \in F$.
(2) $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$.
(3) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.
(4) $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$. *However,* $\mathrm{Tr}(ABC)$ *may not be equal to* $\mathrm{Tr}(ACB)$.
(5) If $A = P^{-1}BP$ *for some invertible matrix* P, *then* $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.

Theorem 4.1.3 (Properties of the Determinant)**.** *Let A, B, C be square matrices with entries from a field F.*

(1) $\mathrm{Det}(AB) = \mathrm{Det}(A)\,\mathrm{Det}(B)$.
(2) If $A \in M_{n\times n}(F)$ *and* $r \in F$, *then*

$$\mathrm{Det}(rA) = r^n\,\mathrm{Det}(A).$$

(3) It A is invertible, then $\mathrm{Det}(A) \neq 0$ *and* $\mathrm{Det}(A^{-1}) = \frac{1}{\mathrm{Det}(A)}$.
(4) $\mathrm{Det}(A) = \mathrm{Det}(A^T)$ *where* A^T *denotes the transpose of* A.

(5) *Suppose* $A = (a_{ij}) \in M_{n\times n}$. *For each* $i, j \in 1, \ldots, n$, *let* A_{ij} *denote the* $(n-1)\times(n-1)$ *matrix obtained by deleting the* i^{th} *row and* j^{th} *column of* A, *calle the* ij minor *of* A. *Then* $(-1)^{i+j}\operatorname{Det}(A_{ij})$ *is called the* ij cofactor of A. *For each fixed* $i \in \{1, \ldots, n\}$, *the determinant of* A *can be computed from the formula*

$$\begin{aligned}\operatorname{Det}(A) &= \sum_{k=1}^{n}(-1)^{i+k}a_{ik}\operatorname{Det}(A_{1k}) \\ &= (-1)^{i+1}a_{i1}\operatorname{Det}(A_{1i}) + \cdots + (-1)^{i+n}a_{in}\operatorname{Det}(A_{1n}).\end{aligned}$$

Example 4.1.4. Consider a matrix $A \in M_{2\times 2}(F)$, given by

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

We will use the definition of the derivative to compute $\operatorname{Det}(A)$. Recall that $S_2 = \{\mathrm{id} = (1), \tau = (12)\}$. The determinant of A is

$$\begin{aligned}\operatorname{Det}(A) &= \sum_{\sigma\in S_2}\epsilon(\sigma)a_{\sigma(1),1}a_{\sigma(2),2} \\ &= \epsilon(\mathrm{id})\, a_{\mathrm{id}(1),1}a_{\mathrm{id}(2),2} + \epsilon(\tau)\, a_{\tau(1),1}a_{\tau(2),2} \\ &= (1)a_{1,1}a_{2,2} + (-1)a_{2,1}a_{1,2} \\ &= a_{11}a_{22} - a_{21}a_{12}.\end{aligned}$$

4.2. Eigenvectors and Eigenvalues

Definition 4.2.1. Let $A \in M_{n\times n}(F)$ for a field F, and let $I_n \in M_{n\times n}(F)$ denote the identity matrix.

(1) An *eigenvalue* of A is an element $\lambda \in F$ if there exits a nonzero vector $v \in F^n$ such that $Av = \lambda v$, or equivalently, if $(A - \lambda I_n)v = 0$. Such a nonzero vector v is called an *eigenvector* of A corresponding to λ.

(2) Suppose A has eigenvalue $\lambda \in F$. The *eigenspace* of A corresponding to λ is the set

$$E_A(\lambda) = \{v \in F^n : Av = \lambda v\} \cup \{0\}.$$

Then $E_A(\lambda)$ is a vector space, and its dimension is called the *geometric multiplicity* of λ.

(3) The *characteristic polynomial* of A is the polynomial $p_A(x) \in F[x]$ defined by

$$p_A(x) = \operatorname{Det}(A - xI_n).$$

Note that A has eigenvalue λ if and only if $p_A(\lambda) = 0$. [1] If A has eigenvalue λ, the highest power of $(x - \lambda)$ that divides $p_A(x)$ is called the *algebraic multiplicity* of λ.

Theorem 4.2.2 (Existence of Eigenvalues)**.** *If* $F = \mathbb{C}$ *(or if* F *is any algebraically closed field), then a square matrix over* F *always has at least one eigenvalue.*

[1] Of course, this needs to be proven.

PROOF.

□

Example 4.2.3.

(1) Suppose

$$A = \begin{pmatrix} -13 & -8 & -4 \\ 12 & 7 & 4 \\ 24 & 16 & 7 \end{pmatrix}.$$

Then

$$p_A(x) = \begin{vmatrix} -13-x & -8 & -4 \\ 12 & 7-x & 4 \\ 24 & 16 & 7-x \end{vmatrix} = \cdots = -(x-3)(x+1)^2.$$

So A has eigenvalues 3 and -1, with algebraic multiplicities 1 and 2, respectively. Finding the geometric multiplicities, we have:

$$\lambda = 3: \qquad A - 3I_3 = \begin{pmatrix} -16 & -8 & -4 \\ 12 & 4 & 4 \\ 24 & 16 & 4 \end{pmatrix} \overset{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

$$E_A(3) = \left\langle \left\{ \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\} \right\rangle = \left\langle \left\{ \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\} \right\rangle$$

$$\lambda = -1: \qquad A + I_3 = \begin{pmatrix} -12 & -8 & -4 \\ 12 & 8 & 4 \\ 24 & 16 & 8 \end{pmatrix} \overset{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E_A(-1) = \left\langle \left\{ \begin{pmatrix} -\frac{2}{3} \\ 1 \\ 0 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} -\frac{1}{3} \\ 0 \\ 1 \end{pmatrix} \right\} \right\rangle.$$

$$= \left\langle \left\{ \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} \right\} \right\rangle.$$

Therefore, the geometric multiplicities for both eigenvalues match their respective algebraic multiplicities. This is not always the case, and in general, for a given eigenvalue, the geometric multiplicity can never exceed the algebraic multiplicity.

4.3. Jordan Canonical Form

Definition 4.3.1. Suppose $\varphi : V \to V$ is a linear transformation on a finite dimensional vector space V. A basis B of V is called a *Jordan basis* for φ if with respect to this basis, the matrix $M_B^B(\varphi)$ is given by

$$M_B^B(\varphi) = \begin{pmatrix} A_1 & \cdots & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix},$$

which is called a *Jordan matrix*, where each A_j is an upper-triangular matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix},$$

which is called a *Jordan block corresponding to* λ_j.

Theorem 4.3.2 (Jordan Canonical Form)**.** *Suppose V is a finite dimensional vector space over $\mathbb{C}$ (or any algebraically closed field). If $\varphi : V \to V$ is a linear transformation, then there is a basis of V that is a Jordan basis for φ.*

Remark 4.3.3.

(1) The above theorem, together with our discussion on "change of basis," we can conclude the following: Given any matrix A (with corresponding linear transformation $\varphi : V \to V$) over $\mathbb{C}$, there exists an invertible matrix P such that
$$P^{-1}AP = J$$
where J is a Jordan matrix for φ. In other words, any matrix over $\mathbb{C}$ is similar to a Jordan matrix, and we refer to this matrix J as the *Jordan normal form of A*.

(2) The matrices J and P have the following properties:
   (a) Each Jordan block within the Jordan matrix J corresponds to an eigenvalue of A. Given an eigenvalue λ of A, the number of Jordan blocks corresponding to λ equals the geometric multiplicity of λ.

(b) The algebraic multiplicity of an eigenvalue λ is equal to the sum of the sizes of all Jordan blocks corresponding to λ.

Example 4.3.4. Let

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

Then $p_A(x) = (x-1)(x-2)(x-4)^2$. So A has eigenvalues 1,2 and 4, with algebraic multiplicities 1,1, and 2, respectively.

Find their geometric multiplicities and the Jordan normal form of A.

4.4. The Spectral Theorem

Definition 4.4.1. Let $A \in M_{m\times n}(\mathbb{C})$ be written in the standard basis for $\mathbb{C}$.

(1) The *adjoint* of A, denoted A^*, is the matrix

$$A^* := \overline{A}^T.$$

In other words, A^* is the conjugate transpose of A.

(2) A is called *normal* if it commutes with its adjoint, i.e. if

$$AA^* = A^*A.$$

(3) A is called *self-adjoint* if $A = A^*$. Note that if A has all real-valued entries and is self-adjoint, then $A = A^T$, i.e. A is *symmetric*.

(4) A basis B for a complex vector space is called *orthonormal* if:
   (a) Each pair of vectors in B are orthogonal (i.e. dot product equal to zero), and
   (b) the length of each vector in B is equal to 1.

Proposition 4.4.2. *If a matrix $A \in M_{n\times n}(\mathbb{C})$ is self-adjoint, then the eigenvalues of A are real.*

Proof.

□

Remark 4.4.3. The Jordan Form tells us that if A is an $n \times n$ over $\mathbb{C}$, A has n distinct eigenvectors if and only if A is diagonalizable. In particular, the diagonal matrix D has the eigenvalues of A on its diagonal, and the change of basis matrix contains the eigenvectors of A.

More precisely, let $v_1, \ldots, v_n$ be n distinct eigenvectors of A, corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$, and define

$$P := [v_1|\cdots|v_n].$$

Then P is invertible, and

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Note that since $v_1, \ldots, v_n$ are distinct eigenvectors, they are linearly independent (prove it!), and hence form a basis for $\mathbb{C}^n$. Since the change of basis matrix is made up of eigenvectors of A, we say that *A is diagonalizable with respect to n distinct eigenvectors.*

The following Spectral Theorem gives us a criteria for when a matrix A is diagonalizable with respect to a set of *orthonormal* eigenvectors of A.

Theorem 4.4.4 ((Complex) Spectral Theorem)**.** *Let $A \in M_{m\times n}(\mathbb{C})$ be written in the standard basis for $\mathbb{C}$. The following are equivalent:*

(1) A is normal.
(2) A is diagonalizable with respect to an orthonormal set of eigenvectors of A.

There is also a version of the Spectral Theorem when A has real entires:

Theorem 4.4.5 ((Real) Spectral Theorem)**.** *Let $A \in M_{m\times n}(\mathbb{R})$ be written in the standard basis for $\mathbb{R}$. The following are equivalent:*

(1) A is symmetric.
(2) A is diagonalizable with respect to an orthonormal set of eigenvectors of A.

Remark 4.4.6. Suppose A is normal (or symmetric).

As with the Jordan normal form, the matrices P and D for which $P^{-1}AP = D$ are such that D has diagonal entries which are the eigenvalues of A, and P contains eigenvectors of A. However, the Spectral Theorem says that P can be chosen so that the eigenvectors of A which make up the columns of P are an orthonormal set. Moreover, this implies that $P^{-1} = P^*$, in which case we call P a *unitary* matrix, (or an *orthogonal* matrix if A is real, in which case $P^{-1} = P^T$). Hence the Spectral Theorem can be stated as follows:

Let A be a square matrix over $\mathbb{C}$ (resp. $\mathbb{R}$). If A is normal (resp. symmetric), then A can be written

$$A = PDP*$$

where the columns of P consist of a set of orthonormal eigenvectors of A, and the matrix D is diagonal with diagonal entries the eigenvalues of A.

Example 4.4.7. The matrix

$$A = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

is not symmetric, but it is normal. The eigenvalues of A are the roots to the characteristic polynomial[2]

$$p_A(x) = x^2 - \mathrm{Tr}(A)x + \mathrm{Det}(A) = x^2 - 4x + 13,$$

[2]This formula only holds for 2×2 matrices.

which are $2 \pm 3i$. Finding eigenvectors,

$$\lambda_1 = 2+3i: \qquad A-(2+3i)I_3 = \begin{pmatrix} -3i & -3 \\ 3 & -3i \end{pmatrix} \overset{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}$$
$$v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$$
$$\lambda_2 = 2-3i: \qquad A-(2-3i)I_3 = \begin{pmatrix} 3i & -3 \\ 3 & 3i \end{pmatrix} \overset{\text{RREF}}{\longrightarrow} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}$$
$$v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

However, these eigenvectors are not length 1. So, we normalize these vectors:

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}$$
$$v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

Now, v_1, v_2 provide a *orthonormal* basis for $\mathbb{C}^2$, and if $P = [v_1|v_2]$, then P^*AP is the diagonal matrix

$$\begin{pmatrix} 2+3i & 0 \\ 0 & 2-3i \end{pmatrix}.$$

4.5. Singular Value Decomposition

We now consider the question of decomposing a rectangular matrix (not necessarily square).

Recall that the Spectral Theorem said that any normal (or symmetric) square matrix A can be written $A = PDP^*$ where D is diagonal and P is a unitary (or orthogonal) matrix (i.e. $PP^* = I$). The following *singular value decomposition* provides a similar decomposition for any *rectangular* matrix.

Definition 4.5.1. Let $A \in M_{m\times n}(\mathbb{C})$. The square roots of the eigenvalues of the square (and normal (or symmetric)) matrix A^*A are the *singular values* of A, denoted $\sigma_1, \ldots, \sigma_r$.

Remark 4.5.2. Note that this definition requires the assumption that the eigenvalues of A^*A are both real and non-negative. First, the eigenvalues of A are real since A^*A is self-adjoint. The second fact requires knowledge of inner products: Let $\langle a, b\rangle$ denote the standard inner product on $\mathbb{C}^n$:

$$\langle a, b\rangle = \langle (a_1, \ldots, a_n), (b_1, \ldots, b_n)\rangle = a_1\overline{b_1} + \cdots + a_n\overline{b_n}.$$

For any nonzero vector $v \in \mathbb{C}^n$

$$\langle A^*Av, v\rangle = \langle Av, Av\rangle \geq 0.$$

Hence if λ is an eigenvalue of A^*A with eigenvector v, we have

$$0 \leq \langle A^*Av, v\rangle = \langle \lambda v, v\rangle = \lambda \langle v, v\rangle,$$

and since $\langle v, v\rangle$ is always nonnegative, then λ must be nonnegative.

Theorem 4.5.3 (Singular Value Decomposition (SVD))**.** *Let $A \in M_{m\times n}(\mathbb{C})$. Then A can be factored into*

$$A = U\Sigma V^* = (\textit{unitary})(\textit{diagonal})(\textit{unitary}).$$

The columns of U $(m \times m)$ are eigenvectors of AA^*, and the columns of V $(n \times n)$ are the eigenvectors of A^*A. The $\mathrm{Rk}(A) = r$ singular *values on the diagonal of Σ $(m \times n)$ are the square roots of the nonzero eigenvalues of both AA^* and A^*A.*

Remark 4.5.4. The following remarks come directly from [**6**, Section 6.3].

(1) U and V give orthonormal bases for all four fundamental subspaces:

$$\begin{aligned}\text{first } r \text{ columns of } U &: \text{column space of } A\\ \text{last } m-r \text{ columns of } U &: \text{nullspace of } A^*\\ \text{first } r \text{ columns of } V &: \text{row space of } A\\ \text{last } n-r \text{ columns of } V &: \text{nullspace space of } A\end{aligned}$$

(2) The SVD chooses the above bases in an extremely special way. They are more than just orthonormal. The equality $AV = U\Sigma$ tells us that when A multiplies a column v_j of V, it produces σ_j times the j^{th} column of U, where σ_j is the j^{th} diagonal entry of Σ.

This is one of the most important aspects of the SVD. The orthonormal basis in U is such that A sends each basis vector to a multiple of a vector that is also a member of an orthonormal basis:

$$Av_j = \sigma_j u_j.$$

(3) U must be the eigenvector matrix for AA^*, and V must be the eigenvector matrix for A^*A. This is because of the following computations:

$$\begin{aligned}AA^* &= (U\Sigma V^*)(V\Sigma^T U^*) = U\Sigma\Sigma^* U^*\\ A^*A &= (V\Sigma^* U^*)(U\Sigma V^*) = V\Sigma^*\Sigma V^*.\end{aligned}$$

For AA^*, the eigenvalue matrix in the middle is $\Sigma\Sigma^*$, which is $m \times m$ with $\sigma_1^2, \ldots, \sigma_r^2$ on the diagonal. The diagonal matrix $\Sigma^*\Sigma$ for A^*A has the same $\sigma_1^2, \ldots, \sigma_r^2$, but it is $n \times n$.

Example 4.5.5.

(1) Let

$$A = \begin{pmatrix}-1\\2\\2\end{pmatrix}.$$

A^*A is a 1×1 matrix, and AA^* is 3×3. They both have eigenvalue 9. The two zero eigenvalues of AA^* leave some freedom for the eigenvectors in columns 2 and 3 of U.

The SVD of A is then

$$\begin{aligned}A = \begin{pmatrix}-1\\2\\2\end{pmatrix} &= \begin{pmatrix}-\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{pmatrix}\begin{pmatrix}3\\0\\0\end{pmatrix} \quad (1)\\ &= U_{3\times 3}\Sigma_{3\times 1}V^T_{1\times 1}.\end{aligned}$$

(2) Find the SVD of the matrix

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

DAY 5

Groups

5.1. Basic Definitions

Definition 5.1.1. A *group* is a nonempty set G together with a binary operation

$$\begin{aligned} * : G \times G &\longrightarrow G \\ (a, b) &\longmapsto a * b \end{aligned}$$

such that

(a) $*$ is associative, i.e., for all $a, b, c \in G$

$$a * (b * c) = (a * b) * c.$$

(b) There exists $e \in G$ (called the *identity* of G) such that for all $a \in G$

$$a * e = a = e * a.$$

(c) For all $a \in G$, there exists $a^{-1} \in G$ (called an *inverse* for a) such that

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

Notation. G or $(G, *)$.

We say that G is *abelian* if $*$ is commutative, i.e., for all $a, b \in G$

$$a * b = b * a.$$

Conventions.

(1) We usually write groups *multiplicatively*, i.e., we write $*$ as $\cdot$ and $a \cdot b$ as ab. We often write the identity as 1 or 1_G.
(2) If G is abelian, we write it *additively*, i.e., we write $*$ as $+$, and the identity element as 0 (also called the *zero element*). The additive inverse of $a \in G$ is denoted by $-a$.

Definition 5.1.2. The *order* of G, denoted by $|G|$ or $\#G$, is the number of elements of G. If $g \in G$, the *order* of g, denoted by $|g|$ or $o(g)$, is the smallest positive integer n such that

$$g^n := \underbrace{g \cdot g \cdots g}_{n \text{ factors}} = 1_G$$

if such an n exists. Otherwise, g has infinite order. (If g is additively written, this is written as $ng := \underbrace{g + g + \cdots + g}_{n \text{ summands}} = 0$.)

Example 5.1.3.

(1) $G = \{1_G\}$ or $G = 1$, the *trivial* group.
(2) $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$ are abelian.
(3) $(\mathbb{Q}^\times) = (\mathbb{Q} - \{0\}, \cdot), (\mathbb{R}^\times) = (\mathbb{R} - \{0\}, \cdot), (\mathbb{C}^\times) = (\mathbb{C} - \{0\}, \cdot), \mathbb{Z}^\times = U(\mathbb{Z}) = (\{\pm 1\}, \cdot)$.
(4) *Integers Modulo* n ($n \in \mathbb{Z}$), denoted $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}/n$, where

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

and

$$\overline{a} = \{r \in \mathbb{Z} : n|(a - r)\} = \{a + n\ell : \ell \in \mathbb{Z}\}.$$

Addition an multiplication, $\overline{a} + \overline{b} := \overline{a \cdot b}, \overline{a} \cdot \overline{b} := \overline{a + b}$ are well-defined.

Then $(\mathbb{Z}/n, +)$ is an additive abelian group, and

$$\begin{aligned}(\mathbb{Z}/n)^\times &= \{\overline{a} \in \mathbb{Z}/n : \overline{a} \text{ has a multiplicative inverse}\} \\ &= \{\overline{a} \in \mathbb{Z}/n : \gcd(a, n) = 1\}\end{aligned}$$

is a multiplicative abelian group of order $\varphi(n)$ (Euler function). For example,

$$\begin{aligned}(\mathbb{Z}/6)^\times &= \{\overline{1}, \overline{5}\}, \text{ and} \\ (\mathbb{Z}/7)^\times &= \{\overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}, \overline{7}\}.\end{aligned}$$

(5) *Direct Product*: Let $(G, *)$ and $(H, \circ)$ be groups. Define

$$G \times H = \{(g, h) : g \in G, h \in H\}$$

with operation given by

$$(g_1, h_1)(g_2, h_2) = (g_1 * g_2, h_1 \circ h_2).$$

Then this is a group, called the *direct product* of G and H.
(6) Let F be a field (eg. $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}/p$ when p is prime). Let $n \in \mathbb{Z}_+$. Then

$$\mathrm{GL}_n(F) := \{A : A \in M_{n\times n}(F) \text{ and } \mathrm{Det}(A) \neq 0_F\}$$

is a group under matrix multiplication, called the *general linear group of degree n over F*. The *special linear group* is the (sub)group

$$\mathrm{SL}_n(F) := \{A \in \mathrm{GL}_n(F) : \mathrm{Det}(A) = 1_F\}.$$

5.2. Symmetric, Cyclic, and Dihedral Groups

Definition 5.2.1.

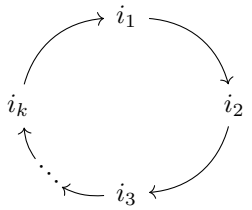
(1) Let $A \neq \varnothing$ be a set. A *permutation of A* is a bijective map $\varphi : A \to A$.
(2) Let $\Omega \neq \varnothing$ be a set. Then

$$S_\Omega := \{\sigma \mid \sigma : \Omega \to \Omega \text{ is a permutation } \}$$

is a group under composition, called the *symmetric group on* Ω.

Special case: $\Omega = \{1, 2, \ldots, n\}$ for some $n \in \mathbb{Z}^+$. Then S_Ω is denoted by S_n and called the *symmetric group of degree n*. Note: $|S_n| = \#S_n = n!$.

Definition 5.2.2. Let $1 \leq k \leq n$. A *cycle of length k* (also called a *k-cycle*) is a permutation $\gamma \in S_n$ of the form



for certain pairwise disjoint $i_1, \ldots, i_k \in \{1, \ldots, n\}$ and $\gamma(j) = j$ for all $j \notin \{i_1, \ldots, i_k\}$.

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

Remark 5.2.3.

(1) Every $\sigma \in S_n$ can be written as a product of disjoint cycles. Hence, the cycles act on pairwise disjoint sets. This decomposition is unique up to permutation of the cycles.
(2) Every $\sigma \in S_n$ can be written as a product of 2-cycles (not necessarily disjoint).
(3) We denote the identity in S_n by 1 or id or (1).
(4) Multiplication (map composition) is from right to left: For example,

$$(1,2,3,4,5)(3,5,6,4)(7,2,1) = (1,7,4,3)(2)(5,6) = (1,7,4,3)(5,6).$$

(5) Disjoint cycles commute.
(6) Non-disjoint cycles do not commute. Hence S_n for $n \geq 3$ is non-abelian.
(7) If $\gamma = (i_1, \ldots, i_k)$ is a k-cycle, then $o(\gamma) = k$ and $\gamma^{-1} = (i_k, \ldots, i_1)$.
(8) If $\sigma \in S_n$ is written as $\sigma = \gamma_1 \cdots \gamma_r$ where $\gamma_1, \ldots, \gamma_r$ are disjoint cycles, then

$$o(\sigma) = \operatorname{lcm}(o(\gamma_1), \ldots, o(\gamma_r)),$$

and $\sigma^{-1} = \gamma_k^{-1} \cdots \gamma_1^{-1}$.

Definition 5.2.4. A 2-cycle is called a *transposition*. An element $\sigma \in S_n$ is called an *even permutation* if σ can be written as a product of an even number of transpositions. Otherwise σ is called an *odd permutation*. [1]

Remark 5.2.5.

(1) Notice that

$$\begin{aligned}(\text{even permutation}) \cdot (\text{even permutation}) &= \text{even}\\ (\text{odd permutation}) \cdot (\text{even permutation}) &= \text{odd}\\ (\text{even permutation}) \cdot (\text{odd permutation}) &= \text{odd}\\ (\text{odd permutation}) \cdot (\text{odd permutation}) &= \text{even}\end{aligned}$$

[1] Note that the definition of even and odd permutations is in fact well defined. In other words, one can show that if $\sigma \in S_n$ can be written as a product of an even (resp. odd) number of transpositions, then *any* presentation of σ has an even (resp. odd) length.

(2) Let $n \geq 2$. Define

$$\begin{aligned} \epsilon : S_n &\longrightarrow \{\pm 1\} \\ \sigma &\longmapsto \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases} . \end{aligned}$$

Then $\epsilon(\sigma)$ is called the *sign of* σ for $\sigma \in S_n$.

Definition 5.2.6.

(1) Let G be an arbitrary group. A subset $S \subseteq G$ is called a set of *generators* of G if every element of G can be written as a finite product of elements of S and their inverses. In this case, we write $G = \langle S \rangle$.
(2) If $G = \langle S \rangle$ and moreover, there exists a collection of *relations* $\{R_\alpha\}_{\alpha \in \Lambda}$ in the elements of S and their inverses, for some index set Λ, such that every relation among the elements of S and their inverses can be obtained from these, we get a *presentation* of G:

$$\langle S \mid R_\alpha, \alpha \in \Lambda \rangle .$$

Definition 5.2.7. A group H is called *cyclic* if there exists $x \in H$ with $H = \langle x \rangle$. Note that $\langle x \rangle$ is in particular abelian.

Lemma 5.2.8. *Let $H = \langle x \rangle$ be a cyclic group. Then $|H| = o(x)$. More precisely,*

(a) If $o(x) = \infty$, then $x^i \neq x^j$ for $i \neq j \in \mathbb{Z}$.
(b) If $o(x) = n < \infty$, then $1_G, x, x^2, \ldots, x^{n-1}$ are pairwise distinct and $H = \{1_G, x, x^2, \ldots, x^{n-1}\}$.

Remark 5.2.9. For a cyclic group H, we have:

- $x^0 = 1_G$ where $0 = 0_{\mathbb{Z}}$.
- $(x^n)^m = x^{nm}$ for all $n, m \in \mathbb{Z}$.
- $x^n x^m = x^{n+m}$ for all $n, m \in \mathbb{Z}$.

If $H = (H, +)$ is written additively, then $\langle x \rangle = \{nx \mid n \in \mathbb{Z}\}$, and

- $0x = 0_G$ where $0 = 0_{\mathbb{Z}}$.
- $m(nx) = (nm)x$ for all $n, m \in \mathbb{Z}$.
- $nx + mx = (n + m)x$ for all $n, m \in \mathbb{Z}$.

Theorem 5.2.10. *Every infinite cyclic group is isomorphic to $(\mathbb{Z}, +)$. Every finite cyclic group of order n is isomorphic to $(\mathbb{Z}/n, +)$.*

Definition 5.2.11. The group of symmetries of a regular n-gon is called the *dihedral group of order* $2n$ and is denoted by D_{2n}. We have

$$D_{2n} = \{s^i r^j \mid i \in 0, 1, j \in 0, 1, \ldots, n-1\}$$

such that $o(s) = 2$ and $o(r) = n$. We have $rs = sr^{-1}$. Note that we also have

$$D_{2n} = \left\langle r, s \mid r^n = 1s^2, rs = sr^{-1} \right\rangle .$$

Be careful: Usually if $g \in G$ where G is a group, then $g^n = 1$ only means that $o(g)$ *divides* n. So the notation $r^n = 1 = s^2$ is sloppy, but we understand it to mean that r has order n and s has order 2 in this context.

When $n = 4$, we can write the elements of D_8 as elements of S_4:

$$\begin{aligned} \mathrm{id} &= (1) & s &= (1,3) \\ r &= (1,2,3,4) & sr &= (1,2)(3,4) \\ r^2 &= (1,3)(2,4) & sr^2 &= (2,4) \\ r^3 &= (1,4,3,2) & sr^3 &= (1,4)(2,3) \end{aligned}$$

5.3. Group Homomorphisms

Definition 5.3.1.

(1) Let $(G, *)$ and $(H, \circ)$ be groups. A map $\varphi : G \to H$ is a *group homomorphism* if for all $a, b \in G$,

$$\varphi(a * b) = \varphi(a) \circ \varphi(b).$$

(2) The *image of* φ is defined as

$$\mathrm{Im}(\varphi) := \varphi(G).$$

This is a group under the operation of H.

(3) The *kernel of* φ is defined as

$$\mathrm{Ker}(\varphi) := \{a \in G \mid \varphi(a) = 1_H\}.$$

This is a group under the operation of G.

(4) A group homomorphism $\varphi : G \to H$ is called a *group isomorphism* if φ is bijective. We say G and H are *isomorphic*, written $G \cong H$.

Remark 5.3.2. If $\varphi : G \to H$ is a group homomorphism, then for all $a \in G$ and for all $n \in \mathbb{Z}$, $\varphi(a^m) = \varphi(a)^m$. In particular, $o(\varphi(a))$ divides $o(a)$ for all $a \in G$.

Lemma 5.3.3. *Let $\varphi : G \to H$ be a group homomorphism.*

(a) φ is surjective if and only if $\mathrm{Im}(\varphi) = H$.

(b) φ is injective if and only if $\mathrm{Ker}(\varphi) = \{1_G\}$.

Example 5.3.4.

(1) $\varphi : \mathbb{Z} \to \mathbb{Z}/n$ where $n \in \mathbb{Z}^+$, defined by $a \mapsto \overline{a}$ is a surjective group homomorphism with $\mathrm{Ker}(\varphi) = n\mathbb{Z}$.

(2) $\varphi : \mathrm{GL}_n(F) \to F^\times$, where $n \in \mathbb{Z}^+, F$ is a field, defined by $A \mapsto \mathrm{Det}(A)$ is a surjective group homomorphism with $\mathrm{Ker}(\varphi) = \mathrm{SL}_n(F)$.

(3) Let e be the base of the natural logarithm. The map

$$\begin{aligned} \exp : (\mathbb{R}, +) &\longrightarrow (\mathbb{R}^+, \cdot) \\ x &\longmapsto e^x \end{aligned}$$

is a group homomorphism. Moreover, exp is bijective since it has an inverse function, $\log_e$. In this example, the group elements are different, and the operations are different, yet the two groups are isomorphic, that is to say that, as groups, they have identical structures.

Example 5.3.5.

(1) $G \cong G$, with isomorphism Id_G.

(2) If $\varphi : G \to H$ is an injective group homomorphism, then $G \cong \mathrm{Im}(\varphi)$.

(3) Let Ω be a size of order $n \in \mathbb{Z}^+$. Then $S_\Omega \cong S_n$.

(4) If $G \xrightarrow{\cong} H$ is an isomorphism, then $|G| = |H|, o(a) = o(\varphi(a))$ for all $a \in G$. Also, G is abelian if and only if H is abelian.
(5) All groups, up to isomorphism, of order ≤ 7:

$$\{1\},\ \mathbb{Z}/2,\ \mathbb{Z}/3,\ \mathbb{Z}/2 \times \mathbb{Z}/2,\ \mathbb{Z}/4,\ \mathbb{Z}/5,\ \mathbb{Z}/6,\ S_3 \cong D_6,\ \mathbb{Z}/7.$$

5.4. Subgroups and Normal Subgroups

Definition 5.4.1. Let G be a group, let H be a subset of G. Then H is called a *subgroup* of G, written $H \leq G$, if H is a group using the operation in G.

Proposition 5.4.2 (Subgroup Criterion)**.** *Let G be a group, let $H \subseteq G$. Then $H \leq G$ if and only if*

(a) $H \neq \varnothing$, and
(b) For all $x, y \in H$, $xy^{-1} \in H$.

Example 5.4.3.

(1) If G is an arbitrary group, it always has the subgroups $\{1_G, G\}$.
(2) If $\varphi : G \to H$ is a group homomorphism, then $\mathrm{Im}(\varphi) \leq H$ and $\mathrm{Ker}(\varphi) \leq G$.
(3) $\mathrm{SL}_n(F) \leq \mathrm{GL}_n(F)$.
(4) $S_n \leq S_m$ for all $n \leq m$.
(5) Let G be a group, S some subset of G. Define

$$H = \{1_G\} \cup \{x_1 \cdots x_n \mid n \in \mathbb{Z}^+, x_i \in S \text{ or } x_i^{-1} \in S\ \forall 1 \leq i \leq n\}.^2$$

Then H is a subgroup of G, since $H \neq \varnothing$ and if $x = x_1 \cdots x_n \in H$ and $y = y_1 \cdots y_m \in H$ (where $x_i \in S$ or $x_i^{-1} \in S$, $y_j \in S$ or $y_j^{-1} \in S$ for all $1 \leq i \leq n,\ 1 \leq j \leq m$), then $xy^{-1} = x_1 \cdots x_n y_m^{-1} \cdots y_1^{-1} \in H$.

Then H is the smallest subgroup of G containing S and is called the *subgroup of G generated by S*, written $H = \langle S \rangle$.

(6) Let G be an arbitrary group. Let $Z(G) := \{a \in G \mid ag = ga\ \forall g \in G\}$. This is a subgroup of G (check the subgroup criterion) called the *center of G*.

Definition 5.4.4. Let G be a group, let $x, g \in G$, let $H \subseteq G$. Then gxg^{-1} is called the *conjugate of x by g*. The set

$$gHg^{-1} := \{ghg^{-1} \mid h \in H\}$$

is called the *conjugate of H by g*.

The set

$$N_G(H) := \{g \in g \mid gHg^{-1} = H\}$$

is called the *normalizer of H in G*.

Remark 5.4.5.

(1) When $H \leq G$, we have $N_G(H) \leq G$. Indeed, $N_G(H) \neq \varnothing$ since $1_G \in N_G(H)$, and if $a, b \in N_G(H)$,

$$(ab^{-1})H(ab^{-1})^{-1} = a(b^{-1}Hb)a^{-1} = aHa^{-1} = H,$$

where the second equality follows since $bHb^{-1} = H \iff H = b^{-1}Hb$.
(2) If $H \leq G$, then $H \leq N_G(H)$.

[2] We only need $\{1_G\}$ when $S \neq \varnothing$

Example 5.4.6. Let $G = S_3$ and $H = \langle(12)\rangle$. Then $H \leq N_H(G)$. By Lagrange's Theorem (5.5.5), either $N_G(H) = H$ or $N_G(H) = G$.

$$(13)(12)(13)^{-1} = (23) \notin H,$$

so $(13) \notin N_G(H)$, and hence $N_G(H) = H$.

Definition 5.4.7. Let G be a group, let $N \leq G$. We say N is a *normal subgroup of* G, written $N \trianglelefteq G$, if for all $g \in G$, $gNg^{-1} = N$.

Theorem 5.4.8. *Let G be a group, $N \leq G$. The following are equivalent:*

(a) $N \trianglelefteq G$.
(b) $N_G(N) = G$.
(c) For all $g \in G$, for all $n \in N$, $gng^{-1} \in N$.
(d) For all $g \in G$, $gN = Ng$.

Proof.

- (a) $\iff$ (b) By definition of $N_G(N)$.
- (a) $\iff$ (c) Let $g \in G$ and $n \in N$. Then $gng^{-1} \in gNg^{-1} \overset{(a)}{=} N$.
- (c) $\iff$ (d) Let $g \in G$ and $n \in N$.

$$gn = (gng^{-1})g \in Ng \text{ by (c)} \implies gN \leq Ng.$$
$$ng = g(g^{-1}ng) \in gN \text{ by (c)} \implies Ng \leq gN.$$

- (d) $\iff$ (a) Let $g \in G$. Then $gNg^{-1} \overset{(d)}{=} (Ng)g^{-1} = N$.

□

Example 5.4.9.

(1) If G is abelian, then $H \leq G$ is normal.
(2) For any group G, the subgroups $\{1_G\}$ and G are always normal. Also $Z(G) \trianglelefteq G$.
(3) If $H \leq G$ and $[G : H] = 2$, then $H \trianglelefteq G$.

Proof. We know that for all $g \in H$, $gH = H$. Hence (since the left cosets of H in G partition G) the only other left coset is $G - H = gH$ for all $g \in G - H$. □

Lemma 5.4.10. *Let $\varphi : G \to H$ be a group homomorphism. Let $K = \mathrm{Ker}(\varphi)$. Then $K \trianglelefteq G$. In fact, we show for all $h \in \mathrm{Im}(\varphi)$, say $h = \varphi(g)$ for some $g \in G$, we have*

$$\varphi^{-1}(h) = \{x \in G \mid \varphi(x) = h\} = gK = Kg.$$

Proof. Let $x \in G$. Then

$$\begin{aligned} x \in \varphi^{-1}(h) &\iff \varphi(x) = h = \varphi(g) \\ &\iff \varphi(g^{-1}x) = 1_H = \varphi(xg^{-1}) \\ &\iff g^{-1}x \in K \text{ and } xg^{-1} \in K. \\ &\iff x \in gK \text{ and } x \in Kg. \end{aligned}$$

□

Proposition 5.4.11. *Let G be a group. Then*

$$\{N \mid N \trianglelefteq G\} = \{\mathrm{Ker}(\varphi) \mid \varphi : G \to H \text{ a group hom. for some group } H.\}.$$

PROOF. " $\supseteq$ " : We proved this in Lemma 5.4.10.
" $\subseteq$ " : Let $N \leq G$. Define

$$\pi : G \longrightarrow G/N$$
$$g \longmapsto gN.$$

Then π is a group homomorphism since for all $a, b \in G$,

$$\pi(ab) = abN = (aN)(bN) = \pi(a)\pi(b).$$

We have

$$\mathrm{Ker}(\pi) = \{g \in G \mid gN = N\} = \{g \in G \mid g \in N\} = N.$$

□

5.5. Cosets and Lagrange's Theorem

Definition 5.5.1. Let G be a group, let $H \leq G$, let $g \in G$. Define

$$gH := \{gh \mid h \in H\},$$

called the *left coset of H in G with representative g*.[3]

Lemma 5.5.2. *Let G be a group, let $H \leq G$. The left cosets of H in G form a partition of G. Moreover, $aH = bH \iff a \in bH \iff b^{-1}a \in H$.*

PROOF. (For left cosets) We have

$$G = \bigcup_{g \in G} gH.$$

Consider aH and bH which have a nontrivial intersection, i.e., there exists $c \in aH \cap bH$. We need to show $aH = bH$.

There exists $h, h' \in H$ such that $c = ah = bh'$. Then

$$a = bh'h^{-1} \implies a \in bH \implies aH \subseteq bH,$$

and similarly, $bH \subseteq aH$. Hence $aH = bH$. This also shows $aH = bH \iff a \in bH$.

Now,

$$\begin{aligned} a \in bH &\iff \text{there exists } h \in H \text{ with } a = bh. \\ &\iff \text{there exists } h \in H \text{ with } b^{-1}a = h. \\ &\iff ba^{-1} \in H. \end{aligned}$$

□

Lemma 5.5.3. *Let G be a group, let $H \leq G$, let $g \in G$. Then $f : H \to gH$, $f(h) = gh$ is bijective.*

PROOF. (For left cosets) f is surjective by definition of coset. If $gh = gh'$, then $h = h'$, so f is injective. □

Definition 5.5.4. Let G be a group, let $H \leq G$. The number of distinct left cosets of G in G is called the *index of H in G*, denoted by $[G : H]$ or $|G : H|$.

Theorem 5.5.5 (Lagrange's Theorem)**.** *Let G be a finite group, let $H \leq G$. Then $|H|$ divides $|G|$, and*

$$[G : H] = \frac{|G|}{|H|}.$$

[3] the *right coset of H in G with representative g* is defined $Hg := \{hg \mid h \in H\}$

PROOF.

□

Corollary 5.5.6. *Let G be a finite group.*

(a) For all $x \in G$, $o(x) \big| |G|$
(b) If $|G| = p$, where p is prime, then G is cyclic.

PROOF.

(a) Let $H = \langle x \rangle$. Then $o(x) = |H|$ divides $|G|$ by Theorem 5.5.5.
(b) If $|G| = p$ is prime, then $|G| > 1$, so there exists $x \in G, x \neq 1_G$. By Theorem 5.5.5, $|\langle x \rangle| \,\big|\, |G| = p$, and hence $|\langle x \rangle| = p$. So $G = \langle p \rangle$.

□

5.6. Quotient Groups

Definition 5.6.1. Let G be a group, let $N \trianglelefteq G$. Define $G/N = \{gN \mid g \in G\}$, and for all $a, b \in G$,

$$(aN) \cdot (bN) = (ab)N.$$

Theorem 5.6.2. *G/N with this operation is a group, called the* quotient group of G modulo N.

PROOF. We first need to show that the operation is well defined. Suppose $aN = a_1N, bN = b_1N$. Then $a_1^{-1}a, b_1^{-1}b \in N$. We need to show $abN = a_1b_1N$, i.e. $(a_1b_1)^{-1}ab \in N$. This is true since $a_1^{-1}ab \in Nb = bN$ so there exists $n \in N$ with $a_1^{-1}ab = bn$, and so

$$(a_1b_1)^{-1}(ab) = b_1^{-1}a_1^{-1}a = b_1^{-1}bn \in N.$$

We also need to check the group axioms. The identity element in G/N is $1_GN = N$ and the inverse of aN in G/N is $a^{-1}N$. Showing the group axioms hold is left as an exercise. □

Remark 5.6.3. $|G/N| = [G : N]$. If G is finite, by Lagrange (5.5.5), $|G/N| = \frac{\#G}{\#H}$.

Example 5.6.4.

(1) If G is an arbitrary group, $\{1_G\}, G \trianglelefteq G$. Then $G/\{1_G\} \cong G$ and $G/G \cong \{1_G\}$.
(2) If G is abelian, all $H \leq G$ are normal. For example, when $G = \mathbb{Z}$, $H = \langle n \rangle = n\mathbb{Z}$, we have $G/N = \mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/n$.
(3) Let G be an arbitrary group and let $Z(G)$ be the center of G. If $H \leq \mathbb{Z}(G)$ then $H \trianglelefteq G$ since for all $h \in H$, $ghg^{-1} = g \in H$ since $h \in H \leq Z(G)$.

    For example, let $G = D_8 = \langle r, s \mid r^4 = 1 = s^2, srs = r^{-1} \rangle$. Then $Z := Z(G) = \langle r^2 \rangle = \{1, r^2\}$. Then

$$\begin{aligned} \overline{G} := G/Z &= \{ \underset{\underset{\{1,r^2\}}{\|}}{Z} , \underset{\underset{\{r,r^3\}}{\|}}{rZ} , \underset{\underset{\{s,sr^2\}}{\|}}{sZ} , \underset{\underset{\{sr,sr^3\}}{\|}}{srZ} , \} \\ &= \{\overline{1}, \overline{r}, \overline{s}, \overline{sr}\}. \end{aligned}$$

Now,

$$\begin{aligned} \overline{r}^2 &= (rZ)^2 = r^2Z = Z = \overline{1} \\ \overline{s}^2 &= (sZ)^2 = s^2Z = Z = \overline{1} \\ \overline{sr}^2 &= (srZ)^2 = (sr)^2Z = Z = \overline{1} \end{aligned}$$

Hence $\overline{G} \cong \mathbb{Z}/2 \times \mathbb{Z}/2$.

Definition 5.6.5. Let G be a group, let $N \trianglelefteq G$. We call $\pi : G \to G/N$, defined by $\pi(g) = gN$, the *natural projection from G onto G/N*. If $\overline{H} \leq G/N$ then the *full preimage of $\overline{H}$ in G* is defined to be

$$\pi^{-1}\left(\overline{H}\right) := \{g \in G \mid gN \in \overline{H}\}.$$

Note. $\overline{H} = \pi^{-1}\left(\overline{H}\right)/N$ and $N \leq \pi\left(\overline{H}\right)$.

Remark 5.6.6. Recall that Lagrange (5.5.5) says: If G is a finite group, then

$$H \leq G \implies \#H \big| \#G.$$

In general, the converse is not true. For example, if $G = A_4$, then $\#G = 12$, but G has no subgroup of order 6 (check). So, what is still true? The following two theorems provide information about prime divisors (and prime power divisors) of a group's order:

Theorem 5.6.7 (Cauchy's Theorem)**.** *If G is a finite group, and p is a prime with $p \big| \#G$, then there exists $x \in G$ with $o(x) = p$.*

Theorem 5.6.8 (Sylow's Theorem)**.** *If G is a finite group of order $p^a \cdot m$, where p is a prime, $a \in \mathbb{Z}^+$, and $p \nmid m$, then there exists $P \leq G$ with $\#P = p^a$. P is called a* Sylow p-subgroup *of G.*

5.7. Isomorphism Theorems

Theorem 5.7.1 (First Isomorphism Theorem)**.** *Let $\varphi : G \to H$ be a group homomorphism. Then* $\operatorname{Ker}(\varphi) \trianglelefteq G$ *and* $G/\operatorname{Ker}(\varphi) \cong \operatorname{Im}(\varphi)$.

PROOF. Let $K := \text{Ker}(\varphi)$. We showed in Lemma 5.4.10 that $K \trianglelefteq G$. Define

$$\begin{aligned} \psi : G/K &\longrightarrow \text{Im}(\varphi) \\ gK &\longmapsto \varphi(g). \end{aligned}$$

- ψ is well-defined: Let $gK = g'K$. Then there exists $k \in K$ with $g = g'k$ and
$$\varphi(g) = \varphi(g'k) = \varphi(g)\varphi(k) = \varphi(g'),$$
so $\psi(gK) = \psi(gK)$.
- ψ is a group homomorphism since φ is a group homomorphism.
- ψ is surjective by the definition of $\text{Im}(\varphi)$.
- ψ is injective since
$$\begin{aligned} \text{Ker}(\psi) &= \{gK \in G/K \mid \varphi(g) = 1_H\} \\ &= \{gK \in G/K \mid g \in K\} \\ &= \{K\} \\ &= \{1_{G/K}\}. \end{aligned}$$

□

Remark 5.7.2. There is a connection between the First Isomorphism Theorem for groups and the Rank-Nullity Theorem. In particular, a linear transformation $\varphi : V \to W$ is a group homomorphism that also commutes with scalars. The First Isomorphism Theorem tells us that we always have an isomorphism **as groups**

$$V/(\text{Ker}(\varphi)) \cong \text{Im}(\varphi).$$

For vector spaces, say $V_1 \supseteq V_2$, the quotient space has dimension

$$\dim(V_1/V_2) = \dim(V_1) - \dim(V_2)$$

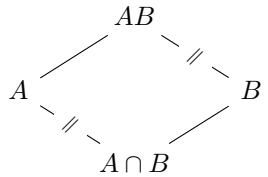
The Rank-Nullity Theorem takes the First Isomorphism Theorem a step further by telling us that

$$\dim V - \dim(\text{Ker}(\varphi)) = \dim(\text{Im}(\varphi)).$$

Moreover, note that there is a "First Isomorphism Theorem" for vector spaces (and modules!), which gives us the vector space isomorphism $V/(\text{Ker}(\varphi)) \cong \text{Im}(\varphi)$, from which the Rank-Nullity Theorem follows immediately.

Theorem 5.7.3 (Second (or Diamond) Isomorphism Theorem)**.** *Let G be a group, let $A, B \leq G$ with $A \leq N_G(B)$. Then $AB \leq G$, $B \trianglelefteq AB$, $A \cap B \trianglelefteq A$, and*

$$AB/B \cong A/(A \cap B)$$



Theorem 5.7.4 (Third (or Cancellation) Isomorphism Theorem)**.** *Let G be a group, let $H, K \trianglelefteq G$ with $H \leq K$. Then $K/H \trianglelefteq G/H$ and*

$$(G/H)/(K/H) \cong G/K.$$

Theorem 5.7.5 (Fourth (or Lattice) Isomorphism Theorem)**.** *Let G be a group, let $N \trianglelefteq G$. Define*

$$\mathcal{G} = \{H \mid N \leq H \leq G\} \quad \text{and} \quad \overline{\mathcal{G}} = \{\overline{H} \mid \overline{H} \leq G/N\}.$$

Then the map

$$\begin{aligned} f : \mathcal{G} &\longrightarrow \overline{\mathcal{G}} \\ H &\longmapsto H/N \end{aligned}$$

is a bijection. Moreover, define $\overline{G} := G/N$. *If* $A, B \in \mathcal{G}$ *define* $\overline{A} := A/N, \overline{B} := B/N$. *We have:*

(1) $A \leq B \iff \overline{A} \leq \overline{B}$.
(2) If $A \leq B$ *then* $[B : A] = [\overline{B} : \overline{A}]$.
(3) $\overline{\langle A, B\rangle} = \langle \overline{A}, \overline{B}\rangle$.
(4) $\overline{A \cap B} = \overline{A} \cap \overline{B}$.
(5) $A \trianglelefteq G \iff \overline{A} \trianglelefteq \overline{G}$.

Important: If G is a group, $N \trianglelefteq G$, $\pi : G \to G/N$ the natural projection, then

- If $\overline{H} \leq G/N$ then $\overline{H} = \pi(\overline{H})/N$.
- If $H \leq G$ then $\pi(H) = \{hN \mid h \in H\} = HN/N \cong H/(H \cap N)$.

5.8. Group Actions

Definition 5.8.1.

(1) Let G be a group, let $M \neq \varnothing$ be a set. A *(left) group action of* G *on* M is a map

$$\begin{aligned} \boldsymbol{.} : G \times M &\longrightarrow M \\ (g, m) &\longmapsto g\boldsymbol{.}m \end{aligned}$$

such that for all $g, h \in G$ and for all $m \in M$:
(a) $g\boldsymbol{.}(h\boldsymbol{.}m = (gh)\boldsymbol{.}m$.
(b) $1_G\boldsymbol{.}m = m$.
A *(right) group action of* G *on* M is a map

$$\begin{aligned} \boldsymbol{.} : G \times M &\longrightarrow M \\ (g, m) &\longmapsto m\boldsymbol{.}g \end{aligned}$$

such that for all $g, h \in G$ and for all $m \in M$:
(a) $(m\boldsymbol{.}g)\boldsymbol{.}h = m\boldsymbol{.}(gh)$.
(b) $m\boldsymbol{.}1_G = m$.

(2) A *permutation representation of* G *on* M is a group homomorphism $\varphi : G \to S_M$, where S_M is the symmetric group on M.

Proposition 5.8.2. *Let* G *be a group, let* $M \neq \varnothing$ *be a set. There is a bijection between the left group actions of* G *on* M *and the permutation representations of* G *on* M.

Notation. Let $\boldsymbol{.} : G \times M \to M$ be a left group action of G on M. Then this action *induces* a permutation representation

$$\begin{aligned} \varphi : G &\longrightarrow S_M \\ g &\longmapsto \sigma_g, \quad \text{where} \quad \sigma_g : M \to M \\ & \qquad\qquad\qquad\quad m \mapsto g\boldsymbol{.}m \end{aligned}$$

φ is called the permutation representation *associated* to the group action.

Definition 5.8.3. The *kernel of a group action* is defined as

$$\{g \in G \mid g.m = m \ \forall \ m \in M\} = \{g \in G \mid \sigma_G = \mathrm{id}_M\} = \mathrm{Ker}(\varphi).$$

We say the action if *faithful* is the kernel is $\{1_G\}$.

Example 5.8.4. Let G be a group, let $M \neq \varnothing$ a set.

(1) The *trivial group action*: For all $g \in G$ for all $m \in M$, $g.m = m$. So the kernel of the trivial action is G.
(2) Let $M = G$.
   (a) G acts on itself by left multiplication:
   $$G \times M \longrightarrow M$$
   $$(g, m) \longmapsto gm$$
   This is called the *left regular action* of G on itself. The kernel of this action is $\{1_G\}$.
   (b) G acts on itself by conjugation:
   $$G \times M \longrightarrow M$$
   $$(g, m) \longmapsto gmg^{-1}.$$
   The kernel is
   $$\{g \in G \mid gmg^{-1} = m \forall \ m \in G\} = Z(G).$$
(3) $\mathrm{GL}_n(F)$ acts on
$$F^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \;\middle|\; a_i \in F \right\}$$
by left multiplication, with kernel $\{I_n\}$.
(4) For all $n \in \mathbb{Z}^+$, D_{2n} and S_n act on $\{1, 2, \ldots, n\}$ with kernel $\{1\}$.

5.8.1. Stabilizers, Normalizers, and Centralizers.

Definition 5.8.5. Given a group action of G on M, and $m \in M$, define

$$G_m := \{g \in G \mid g.m = m\}$$

called the *stabilizer of m in G*.

Note 5.8.6. The kernel of a group action is

$$\{g \in G \mid g.m = m \ \forall \ m \in M\} = \bigcap_{m \in M} G_m,$$

the intersection of all the stabilizers.

Claim 5.8.7. For all $m \in M$, $G_m \leq G$.

Proof. Use the subgroup criterion: $G_m \neq \varnothing$ since $1_G \in G_m$. Let $g, h \in G_m$. Then $g.m = m = h.m$, and so

$$(gh^{-1}).m = (gh^{-1}).(h.m) = (gh^{-1}h).m = g.m = m.$$

Hence $gh^{-1} \in G_m$. □

Definition 5.8.8. Let G be a group.

(1) Let $M = \mathcal{P}(G) = \{A \mid A \subseteq G\}$. G acts on $\mathcal{P}(G)$ by conjugation: For all $A \in \mathcal{P}(G)$, for all $g \in G$, $g.A := gAg^{-1}$. The stabilizer of A under this action
$$G_A := \{g \in G \mid gAg^{-1} = A\}$$
is called the *normalizer of A in G*, denoted $N_G(A)$. Since $N_G(A) = G_A$, $N_G(A) \leq G$.

(2) Let $A \in \mathcal{P}(G)$, let $N_G(A)$ act on A by conjugation: For all $a \in A$, for all $g \in N_G(A)$, $g.a := gag^{-1}$. The kernel of this action is
$$\{g \in N_G(A) \mid gag^{-1} = a \ \forall \ a \in A\}$$
is called the *centralizer of A in G*, denoted by $C_G(A)$.

Remark 5.8.9. Note that $C_G(A) \leq N_G(A) \leq G$ and $Z(G) = C_G(G)$. If $A \leq G$, then $Z(A) = C_G(G) \cap A$.

Example 5.8.10.

(1) If G is abelian, then for all $A \subseteq G$, $N_G(A) = G = C_G(A)$.

(2) $G = D_8 = \langle r, s \mid r^4 = s^2 = 1, srs = r^{-1}\rangle$, $A = \langle r\rangle$. Since A is abelian, $A \leq C_G(A)$. We have $srs^{-1} = srs = r^{-1} \neq r$ and so $s \notin C_G(A)$. Since $A \leq C_G(A)$ and $[G : A] = 2$, $C_G(A) = A$ or G. Since $s \notin C_G(A)$, $C_A(A) = A$. Since $A \trianglelefteq G$ (as a subgroup of index 2), $N_G(A) = G$.

Definition 5.8.11. Let G be a group acting on a set $M \neq \varnothing$. We define a relation $\sim$ on M by: If $m_1, m_2 \in M$ then $m_1 \sim m_2$ if there exists $g \in G$ with $m_1 = g.m_2$. One can show that $\sim$ is an equivalence relation.

5.8.2. Orbits.

Definition 5.8.12. If G acts on M and $\sim$ is the corresponding equivalence relation , then the equivalence class containing $m \in M$ is called the *G-orbit of M*, denoted by $G.m = \{g.m \mid g \in G\}$. (Other notation: $\mathcal{O}_m$).

If there is only one orbit, we call the action *transitive*, i.e. for all $m_1, m_2 \in M$, there exists $g \in G$ with $m_1 = g.m_2$, i.e. $\mathcal{O}_m = G.m = M$ for all $m \in M$.

Careful! We have $G_m \leq G$ (stabilizer subgroup), but $G.m \subseteq M$ (G-orbit).

Lemma 5.8.13 (Orbit Lemma)**.** *Given a (left) group action of G on M, we have for all $m \in M$*
$$|G.m| = [G : G_m].$$

Proof. Let $\mathcal{C} = \{gG_m \mid g \in G\}$ be the set of left cosets of G_m in G. Define
$$\begin{aligned} f : G.m &\longrightarrow \mathcal{C} \\ g.m &\longmapsto gG_m. \end{aligned}$$
f is well defined and injective: Let $g, h \in G$. Then
$$\begin{aligned} g.m = h.m &\iff (h^{-1}g).m = m \\ &\iff h^{-1}g \in G_m \\ &\iff gG_m = hG_m. \end{aligned}$$
f is surjective by the definition of $\mathcal{C}$. So $|G.m| = |\mathcal{C}| = [G : G_m]$. □

Example 5.8.14.

(1) $G = S_n$ acts on $M = \{1, \ldots, n\}$ by: For all $\sigma \in S_n$, for all $i \in \{1, \ldots, n\}$, $\sigma.i := \sigma(i)$. This is a left group action (exercise). The kernel is

$$\{\sigma \in S_n \mid \sigma(i) = i \ \forall\ 1 \leq i \leq n\} = \{\mathrm{id}\},$$

so this is a faithful action.

Stabilizers: Let $i \in \{1, \ldots, n\}$. Then

$$G_i = \{\sigma \in S_n \mid \sigma(i) = i\} = S_{\{1,2,\ldots,i-1,i+1,\ldots,n\}} \cong S_{n-1}.$$

Orbits: This action is transitive since for all $i \neq j$ in $\{1, \ldots, n\}$ the transposition $\tau = (i, j) \in S_n$ gives $\tau.i = \tau(i) = j$.

Note that by the Orbit Lemma (5.8.13), for all $i \in \{1, \ldots, n\}$,

$$|G.i| = [G : G_i] = \frac{\#G}{\#G_i} = \frac{n!}{(n-1)!} = n.$$

Hence we see each orbit has length n, so there can only be one orbit.

(2) For $n \geq 2$, let $\sigma \in S_n$. Then $G := \langle \sigma \rangle$ acts on $\{1, \ldots, n\}$ by

$$\sigma^a.i = \sigma^a(i) \quad \forall\ i \in \{1, \ldots, n\},\ \forall\ a \in \mathbb{Z}.$$

The G-orbits in $\{1, \ldots, n\}$ give the cycles in the cycle decomposition of σ:

$$G.i = \{i, \sigma(i), \sigma^2(i), \sigma^2(i), \ldots, \sigma^{r-1}(i)\}$$

where r is the smallest positive integer with $\sigma^r \in G_i$ (i.e $\sigma^r(i) = i$). Since the G-orbits partition $\{1, \ldots, n\}$, the cycle decomposition is unique up to permutation of the cycles.

5.9. Sylow's Theorems

Definition 5.9.1. Let p be a prime number.

(a) A group G is called a *p-group* if G is not trivial and every element in G has order a power of p. If G is finite, then G is a p-group if and only if $|G| = p^\alpha$ for some $\alpha \in \mathbb{Z}^+$.

(b) If G is a finite group, $|G| = p^\alpha \cdot m$, where $\alpha \in \mathbb{Z}^+ \cup \{0\}$ and $p \nmid m$, then a subgroup $P \leq G$ is called a *Sylow p-subgroup* of G if $|P| = p^\alpha$. We write

$$\mathrm{Syl}_p(G) = \{P \leq G \mid P \text{ is a Sylow } p - \text{subgroup of} G\},$$

and $n_p(G) := |\mathrm{Syl}_p(G)|$.

Theorem 5.9.2 (Sylow's Theorems)**.** *Let G be a finite group, $|G| = p^\alpha \cdot m$, where $\alpha \in \mathbb{Z}^+ \cup \{0\}$ and $p \nmid m$.*

(1) $\mathrm{Syl}_p(G) \neq \varnothing$.

(2) If $P \in \mathrm{Syl}_p(G)$ and $Q \leq G$ is a p-subgroup (i.e. $|Q| = p^\beta$ for $\beta \leq \alpha$) then there exists $g \in G$ with $Q \leq gPg^{-1}$. In particular, all Sylow p-subgroups of G are conjugate.

(3) $n_p(G) = [G : N_G(P)]$ for all $P \in \mathrm{Syl}_p(G)$. Moreover, $n_p(G) \equiv 1 \mod p$ and $n_p(G) \mid m$.

Corollary 5.9.3. *Let $P \in \mathrm{Syl}_p(G)$. Then $n_p(G) = 1 \iff P \trianglelefteq G$*

Example 5.9.4.

(1) Dummit & Foote, pg. 147, # 24: Let $|G| = 231$. Show that $Z(G)$ contains a Sylow 11-subgroup of G, and a Sylow 7-subgroup is normal in G.

PROOF. $|G| = 231 = 3 \cdot 7 \cdot 11$

- $n_3(G) \equiv 1 \mod 3$ and $n_3(G) \mid 7 \cdot 11$, so $n_3(G) \in \{1, 7\}$.
- $n_7(G) \equiv 1 \mod 7$ and $n_7(G) \mid 3 \cdot 11$, so $n_7(G) \in \{1\}$, and hence a Sylow 7-subgroup of G is normal in G.
- $n_{11}(G) \equiv 1 \mod 11$ and $n_{11}(G) \mid 3 \cdot 7$, so $n_{11}(G) \in \{1\}$, and hence a Sylow 11-subgroup of G is normal in G.

Let $P \in \mathrm{Syl}_{11}(G)$. Then $P \trianglelefteq G$. Since $N_G(P)/C_G(P)$ is isomorphic to a subgroup of $\mathrm{Aut}(P)$ and $P \cong \mathbb{Z}/11$, then

$$G/C_G(P) = N_G(P)/C_G(P) \cong H \leq \mathrm{Aut}(P) \cong (\mathbb{Z}/11)^\times \cong \mathbb{Z}/10.$$

Since P is abelian, $P \leq C_G(P)$ and hence $|G/C_G(P)|$ divides $\frac{|G|}{|P|} = 21$. Since $\gcd(21, 10) = 1$, this means H is trivial. So $|G/C)G(P)| = 1$ i.e. $C_G(P) = G$, and hence $P \leq Z(G)$. □

(2) Dummit & Foote, pg. 147, # 13: $|G| = 56$. Show there exists a prime p dividing $|G|$ such that G has a normal Sylow p-subgroup.

PROOF. $|G| = 2^3 \cdot 7$.

- $n_2(G) \equiv 1 \mod 2$ and $n_2(G) \mid 7$, so $n_2(G) \in \{1, 7\}$.
- $n_7(G) \equiv 1 \mod 7$ and $n_7(G) \mid 2^3$, so $n_7(G) \in \{1, 8\}$.

If $n_7(G) = 1$, then a Sylow 7-subgroup is normal in G. Now suppose $n_7(G) = 8$. Write

$$\mathrm{Syl}_7(G) = \{P_1, \ldots, P_8\}.$$

Note that if $i \neq j$ in $\{1, \ldots, 8\}$ then $P_i \cap P_j = \{1_G\}$ since P_i, P_j are isomorphic to $\mathbb{Z}/7$ and $P_i \neq P_j$. So

$$\#(\text{elements of order 7 in } G) = 8 \cdot (7 - 1) = 48.$$

So we have $56 - 48 = 8$ elements in G left that do *not* have order 7. We have space for only one Sylow 2-subgroup of order 8, meaning $n_2(G) = 1$, and hence a Sylow 2-subgroup is normal. □

(3) If G is a group of order 30, show that G contains a subgroup of order 15.

PROOF. $|G| = 2 \cdot 3 \cdot 5$.

- $n_3(G) \equiv 1 \mod 3$ and $n_3(G) \mid 10$, so $n_3(G) \in \{1, 10\}$.
- $n_5(G) \equiv 1 \mod 5$ and $n_5(G) \mid 6$, so $n_5(G) \in \{1, 6\}$.

Let $P \in \mathrm{Syl}_3(G), Q \in \mathrm{Syl}_5(G)$. If $P \trianglelefteq\in G$ or $Q \trianglelefteq G$, then we know $PQ \leq G$. Moreover, $|P \cap Q| = 1$ since $P \cap Q \leq P$ and $P \cap Q \leq Q$, so $|P \cap Q|$ divides both 3 and 5. So

$$|PQ| = \frac{|P| \cdot |Q|}{|P \cap Q|} = \frac{3 \cdot 5}{1} = 15.$$

Therefore, it suffices to prove that either $n_3(G) = 1$ or $n_5(G) = 1$. By contradiction, assume $n_3(G) > 1$ and $n_5(G) > 1$, i.e. $n_3(G) = 10$ and $n_5(G) = 6$. Then

$$\mathrm{Syl}_3(G) = \{P_1, \ldots, P_{10}\} \quad \text{and} \quad \mathrm{Syl}_5(G) = \{Q_1, \ldots, Q_6\}.$$

For all $i \neq j$ in $\{1, \ldots, 10\}$, we have $P_i \cap P_j = \{1_G\}$, and for all $k \neq \ell$ in $\{1, \ldots, 6\}$, we have $Q_k \cap Q_\ell = \{1_G\}$. So there are $10 \cdot (3-1) = 20$ elements of G of order 3, and $6 \cdot (5-1) = 24$ elements in G of order 5. Hence there are more than $20 + 24 = 44$ elements in G, contradicting $|G| = 30$. So either $n_3(G) = 1$ or $n_5(G) = 1$. □

Bibliography

[1] David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley and Sons, 2003.

[2] P.A. Grillet. *Abstract Algebra*. Graduate Texts in Mathematics. Springer New York, 2007.

[3] J.R. Munkres. *Topology*. Featured Titles for Topology. Prentice Hall, Incorporated, 2000.

[4] S. Roman. *Advanced Linear Algebra*. Graduate texts in mathematics. Springer-Verlag, 1992.

[5] D. Stewart and S. Oliveira. *Building Proofs: A Practical Guide*. World Scientific Publishing Company, 2015.

[6] Gilbert Strang. *Linear algebra and its applications*. Thomson, Brooks/Cole, Belmont, CA, 2006.