Lecture Notes For

Mathematics 150A

Dr. Tyler J. Evans

Fall 2000

Contents

Chapter 1

Preliminaries

1.1 Lecture 1: Sets and mappings

We assume that the reader is familiar with the language and elementary properties of sets and functions between sets. For convenience and completeness, we briefly recall some of the material that will be prevalent in our everyday work.

For notation, we will denote sets with capital letters $A, B, \ldots$ and write $a \in A$ if a is an element of the set A. If all of the elements of a set B are also elements of a set A, then we say B is a **subset** of A and we write $B \subseteq A$. Every set has a subset consisting of no elements called the **empty set**, and we denote this subset by $\emptyset$. If A is any set, then both A and $\emptyset$ are **improper subsets** of A. If B is a non-empty subset of A but $B \neq A$, then we write $B \subset A$ and we say B is a **proper subset** of A. We will work repeatedly with familiar sets of numbers so that we fix the following notations once and for all:

$\mathbb{Z}$ denotes the set of integers.

$\mathbb{Q}$ denotes the set of rational numbers.

$\mathbb{R}$ denotes the set of real numbers.

$\mathbb{C}$ denotes the set of complex numbers.

The superscript $\times$ will always mean the subset of non-zero elements of these sets so that $\mathbb{Z}^\times$ is the set of non-zero integers, etc....

Two basic operations on sets are union and intersection. If A and B are sets, then the **union** of A

and B is the set

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

and the **intersection** is the set

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

A **function** or a **mapping** from a set A to a set B is a rule or correspondence written $f : A \to B$ that assigns to each element $a \in A$ a unique element $b = f(a) \in B$. If $C \subseteq A$ and $D \subseteq B$, then we have two important subsets

$$f(C) = \{f(c) : c \in C\} \subseteq B$$

and

$$f^{-1}(D) = \{a \in A : f(a) \in D\} \subseteq A$$

called the **image** of C and the **inverse image** of D respectively. A function $f : A \to B$ is called **one-to-one** or **injective** if $f(a) = f(a')$ implies $a = a'$. We say f is **onto** or **surjective** if $f(A) = B$. If f is both injective and surjective, we say f is bijective.

Example 1.1.1 *The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is neither injective nor surjective. Here the image $f(\mathbb{R}) = \{y \in R : y = x^2\} = \{y \in \mathbb{R} : y \geq 0\}$. The function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^3$ is a bijection.*

If A and B are two sets, then the **cross product** of A and B is the set

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

A **binary operation** on a set A is a function $\mu : A \times A \to A$. We usually write the image $\mu(a, a')$ simply as aa' and we refer to μ as a **multiplication on** A. Note that μ may or may not have anything to do with ordinary multiplication or real or complex numbers! Indeed, the set A may not even be a subset of $\mathbb{C}$.

We conclude this lecture by recounting the very important notion of an equivalence relation. Recall that a **relation** R on a set A is a subset of the cross product $A \times A$. We write $a \sim b$ if $(a, b) \in R$ and sometimes just refer to the relation $\sim$. A relation is called **reflexive** if $a \sim a$ for all $a \in A$;

symmetric if $a \sim b$ implies $b \sim a$; and **transitive** if $a \sim b$ and $b \sim a$ implies $a \sim c$. A relation R that is simultaneously reflexive, symmetric and transitive is called an **equivalence relation** on A. Recall a **partition** of a set A is a collection of disjoint subsets of A such that each element of A belongs to one and only one of the subsets. These subsets are called the **cells** of the partition. The relationship between equivalence relations and partitions is summarized in the following theorem.

Theorem 1.1.2 *If A is a non-empty set and $\sim$ is an equivalence relation on A, then the collection of subsets*

$$\overline{a} = \{x \in A : x \sim a\}$$

forms a partition of A. Moreover, given a partition of A, the relation $\sim$ defined on A by $a \sim b$ if and only if a and b belong to the same cell is an equivalence relation on A whose induced partition is the given one.

Proof. Exercise. ■

The cells $\overline{a}$ are called **equivalence classes** and $a \in \overline{a}$ is called a **representative** of the equivalence class $\overline{a}$. The set of all equivalence classes for a given equivalence relation $\sim$ is denoted $A/\sim$.

Example 1.1.3 *Fix a positive integer n and define a relation $\sim_n$ on $\mathbb{Z}$ by $a \sim_n b$ iff. $n|(a-b)$. Then $\sim_n$ is an equivalence relation on $\mathbb{Z}$ and we write $\mathbb{Z}_n = \mathbb{Z}/\sim_n$. Note that the equivalence class $\overline{a}$ for $a \in \mathbb{Z}$ consists of all integers $b \in \mathbb{Z}$ with the same remainder when divided by n. In particular we have*

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

1.2 Lecture 2: Matrices

In this lecture, we wish to briefly recall the essentials of matrix algebra. Matrices will serve as a primary source of examples in the (near) future so that is is imperative that each student have a firm grasp of the material presented here.

If m and n are positive integers, then the rectangular array

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

of mn numbers a_{ij} is called an $m \times n$ **matrix** with **entries** a_{ij}. We call the matrix A **rational**, **real** or **complex** if the entries come from $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ respectively. The indices i and j in a_{ij} are called the **row** and **column** indices respectively. We sometimes use the shorthand notation $A = [a_{ij}]$ to denote the matrix A. If $A = [a_{ij}]$ and $B = [b_{ij}]$ are two $m \times n$ matrices, then their **sum** is the $m \times n$ matrix $A + B$ whose (ij)-entry is $a_{ij} + b_{ij}$. If $\alpha \in \mathbb{C}$, then αA is the $m \times n$ matrix whose (ij)-entry is αa_{ij}. αA is called a **scalar multiple** of A.

We also recall here the definition of matrix multiplication. If If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the product $C = AB$ is defined to be the $m \times p$ matrix with (ij)-entry

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

That is, the entry c_{ij} is computed by taking the ordinary dot product of the i^{th} row of A with the j^{th} column of B. We note that in general, matrix multiplication is not commutative. That is, $AB \neq BA$.

For any positive integer n, we let I_n denote the $n \times n$ **identity matrix** whose (ij)-entry is δ_{ij} (the Kronecker delta). That is,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

We summarize the basic properties of matrix algebra in the following theorem. The proofs are left to the reader. In the statement of the theorem, we assume that the matrices involved are shaped appropriately so that all operations make sense.

Theorem 1.2.1 *If A, B and C are matrices and α, β are scalars, then:*

(1) $A + B = B + A$;
(2) $(A + B) + C = A + (B + C)$;
(3) $\alpha(A + B) = \alpha A + \alpha B$;
(4) $(\alpha + \beta)A = \alpha A + \beta A$;
(5) $A(B + C) = AB + AC$;
(6) $(A + B)C = AC + BC$;
(7) $(AB)C = A(BC)$;
(8) $AI_n = I_n A = A$;
(9) $\alpha(AB) = (\alpha A)B = A(\alpha B)$;
(10) $(\alpha\beta)A = \alpha(\beta A)$.

We conclude this lecture by recalling the very important notion of determinant. First, if $A = [a_{11}]$ is 1×1, define $\det A = a_{11}$. Next, if

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

then we define $\det A = a_{11}a_{22} - a_{12}a_{21}$. Finally, if A is an $n \times n$ matrix, we inductively define

$$\det A = a_{11}A_{11} - a_{12}A_{12} + \cdots \pm a_{1n}A_{1n}$$

where A_{ij} is the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and j^{th} column. If we denote the set of all $n \times n$ matrices (with real entries) by $\mathbb{R}^{n \times n}$, then we can think of the determinant as a function

$$\det : \mathbb{R}^{n \times n} \to \mathbb{R}.$$

We recall the following properties of the function det, all of which can be deduced from the previous definition.

Theorem 1.2.2 *If A and B are $n \times n$ matrices and α is a scalar, then:*

1. $\det(AB) = \det A \det B$.
2. $\det I_n = 1$.
3. $\det A = \det A^T$ *where* A^T *is the* **transpose** *of* A.
4. A *is invertible iff.* $\det A \neq 0$ *and* $\det(A^{-1}) = 1/\det A$.
5. *Interchanging two rows or columns of* A *multiplies* $\det A$ *by* -1.
6. *Multiplying a row of* A *by a scalar* α *multiplies* $\det A$ *by* α.
7. *Replacing a row with the row plus a scalar multiple of another row does not change* $\det A$ *at all.*

1.3 Lecture 3: Permutations and permutation matrices

The purpose of this lecture is to once again provide a rich source of examples for us once we begin studying our first topic—group theory. In addition to the matrix groups, another deep source of examples of groups are groups of permutations.

Definition 1.3.1 *If S is a set, a bijection $p : S \to S$ is called a* **permutation** *of S.*

Example 1.3.2 *If $S = \{1, 2, 3\}$, then the mapping $p : S \to S$ defined by*

$$\begin{aligned} 1 &\mapsto 2 \\ 2 &\mapsto 3 \\ 3 &\mapsto 1 \end{aligned}$$

is a permutation of S.

Proposition 1.3.3 *If $p, q : S \to S$ are permutations, then $pq = p \circ q$ is also a permutation of S. The identity map $1_S : S \to S$ defined by $1_S(s) = s$ for all $s \in S$ is a permutation. If $p : S \to S$ is a permutation, then there is a unique permutation $q : S \to S$ such that $pq = qp = 1_S$.*

Proof. Exercise. ∎

Definition 1.3.4 *A* **permutation matrix** *is a matrix P such that left multiplication by P is a permutation of the rows of the matrix. That is, the rows of PA are precisely the rows of A in some order.*

Example 1.3.5 *Let P be the 3×3 matrix defined by*

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then we note for any column vector $(x_1, x_2, x_3)^T$, we have

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix}.$$

Notice that the entry in the first position is sent to the second, the entry in the second position is sent to the third and the entry in the third position is sent to the first. Thus the matrix P acts just like the permutation of $\{1, 2, 3\}$ in our previous example. This is why P is called a permutation matrix.

A very important remark is in order here. Namely, when we permute the entries of a vector $(x_1, \ldots, x_n)^T$ with a permutation p, the indices are permuted in the opposite way. For example, for

the permutation p above, the indices of the vector $(x_1, x_2, x_3)^T$ are permuted

$$\begin{aligned} 1 &\mapsto 3 \\ 2 &\mapsto 1 \\ 3 &\mapsto 2 \end{aligned}$$

which is just the inverse of the permutation p. Therefore there are two ways to associate a permutation p of the set $\{1, \ldots, n\}$ to an $n \times n$ permutation matrix P: p is the permutation that describes how P permutes the entries of the vector or p is the permutation that describes how P permutes the indices of the entries. These two permutations are inverse to each other. We choose the former so that if P is a permutation matrix and $X = (x_1, \ldots, x_n)^T$ is a vector, we have

$$PX = \begin{bmatrix} x_{p^{-1}(1)} \\ \vdots \\ x_{p^{-1}(n)} \end{bmatrix}.$$

We let e_i denote the vector in $\mathbb{R}^n$ with a single 1 in the i^{th} entry and 0 elsewhere. The collection $\{e_1, \ldots, e_n\}$ is called the **standard basis for** $\mathbb{R}^n$. We then have the following proposition.

Proposition 1.3.6 *Let P be the permutation matrix associated with the permutation p.*

1. *The j^{th} column of P is the vector $e_{p(j)}$.*
2. *P is the sum of n* **matrix units***: $P = \sum_{j=1}^{n} e_{p(j)j}$.*

Proof. (1.) By definition of matrix multiplication, the product Pe_j is just the j^{th} column of P. However, since p is defined by permuting the entries of the vector on which P acts, we see that Pe_j is also a standard basis vector whose non-zero entry is in the $p(j)^{\text{th}}$ position by definition. That is $Pe_j = e_{p(j)}$ and (1) follows.
(2) This is obvious from (1) since by definition, e_{ij} is the $n \times n$ matrix with all entries zero except the (ij)-entry which is a 1. ∎

We remark that one corollary of this proposition is that every permutation matrix P can be obtained from I_n by applying the corresponding permutation p to the *rows* of I_n.

Proposition 1.3.7 *1. If p and q are permutations with permutation matrices P and Q respectively, then the permutation matrix of the permutation pq is the matrix product PQ.*

2. *A permutation matrix is invertible and $P^{-1} = P^T$.*

Proof. The notation pq means composition of functions so that in particular we have

$$pq(j) = p(q(j)).$$

Since the matrix P operates by permuting the rows according to p and Q operates by permuting the rows according to q, the associative law for matrix multiplication tells us that PQ will operate by permuting the rows according to pq:

$$(PQ)X = P(QX).$$

Therefore the matrix of pq is PQ. The proof of (2) is left as an exercise. ■

Definition 1.3.8 *Using a familiar property of the determinant and (2) from the previous theorem, it is easy to see that* $\det P = \pm 1$ *for all permutation matrices* P*. Therefore we define the* **sign** *of a permutation* p *to be the determinant of the permutation matrix the represents* P*. That is*

$$\operatorname{sign} p = \det P = \pm 1.$$

A permutation p *is called* **even** *if* $\operatorname{sign} p = 1$ *and* **odd** *if* $\operatorname{sign} p = -1$*.*

1.4 Lecture 4: Complex numbers

In this final preliminary lecture, we briefly recall some elementary facts about the complex plane $\mathbb{C}$. First, we recall that it is natural to identify $\mathbb{C}$ with the usual Euclidean plane $\mathbb{R}^2$ under the correspondence $(x, y) \leftrightarrow z = x + iy$. The (real) number x is called the **real part** of the complex number z and the (real) number y is called the **imaginary part** of z. We can therefore consider the real numbers $\mathbb{R}$ as a subset of $\mathbb{C}$; namely $\mathbb{R}$ consists of all those complex numbers with zero imaginary part. If $z_1 = x_1 + iy_2$ and $z_2 = x_2 + iy_2$ are two complex numbers, then we define their **sum** and **product** by

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$$

and

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Note that the product formula is derived by formally multiplying out and using the relation $i^2 = -1$.

If $z = x + iy \in \mathbb{C}$, we define the **complex conjugate** of z as $\overline{z} = x - iy$. Note that $z = \overline{z}$ iff. $z \in \mathbb{R}$. We define the **norm** (or **length** or **modulus**) of z to be the real number $|z| = \sqrt{x^2 + y^2} = \sqrt{z\overline{z}}$. Geometrically, the norm of z is the distance from z to the origin in the plane. It is easy to show that if $z \neq 0$, then $z^{-1} = \overline{z}/|z|^2$. If θ is the angle made between z and the positive x-axis (θ is measured counter-clockwise), then the **polar form** of the complex number z is

$$z = |z|(\cos\theta + i\sin\theta).$$

It is easy to show, using familiar trigonometric identities, that

$$z_1 z_2 = |z_1 z_2|(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2))$$

so that in multiplying two complex numbers, we simply multiply their lengths and add their angles. We usually write $e^{i\theta} = \cos\theta + i\sin\theta$ so that in this notation we have $z = |z|e^{i\theta}$ and

$$z_1 z_2 = |z_1 z_2| e^{i(\theta_1 + \theta_2)}.$$

Note that if $z = re^{i\theta}$, then $z^n = r^n e^{in\theta}$. In particular, the solutions of the polynomial $z^n = 1$ must all have the form $\xi_k = e^{2\pi ik/n}$ for $k = 0, 1, \ldots, n-1$. The elements of the collection $U_n = \{1, \xi_1, \ldots, \xi_{n-1}\}$ are called the n^{th} **roots of unity**. U_n will be an important example for us in the future.

Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ denote the set of complex numbers z that have length 1. S^1 is called the **unit circle** or the **1-sphere**. Note that $U_n \subset S^1$ for every $n \geq 1$. n fact, U_n can be visualized as n equally spaced points on S^1. We close this lecture by remarking that the formula $|zw| = |z||w|$ shows that S^1 is closed under multiplication.

Chapter 2

Elements of Group Theory

2.1 Lecture 5: The definition of a group

Before we become weighted down with details, let us look ahead and discuss briefly what our subject matter is all about. In abstract algebra, the test of time has shown that certain operations on sets, that is ways of combining two elements of the set to get a third, have great importance in many branches of mathematics. By studying the essential common properties of these operations abstractly, it is possible to prove general theorems that when applied to a specific concrete situation expose deep insight into the matter at hand that would have otherwise been obscured by the abundance of extraneous information. The operations and properties that are essential to the system are discovered over time by noticing similar behaviors in different concrete situations. For example, the notion of a group (our first topic of study) arose in case after case in special problems throughout the end of the eighteenth and the beginning of the nineteenth centuries. Some of these problems were about matrices and permutations. The abstract notion of a group was not introduced until relatively late in the nineteenth century however. To quote I.N. Herstein, "Amongst mathematicians, neither the beauty nor the significance of the first example we have chosen to discuss—groups—is disputed".

Definition 2.1.1 *Let S be a non-empty set. A function $\mu : S \times S \to S$ is called a* **binary operation** *on S or a* **law of composition** *on S.*

We usually do not use the functional notation $\mu(a, b)$ to denote the image of the pair (a, b) under μ. Rather, we choose a symbol to denote the operation and write a and b on either side of this symbol.

Sometimes we simply juxtapose a and b. Therefore the notations $ab, a+b, a*b, a\circ b, a\cdot b, \ldots$ may all be used to denote the element $\mu(a,b) \in S$. Furthermore, we will often refer to the binary operation as the symbol itself. That is, we may say "let S be a set with a binary operation $*$" as opposed to "let S be a set with a binary operation $\mu : S \times S \to S$.

Example 2.1.2 *1. Let $S = \mathbb{Z}$ denote the set of integers and define $\mu(n,m) = n+m$. That is μ is just ordinary addition of integers.*

2. Let $S = \mathbb{R}^\times$ be the set of non-zero real numbers and define $\mu(a,b) = ab$. Then μ is a binary operation.

3. Ordinary matrix multiplication $\mu(A,B) = AB$ is a binary operation on the set $S = \mathbb{R}^{n\times n}$.

4. Let $\mathrm{Aut}(S)$ denote the set of all permutations of S and define $\mu : \mathrm{Aut}(S) \times \mathrm{Aut}(S) \to \mathrm{Aut}(S)$ by $\mu(p,q) = p \circ q$ (function composition). Then μ is a binary operation on $\mathrm{Aut}(S)$.

Definition 2.1.3 *Let S be a non-empty set and μ a binary operation on S written $\mu(a,b) = ab$. Then we say μ is* **associative** *if for all $a,b,c \in S$:*

$$(ab)c = a(bc).$$

We say μ is **commutative** *if for all $a,b \in S$:*

$$ab = ba.$$

All of the binary operations in the above example are associative, but only the first two are commutative. We remark that if a binary operation is associative, then we can show using mathematical induction that the expression

$$a_1 a_2 \cdots a_n$$

is unambiguous for all positive integers n.

Definition 2.1.4 *If S is a non-empty set with a binary operation, then an element $e \in S$ is called an* **identity** *for the operation if*

$$ae = ea = a$$

for all $a \in S$.

Proposition 2.1.5 *An identity for a binary operation on S is unique.*

Proof. Suppose that both e and e' are identities. Then since e is an identity, $ee' = e'$. Similarly, since e' is an identity $ee' = e$ and it follows that $e = e'$. ∎

If the binary operation on S is written ab, we will sometimes use the symbol 1 to denote the identity element in S. We will use the symbol 0 if the operation is written $a + b$.

Example 2.1.6 *Each of the four binary operations above has an identity. They are respectively* $0 \in \mathbb{Z}$, $1 \in \mathbb{R}^\times$, $I_n \in \mathbb{R}^{n\times n}$ *and* $1_S \in \mathrm{Aut}(S)$.

Definition 2.1.7 *Suppose that S is a set with a binary operation with an identity e. An element $a \in S$ is called* **invertible** *if there exists an element $a' \in S$ such that*

$$aa' = a'a = e.$$

As mentioned at the outset of this lecture, mathematicians at the turn of the nineteenth century were busy studying different sets with different binary operations that had certain properties in common. Eventually, the essential properties from all of these examples were distilled down into the following abstract definition.

Definition 2.1.8 (Group) *A* **group** $(G, *)$ *is a non-empty set G together with a binary operation $*$ on G such that the following three axioms hold:*

G1. *The binary operation $*$ is associative.*

G2. *There is an identity element $e \in G$.*

G3. *Every element $g \in G$ has an inverse.*

Example 2.1.9 *We have already discussed many examples of groups. Let's list a few here.*

1. $G = \mathbb{Z}$ *with the usual addition. (or* $G = \mathbb{Q}, \mathbb{R}, \mathbb{C}$*)*
2. $G = \mathbb{Q}^\times$ *with the usual multiplication. (or* $G = \mathbb{R}^\times, \mathbb{C}^\times$*)*
3. $G = \mathrm{GL}_n(\mathbb{R})$ *is the set of invertible $n \times n$ matrices with real entries under usual matrix multiplication. (or* $G = \mathrm{GL}_n(\mathbb{Q}), \mathrm{GL}_n(\mathbb{C})$*) The groups* GL_n *are called the* **general linear groups**.

4. $G = \mathrm{Aut}(S)$ *under composition of functions.*

Note that the set of positive integers does not form a group under addition as there is no zero element. Note that the set $\mathbb{R}^{n\times n}$ of $n \times n$ matrices is not a group under multiplication since every element does not have an inverse. However, $\mathbb{R}^{n\times n}$ is a group under matrix addition. In the next lecture, we will investigate the elementary properties of a group G that follow immediately from the definition. After this, we will look at these interpretations of these results in our known examples.

2.2 Lecture 6: Elementary properties of groups

In this lecture we see what properties of groups we can deduce from the definition as well as introduce more examples of groups. We will (slightly) abuse notation and refer to a group G rather than $(G, *)$. Also, we will write the product of two elements a and b in an arbitrary group as ab.

Proposition 2.2.1 *If G is a group, then the identity element e is unique. If $g \in G$, then the inverse of g is unique.*

Proof. We have already seen that the identity element of a binary operation is unique. If g' and g'' are both inverses to g, then multiplying the equation $e = gg'$ on the left by g'' and using the associative law gives

$$g'' = g''(gg') = (g''g)g' = eg' = g'.$$

■

We will denote the inverse of an element $g \in G$ by g^{-1} so that $gg^{-1} = g^{-1}g = e$.

Definition 2.2.2 *A group G is* **abelian** *if $ab = ba$ for all $a, b \in G$.*

Example 2.2.3 *The groups $\mathbb{Z}^+, \mathbb{Q}^+, \mathbb{R}^+, \mathbb{C}^+, \mathbb{Q}^\times, \mathbb{R}^\times, \mathbb{C}^\times$ are all abelian. Here the $+$ means the binary operation is the usual addition. The groups $\mathrm{GL}_n(\mathbb{Q}), \mathrm{GL}_n(\mathbb{R}), \mathrm{GL}_n(\mathbb{C})$ are not abelian.*

If G is a group and $g \in G$, then the associative law makes the expression

$$g^n = \overbrace{g \cdots g}^{n}$$

unambiguous for positive integers n. If we further define $g^0 = e$ and $g^{-n} = (g^{-1})^n$, it is easy to see that we have the familiar rules of exponents:

$$g^n g^m = g^{n+m} \qquad \text{and} \qquad (g^n)^m = g^{nm}.$$

It is not a good idea to use fraction notation a/b in a group since if G is not abelian, we do not know if a/b means $b^{-1}a$ or ab^{-1} and these elements need not be the same. We leave it as an exercise to show that $(ab)^{-1} = b^{-1}a^{-1}$ for all $a, b \in G$ and more generally if $a_1, \ldots, a_n \in G$, then

$$(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} \cdots a_2^{-1} a_1^{-1}.$$

The following two theorems are useful for computations in a group.

Theorem 2.2.4 (Cancellation Laws) *If G is a group and $a, b, c \in G$ satisfy $ab = ac$, then $b = c$. Similarly if $ba = ca$, then $b = c$.*

Proof. If $ab = ac$, then multiplying on the left by a^{-1} and using the associative law shows that $(a^{-1}a)b = (a^{-1}a)c$ so that $b = c$. The right cancellation law is left as an exercise. ∎

Theorem 2.2.5 *If G is a group and $a, b \in G$, then the equations $ax = b$ and $ya = b$ have unique solutions in G.*

Proof. Let $a, b \in G$ and define $x = a^{-1}b$. Then $x \in G$ and

$$ax = a(a^{-1}b) = (aa^{-1})b = eb = b$$

so that x is a solution to $ax = b$. This shows existence. If x_1 and x_2 are both solutions to $ax = b$, then $ax_1 = b = ax_2$ so that $x_1 = x_2$ by the cancellation law and therefore our solution is unique. The existence and uniqueness of solutions to $ya = b$ is left as an exercise. ∎

We conclude this lecture with an example that will be of great importance throughout our course.

Example 2.2.6 Let us denote the set of permutations of the set $\{1, 2, \ldots, n\}$ by S_n. We know S_n is a group under the binary operation of function composition. Note that there are $n!$ elements in S_n. The structure of S_n becomes very complicated as n gets large. We want to look at the case $n = 3$. Recall the 3×3 permutation matrix

$$x = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

which permutes the *entries* of the vector $(v_1, v_2, v_3)^T$ cyclicly by $1 \mapsto 2 \mapsto 3 \mapsto 1$. We let y be the permutation that interchanges the first two entries and fixes the third:

$$y = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The reader can verify that the $3! = 6$ permutations of $\{1, 2, 3\}$ are

$$S_3 = \{1, x, x^2, y, xy, x^2y\}.$$

Remember the "1,2,3" in $\{1, 2, 3\}$ represent the entries positions of the vector, not the indices. A hint for this exercise is to show that $x^3 = 1, y^2 = 1$ and $yx = x^2y$. Together these rules let you write any product for x and y as $x^j y^i$ with $0 \leq j \leq 2$ and $0 \leq i \leq 1$. Since there are only 6 elements in S_3, your done as soon as you show these six are all distinct. We remark that this is our first explicit example of a finite group, and it is not abelian. We will see that S_3 is the smallest non-abelian finite group.

2.3 Lecture 7: Subgroups

An important principle in mathematics is that one can gain insight into the structure of an object by studying subsets of that object that are themselves the same type of object. We do this now in the case that the object is a group.

Definition 2.3.1 (Subgroup) *If G is a group, a non-empty subset H of G is called a* **subgroup** *of G if H is itself a group under the induced operations in G. If H is a subgroup of G, we write $H \leq G$ and $H < G$ if $H \neq G$. Of course $H = G$ is a subgroup of G called the* **improper subgroup**. *Also, then set $H = \{e\}$ consisting of the identity element alone is a subgroup called the* **trivial subgroup**. *Any other subgroup $\{e\} \subset H \subset G$ is called* **proper**.

Example 2.3.2

1. Let $\mathrm{SL}_n(\mathbb{R})$ denote the subset of $\mathrm{GL}_n(\mathbb{R})$ that consists of all matrices $A \in \mathrm{GL}_n(\mathbb{R})$ such that $\det A = 1$. Then $\mathrm{SL}_n(\mathbb{R})$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$ called the **special linear group over** $\mathbb{R}$.

2. The subset S^1 of complex numbers of length 1 is a subgroup of $\mathbb{C}^\times$.

3. The n^{th} roots of unity U_n are a subgroup of $\mathbb{C}^\times$.

4. The set of positive integers is not a subgroup of $\mathbb{Z}^+$.

5. The set S^1 is not a subgroup of $\mathbb{C}^+$.

The following theorem is useful in showing a subset of a group G is a subgroup. In fact, it is the definition of subgroup given in your text.

Theorem 2.3.3 *A subset H of a group G is a subgroup of G if and only if*

1. *H is closed under the binary operation in G. That is, if $a, b \in H$, then $ab \in H$.*
2. *The identity element $e \in H$.*
3. *For every element $a \in H$, the inverse $a^{-1} \in H$.*

Proof. ($\Longrightarrow$) If H is a subgroup of G, then in particular H is a group under the binary operation in G so that H is closed under this operation. Moreover, since H is non-empty, there must be an element $a \in H$ so that $a^{-1} \in H$ since H is a group and hence $aa^{-1} = e \in H$. Finally condition (3) follows immediately since H is a group.
($\Longleftarrow$) Suppose that (1), (2) and (3) hold. Then H has a binary operation by (1) and this operation is associative since G is a group. Conditions (2) and (3) are precisely the second and third group axioms respectively so that H is a group under the induced operation from G and hence H is a subgroup by definition. ■

Definition 2.3.4 *If G is a group and the set G, is finite, then we say G is a* **finite group**. *If G is a finite group, the* **order of** *G is the number of elements in G. The order of G is denoted by $|G|$. If G is not a finite group, we say the order of G is infinite.*

Example 2.3.5 Let A_n denote the subset of S_n that consists of all even permutations. That is

$$A_n = \{p \in S_n : \operatorname{sign} p = 1\}.$$

We claim A_n is a subgroup of S_n called the **alternating group**. We will prove this using our new theorem. First, suppose that $p, q \in A_n$ and let P and Q be the corresponding permutation matrices. We know the permutation matrix for pq is PQ and moreover, $\det(PQ) = \det P \det Q = 1 \cdot 1 = 1$ so that $\operatorname{sign}(pq) = 1$. Therefore A_n is closed under the binary operation in S_n. The matrix of the identity permutation is I_n and $\det I_n = 1$ so that the identity permutation is in A_n. Finally, we know that the matrix for the permutation p is $P^{-1} = P^T$ and $\det P = \det P^T$ so that if $p \in A_n$, $p^{-1} \in A_n$ and hence $A_n \leq S_n$ as claimed.

We end this lecture with a discussion of a very important class of subgroups called cyclic subgroups. If $H \leq G$ and $a \in H$, then by our theorem, we see that $a^n \in H$ for every integer n. That is a subgroup containing a must contain the set

$$\{a^n : n \in \mathbb{Z}\}.$$

Note that this set need not be infinite. Indeed if $a = x \in S_3$, this set has the three elements $\{1, x, x^2\}$ precisely because $x^3 = 1$. We are hinting around the following theorem.

Theorem 2.3.6 *If G is a group and $a \in G$, then the set*

$$H = \langle a \rangle = \{a^n : n \in \mathbb{Z}\}$$

is a subgroup of G and is the smallest subgroup of G that contains a. That is, every subgroup K of G that contains a also contains $\langle a \rangle$.

Proof. To prove the first assertion, note that $a^n a^m = a^{n+m}$ so that $\langle a \rangle$ is closed under the binary operation in G. Also, $a^0 = e \in \langle a \rangle$ by definition. Finally, if $a^n \in \langle a \rangle$, then $(a^n)^{-1} = a^{-n} \in \langle a \rangle$ so that $\langle a \rangle$ is a subgroup. It remains to show that it is the smallest subgroup containing a. But our remarks before the theorem show that if K is a subgroup containing a, then $\langle a \rangle \subseteq K$ and this is precisely what it means to be the smallest subgroup containing a. ∎

Definition 2.3.7 *If G is a group and $a \in G$, the subgroup $\langle a \rangle$ is called the* **cyclic subgroup generated by** *a. The element a is called a* **generator** *of $\langle a \rangle$. If the order of $\langle a \rangle$ is n, then we say the element a has* **order** *n. Otherwise we say a has* **infinite order**. *We say that the group G is* **cyclic** *if $G = \langle a \rangle$ for some $a \in G$. In this case we say a* **generates** *G.*

We remark that a cyclic subgroup generated by $a \in G$ may be infinite or finite. Too see this, take $\langle 1 \rangle \leq \mathbb{Z}^+$ and $\langle x \rangle \leq S_3$. Note that $\langle e \rangle = \{e\}$ is the trivial subgroup. The group $\mathbb{Z}^+$ is cyclic. You can check directly that the group S_3 is not cyclic. This also follows from the following general theorem.

Theorem 2.3.8 *If G is a cyclic group, then G is abelian.*

Proof. If G is cyclic, then $G = \langle a \rangle$ for some $a \in G$. Therefore two arbitrary elements of G have the form a^n and a^m for some integers n, m and we compute

$$a^n a^m = a^{n+m} = a^{m+n} = a^m a^n$$

so that G is abelian. ∎

Example 2.3.9 The reader can check that the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ has order 6 in $\mathrm{GL}_2(\mathbb{R})$ and hence

generates a cyclic subgroup of order 6. The matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has infinite order however since

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}.$$

2.4 Lecture 8: Subgroups of the group $\mathbb{Z}^+$

We saw in the last lecture that the additive group $\mathbb{Z}^+$ is cyclic with generator 1. In an additive group, we usually write $na = a + \cdots + a$ (n times) instead of a^n. Our first goal in the current lecture is to show that all subgroups of $\mathbb{Z}^+$ are cyclic. Our proof will depend on the following important theorem from number theory. The proof is left to the reader.

Theorem 2.4.1 (Division with remainder) *If n is a positive integer and m is an arbitrary integer, then there exist unique integers q and r such that*

$$m = nq + r$$

and $0 \leq r < n$. ■

If n is a fixed integer, let

$$n\mathbb{Z} = \{a \in \mathbb{Z} : a = nk \text{ for some } k \in \mathbb{Z}\}.$$

Theorem 2.4.2 *Every subgroup of the additive group $\mathbb{Z}^+$ has the form $n\mathbb{Z}$ for some integer n.*

Proof. We will leave the verification that $n\mathbb{Z}$ is a subgroup as an exercise. We therefore proceed to show that every subgroup $H \leq \mathbb{Z}^+$ is of this form. If $H = \{0\}$ is the trivial subgroup, then $H = 0\mathbb{Z}$ and we are done. Otherwise there is a non-zero element $n \in H$. Without loss of generality, we may assume that n is the smallest positive integer in H (why?). We claim $H = n\mathbb{Z}$. Noting that $n\mathbb{Z} = \langle n \rangle$ we see that $n\mathbb{Z} \subseteq H$. To show the opposite inclusion, let $m \in H$ be an arbitrary element and write $m = nq + r$ where $0 \leq r < n$ by division with remainder. Note that since $n \in H$, $nq \in H$ so that $r = m - nq \in H$. But $r < n$ and n is the smallest positive integer in H so that we must have $r = 0$ and hence $m = nq \in n\mathbb{Z}$. Therefore $H \subseteq n\mathbb{Z}$ and hence $H = n\mathbb{Z}$. ■

Actually, our previous result is a special case of a more general theorem which states that any subgroup of a cyclic group is cyclic. The proof given above is easily modified to prove

Theorem 2.4.3 *Every subgroup of a cyclic group is cyclic.*

Sketch of proof. Let $G = \langle a \rangle$ and suppose $H \leq G$. If H is trivial, it is cyclic. Otherwise there is an element $a^n \in H$ with $n > 0$. It follows that there is a smallest positive integer m such that $a^m \in H$. Show that $\langle a^m \rangle = H$ using Division with remainder. ■

Proposition 2.4.4 *Let $n, m \in \mathbb{Z}$ be non-zero integers and define*

$$n\mathbb{Z} + m\mathbb{Z} = \{a \in \mathbb{Z} : a = nr + ms \text{ for some } r, s \in \mathbb{Z}\}.$$

Then $n\mathbb{Z} + m\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

Proof. Exercise.

Now, the previous two theorems together imply that given two non-zero integers n and m, there is a (positive) integer d such that $d\mathbb{Z} = n\mathbb{Z} + m\mathbb{Z}$. The integer d is called the **greatest common divisor (gcd) of n and m**. We state this formally as a proposition.

Proposition 2.4.5 *If n and m are non-zero integers and d is the positive integer such that $d\mathbb{Z} = n\mathbb{Z} + m\mathbb{Z}$, then*

1. *There exist integers r and s such that $d = nr + ms$.*
2. *d divides n and m.*
3. *If e is another integer dividing both n and m, then e divides d.*

Proof. Statement (1) simply asserts that $d \in n\mathbb{Z} + m\mathbb{Z}$. Similarly, taking $r = 1$ and $s = 0$ shows that $n \in d\mathbb{Z}$ so that $n = dk$ for some integer k so that d divides n by definition. Similarly d divides m. Finally, if e divides n and m, then $n = ek$ and $m = el$ for some integers k and l. But then we have

$$d = nr + ms = ekr + els = e(kr + ls)$$

which shows that e divides d. ■

We conclude this lecture with a return to the topic of cyclic groups. Let G be a group and let $H = \langle a \rangle$ be the cyclic subgroup generated by $a \in G$ so that

$$H = \{\cdots, a^{-2}, a^{-1}, e, a, a^2, \cdots\}.$$

If all of the elements in this list are distinct, then H is an infinite cyclic group. Otherwise we must have $a^n = a^m$ for some integers n and m with $n > m$ so that $x^{n-m} = e$ and $n - m > 0$. Therefore there is a non-zero power of a equal to e. We want to show that the smallest such power is the order of a.

Lemma 2.4.6 *The set S of integers n such that $a^n = e$ is a subgroup of $\mathbb{Z}$.*

Proof. If $a^n = e$ and $a^m = e$, then $a^n a^m = a^{n+m} = e$ too so that $n + m \in S$. Clearly $0 \in S$ and if $a^n = e$, then $a^{-n} = a^{-n}a^n = e$ so that $-n \in S$ and hence S is a subgroup of $\mathbb{Z}$. ■

Proposition 2.4.7 *Let G be a group and let $a \in G$ have finite order m. Then m is the smallest positive integer satisfying $a^m = e$. Moreover, $a^n = e$ iff. m divides n.*

Proof. By definition, the order of a is the order of the cyclic subgroup generated by a. By the lemma, the set of integers n such that $a^n = e$ is of the form $m'\mathbb{Z}$. Of course m' is the smallest positive element of $m'\mathbb{Z}$ so that m' is the smallest positive integer such that $a^{m'} = e$. It follows that the elements of the set $\{e, a, a^2, \ldots, a^{m'-1}\}$ are all distinct. Moreover, if $a^n \in \langle a \rangle$, then we can write $n = m'q + r$ with $0 \leq r < m'$ so that

$$a^n = a^{m'q+r} = (a^{m'})^q a^r = a^r$$

so that $a^n \in \{e, a, a^2, \ldots, a^{m'-1}\}$. It follows that $\langle a \rangle = \{e, a, a^2, \ldots, a^{m'-1}\}$ and hence $m' = m$ is the order of a by definition. Finally we note that $a^n = e$ iff. $n \in m\mathbb{Z}$ iff. $n = mk$ iff. m divides n. ■

2.5 Lecture 9: The dihedral groups D_n.

In the last lecture, we were able to completely classify the subgroups of the additive group $\mathbb{Z}^+$. In this lecture, we will study an family of subgroups of the symmetric group S_n called the **dihedral groups**. We begin with a slightly different way to view the group S_3.

Example 2.5.1 Recall that the group S_3 consists of the six permutations $\{1, x, x^2, y, xy, x^2y\}$ where x is the cyclic permutation $1 \mapsto 2 \mapsto 3 \mapsto 1$ and y interchanges 1 and 2 while fixing 3. There is a natural correspondence between the elements of S_3 and the ways in which two copies of an equilateral triangle with vertices labeled $1, 2$ and 3 (see Figure 2.1) can be placed with one on top of the other. For this reason, S_3 is sometimes also called the **group of symmetries of an equilateral triangle**.

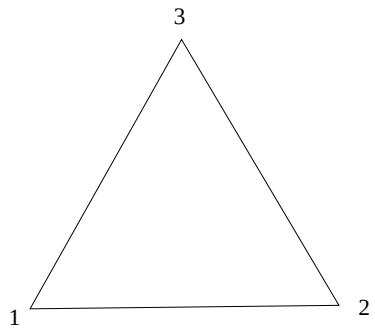


Figure 2.1: S_3 is the group of symmetries of a labeled equilateral triangle. The cyclic permutation x represents a rotation of the triangle through an angle of $2\pi/3$ radians counter-clockwise and the transposition y represents a reflection of the triangle through the median passing through the vertex 3.

Note that if we label the vertices as in Figure 2.1, then x represents a rotation of the triangle through an angle of $2\pi/3$ radians and y represents a reflection of the triangle through the median passing through the vertex 3. In the same way, we can define D_n to be the group of symmetries of a regular n-gon. D_n will have n "rotational" symmetries which are all multiplies of a rotation through $2\pi/n$ radians and n reflections through lines of symmetry. Therefore the order of D_n is $2n$. Note that if $n > 3$, then D_n is a proper subgroup of S_n since $2n < n!$.

To clarify these matters, we will compute the group D_4 - the symmetries of the square. To do so, we will introduce another notation for an element $p \in S_n$ that facilitates computations. If $p \in S_n$ is a permutation of $\{1, 2, \ldots, n\}$, we will write

$$p = \begin{pmatrix} 1 & 2 & \cdots & n \\ p(1) & p(2) & \cdots & p(n) \end{pmatrix}.$$

In this notation, the element $x \in S_3$ is written

$$x = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

We multiply two elements in this notation by composing the functions from *right to left*. For example, we have

$$yx = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

This computation shows that the product yx is the reflection of the triangle in the median through the vertex 1. The reader should explain this to him or herself geometrically. If we label the vertices

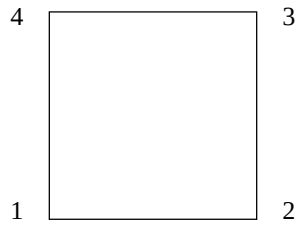


Figure 2.2: D_4 is the group of symmetries of a labeled square. There are four rotational symmetries and 4 lines of reflection.

of a square as in Figure 2.2, then the 8 symmetries of the square are:

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \mu_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$
$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$
$$\rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$
$$\rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}, \quad \delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

We are naively using the notations ρ_i for *rotations*, μ_j for *mirror image* in the lines through opposite midpoints and δ_j for *diagonal flips*. In particular, ρ_k is a rotation through $2\pi k/4$ radians counter-clockwise and $\rho_1^4 = \rho_0 = e$. If we temporarily let $\rho_1 = x$ and $\delta_1 = y$, then we leave as an exercise the verification that $x^4 = 1$, $y^2 = 1$ and $yx = x^3y$ (where 1 denotes the identity permutation) so that

$$D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y\}.$$

Compare this to our investigation of the group S_3 (which happens to equal D_3). Finally we remark that the reader should explain to him or herself geometrically why both products yx and x^3y are equal to μ_2 in the naive notations.

Now, if we label the vertices if an n-gon counter-clockwise, then the permutations $x, y \in S_n$ defined by

$$x = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ n-1 & n-2 & \cdots & 1 & n \end{pmatrix}$$

satisfy the relations $x^n = 1$, $y^2 = 1$ and $yx = x^{n-1}y$. It follows that for $n \geq 3$,

$$D_n = \{1, x, x^2, \ldots, x^{n-1}, y, xy, x^2y, \ldots, x^{n-1}y\}.$$

Of course, geometrically x represents a rotation of the n-gon by an angle of $2\pi/n$ radians counter-clockwise and y represents a reflection through the angle bisector of the vertex labeled n. The reader is invited to investigate the group D_2, the symmetries of a "regular 2-gon".

2.6 Lecture 10: Homomorphisms

In this section we begin to study the mappings between groups that are compatible with the group structure. Such mappings are called homomorphisms and they are a central topic in abstract algebra. In fact, a guiding principle in mathematics says that one can gain insight into the structure of an object by studying the mappings of the object into itself that preserve the structure. In this lecture, we will write all arbitrary groups multiplicatively and denote the identity elements with the symbol 1. We begin with the main definition.

Definition 2.6.1 (Homomorphism) *If G and G' are two groups, a mapping $\varphi : G \to G'$ is called a* **(group) homomorphism** *if for all $a, b \in G$, we have*

$$\varphi(ab) = \varphi(a)\varphi(b).$$

A remark on this formula is in order. Namely, we are using the same notation to denote two (possibly) different binary operations: the one in G and the one in G'. However, since $\varphi : G \to G'$, there is no real ambiguity since $\varphi(a)$ and $\varphi(b)$ are both elements of G' so that the product $\varphi(a)\varphi(b)$ can only mean the group product in G'.

Here are some examples of homomorphisms defined on familiar groups. In each example, the reader should carefully verify that the homomorphism property is valid for the defined function.

Example 2.6.2 *1.* Define $\varphi : \mathbb{Z}^+ \to \mathbb{Z}^+$ by $\varphi(n) = 2n$. Then φ is a homomorphism since for any integers $n, m \in \mathbb{Z}$, we have

$$\varphi(n + m) = 2(n + m) = 2n + 2m = \varphi(n) + \varphi(m).$$

Note that we use additive notation in the homomorphism property since this is how we write the group law in $\mathbb{Z}^+$.

2. For any two groups G and G', the map $\varphi : G \to G'$ defined by $\varphi(a) = 1$ for all $a \in G$ is a group homomorphism called the **trivial homomorphism**.

3. Let $U_2 = \{1, -1\} \subset \mathbb{C}^\times$ be the subgroup of $\mathbb{C}^\times$ consisting of plus and minus 1 and define a map sign : $S_n \to U_2$ from the symmetric group S_n to U_2 sending a permutation $p \in S_n$ to $\operatorname{sign} p$. Then sign is a homomorphism.

4. A familiar property of the determinant shows that the map det : $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is a homomorphism.

Now that we have seen a few examples of group homomorphisms, let us see what properties we can deduce from the definition.

Proposition 2.6.3 *If $\varphi : G \to G'$ is a homomorphism from a group G into a group G', then*

1. *If $1 \in G$ is the identity, then $\varphi(1) \in G'$ is the identity in G'.*

2. *If $a \in G$, then $\varphi(a)^{-1} = \varphi(a^{-1})$.*

3. *If $H \leq G$ is a subgroup of G, then $\varphi(H) \leq G'$ is a subgroup of G'.*

4. *If $K \leq G'$ is a subgroup of G', then $\varphi^{-1}(K) \leq G$ is a subgroup of G.*

Proof. (1) Note that for any $a \in G$,

$$\varphi(a) = \varphi(a \cdot 1) = \varphi(a)\varphi(1).$$

The cancellation law then implies that $1 = \varphi(1)$.
(2) Note that

$$\varphi(a)\varphi(a^{-1}) = \varphi(aa^{-1}) = \varphi(1) = 1$$

and similarly $\varphi(a^{-1})\varphi(a) = 1$ so that $\varphi(a^{-1}) = \varphi(a)^{-1}$ by uniqueness of inverses.
The proofs of (3) and (4) are exercises. ∎

If $\varphi : G \to G'$ is a group homomorphism, we use the notation $\operatorname{im}\varphi$ in addition to $\varphi(G)$ to denote the image of φ. A group homomorphism is called **injective, surjective** or **bijective** according to whether φ is injective, surjective or bijective as a map of sets. A bijective group homomorphism is called an **isomorphism**. The last theorem can be summarized by saying that group homomorphism

take identities to identities, inverses to inverses, subgroups to subgroups and the inverse image of a subgroup is a subgroup.

Every group homomorphism $\varphi : G \to G'$ has two important subgroups associated to it. One on them is $\operatorname{im}\varphi$, and the other is $\varphi^{-1}(\{1\})$.

Definition 2.6.4 *If $\varphi : G \to G'$ is a group homomorphism, we define the* **kernel of** φ, *written* $\ker\varphi$, *to be the subset of G that is mapped to the identity in G'. That is*

$$\ker\varphi = \{a \in G : \varphi(a) = 1\}.$$

Proposition 2.6.5 *If $\varphi : G \to G'$ is a group homomorphism, then $\ker\varphi \leq G$ is a subgroup of G.*

Proof. $\ker\varphi = \varphi^{-1}(\{1\})$. ∎

Example 2.6.6 *1.* The kernel of the map $\varphi : n \mapsto 2n$ from $\mathbb{Z}^+$ to $\mathbb{Z}^+$ is $\{0\}$.

2. The kernel of the trivial homomorphism $G \to G'$ is G.

3. The kernel of the sign homomorphism $\operatorname{sign} : S_n \to U_2$ consists of all permutations $p \in S_n$ such that $\operatorname{sign} p = 1$. This is the set of all even permutations so that $\ker(\operatorname{sign}) = A_n$; the alternating group.

4. The kernel of $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ *consists of those matrices* $A \in \mathrm{GL}_n(\mathbb{R})$ *such that* $\det A = 1$. *Thus* $\ker(\det) = \mathrm{SL}_n(\mathbb{R})$ *is the special linear group. Note that this gives us another proof that* $\mathrm{SL}_n(\mathbb{R})$ *is a subgroup of* $\mathrm{GL}_n(\mathbb{R})$.

The kernel of a group homomorphism $\varphi : G \to G'$ has another very important property. Namely, if $a \in \ker\varphi$ and $b \in G$ is any element, then the **conjugate element** bab^{-1} is also in $\ker\varphi$. To see this, just compute

$$\varphi(bab^{-1}) = \varphi(b)\varphi(a)\varphi(b)^{-1} = \varphi(b)\varphi(b)^{-1} = 1.$$

We abstract this property and make the following definition.

Definition 2.6.7 (Normal subgroup) *A subgroup N of a group G is called* **normal** *if for every $a \in N$ and every $b \in G$, the conjugate $bab^{-1} \in N$.*

As we have just proved, the kernel of a homomorphism is always a normal subgroup. In particular, the alternating group A_n is a normal subgroup of S_n and $\mathrm{SL}_n(\mathbb{R})$ is a normal subgroup of $\mathrm{GL}_n(\mathbb{R})$. Any subgroup of an abelian group is normal since $bab^{-1} = a$ in an abelian group. Subgroups of non-abelian groups need not be normal. For an example, let $H = \langle y \rangle = \{1, y\}$ be the cyclic subgroup of S_3 generated by y. Then H is not normal since

$$xyx^{-1} = xyx^2 = x^3yx = yx = x^2y \notin H.$$

Definition 2.6.8 (Center) *If G is a group, the subset*

$$Z = Z(G) = \{z \in G : za = az \text{ for all } a \in G\}$$

is called the **center of** G.

Proposition 2.6.9 *The center $Z(G)$ of a group G is a normal subgroup of G.*

Proof. Exercise. ∎

2.7 Lecture 11: Isomorphisms

Our limited experience with groups has already shown us that sometimes different looking groups can behave algebraically the same. For example, the cyclic subgroup of D_4 generated by the rotation ρ_1 is identical to the group U_4 if we make the identification $\rho_1 \leftrightarrow i$ where $i = e^{i\pi/2}$. Here we mean much more than both of these groups have order 4. Indeed, we mean that this assignment preserves the group structures as well. The goal of the current lecture is to make mathematically precise the notion of two groups G and G' being the same, or *isomorphic*. We begin by recalling the definition.

Definition 2.7.1 (Isomorphism) *A group homomorphism $\varphi : G \to G'$ is called an* **isomorphism** *if φ is a bijection. If there exists an isomorphism $\varphi : G \to G'$, we say that G is* **isomorphic to** G' *and we write $G \simeq G'$.*

Since an isomorphism is a bijection of the sets G and G' which preserves the group structures, it is clear that if G is isomorphic to G', then we can think of G' as the group G with the elements renamed by φ.

Let us digress for a moment to discuss the notion of an inverse map. If $f : A \to B$ is a bijection from a set A onto a set B, there is a natural map $f^{-1} : B \to A$ which is also a bijection defined

by reversing the arrows for the mapping f. That is, $f^{-1}(b) = a$ iff. $f(a) = b$. The map f^{-1} is called the **inverse map** to f. Do not confuse the inverse of a map with the inverse image of a set. The former is defined only when the map f is a bijection whereas the latter is always defined. It is clear that if f^{-1} exists, then $f^{-1} \circ f = 1_A$ and $f \circ f^{-1} = 1_B$. We leave it as an exercise to show that the inverse of a bijective group homomorphism is also a group homomorphism. If $\mathfrak{G}$ denotes the collection of all groups, the relation $\simeq$ on $\mathfrak{G}$ defined by $G \simeq G'$ iff. G is isomorphic to G' is an equivalence relation on $\mathfrak{G}$. We leave the proof of this fact to the reader.
The following proposition is useful in showing that a group homomorphism is injective.

Proposition 2.7.2 *If $\varphi : G \to G'$ is a group homomorphism, then φ is injective if and only if* $\ker \varphi = \{1\}$.

Proof. ($\Longrightarrow$) We know $1 \in \ker \varphi$ and if φ is injective, then this is the only element of $\ker \varphi$.
($\Longleftarrow$) Suppose that $\ker \varphi = \{1\}$ and $\varphi(a) = \varphi(b)$. Then we compute

$$\varphi(ab^{-1}) = \varphi(a)\varphi(b)^{-1} = 1$$

so that $ab^{-1} \in \ker \varphi$. Therefore $ab^{-1} = 1$ so that $a = b$ and hence φ is injective as claimed. ■

To show that two groups G and G' are isomorphic, you must:

1. Define a group homomorphism $\varphi : G \to G'$.
2. Show that $\ker \varphi = \{1\}$.
3. Show that $\operatorname{im} \varphi = G'$.

In practice this is usually easy to do if you have a feeling for why the two groups are isomorphic. Let's look at some examples.

Example 2.7.3 Let $G = \mathbb{R}^+$ be the additive group of real numbers and let $G' = \mathbb{R}^\times_{>0}$ be the multiplicative group of positive real numbers. Then G is isomorphic to G'. To see this, we first define a map $\varphi : G \to G'$ by

$$\varphi(x) = e^x.$$

Note that $e^x > 0$ for all $x \in \mathbb{R}$. Now, if $x, y \in \mathbb{R}$, then

$$\varphi(x + y) = e^{x+y} = e^x e^y = \varphi(x)\varphi(y)$$

so that φ is a group homomorphism. Moreover, we note that if $\varphi(x) = 1$, then $e^x = 1$ so that $x = \ln 1 = 0$ and hence φ is injective. Finally, given $y \in G'$, we must have $y > 0$ so that $x = \ln y \in \mathbb{R}$ and easily $\varphi(x) = y$.

Here is an important fact about infinite cyclic groups.

Theorem 2.7.4 *Every infinite cyclic group is isomorphic to the additive group $\mathbb{Z}^+$.*

Proof. Let H be an infinite cyclic group with generator a and write the group operation in H multiplicatively so that $H = \{a^n : n \in \mathbb{Z}\}$. This suggests that we should define a map $\varphi : \mathbb{Z} \to H$ by

$$\varphi(n) = a^n.$$

Note that if $n, m \in \mathbb{Z}$, then

$$\varphi(n + m) = a^{n+m} = a^n a^m = \varphi(n)\varphi(m)$$

so that φ is a group homomorphism. Since H is infinite, $a^n = e$ if and only if $n = 0$ so that $\ker \varphi = \{0\}$ and hence φ is injective. Obviously φ is surjective so that φ is an isomorphism and the proof is complete. ■

Since isomorphism is an equivalence relation, this theorem implies that any two infinite cyclic groups are isomorphic (they are both isomorphic to $\mathbb{Z}^+$). Therefore we will say that there is only one infinite cyclic group up to isomorphism. We remark that this very powerful result is deduced from only basic definitions. This fact is important enough to be stated again.

Corollary 2.7.5 *Any two infinite cyclic groups are isomorphic.* ■

What about finite cyclic groups? You already know an example of a cyclic group of order n for every positive integer n, namely the n^{th} roots of unity U_n. Another example is the cyclic subgroup generated by the rotation of $2\pi/n$ radians in the dihedral group D_n. Are these two cyclic groups of order n really different? An inspection of their multiplication tables shows that they are simply the same group with different names for the elements. In particular, we have the following theorem.

Theorem 2.7.6 *Any two cyclic groups of (finite) order n are isomorphic.*

Proof. Since isomorphism is an equivalence relation, it suffices to show that if H is a cyclic group of order n, then H is isomorphic to U_n. Recall that as a set, $U_n = \{1, \xi_1, \ldots, \xi_{n-1}\}$ where

$$\xi_k = e^{2\pi i k/n}.$$

Recalling that $H = \{e, a, a^2, \ldots, a^{n-1}\}$, we can define a map $\varphi : U_n \to H$ by $\varphi(\xi_k) = a^k$. Since we multiply in U_n and H by adding exponents (modulo n), the map φ is a homomorphism. Moreover, φ is obviously a bijection so that φ is a group isomorphism. ■

Now that we have talked about how to show two groups are isomorphic, we turn to the opposite problem; how to show that two groups are not isomorphic. Let us say that any property preserved by an isomorphism is a **structural property of a group**. That is if P is a structural property and a group G has the property P, then any group that is isomorphic to G also has the property P. You will be asked to show that the following properties are structural properties of a group.

1. The group is cyclic.
2. The group is abelian.
3. The group has order n.
4. The group has exactly two element of order 6.
5. The equation $x^2 = a$ has a solution for every a in the group.

Of course this is only a partial list of possible structural properties of groups. We end this lecture with some examples of non-isomorphic groups.

Example 2.7.7 *1.* The group U_4 is not isomorphic to the group D_4 since $|U_4| = 4$ and $|D_4| = 8$.

2. Both the groups U_6 and S_3 have order 6, but they cannot be isomorphic since U_6 is abelian whereas S_3 is non abelian.

3. Note that in the multiplicative group of real numbers, the equation $x^2 = a$ has a solution $x = \sqrt{a}$ for every a. However, the equation $x^2 = 2$ has no solution in the multiplicative group of positive rational numbers and hence these groups are not isomorphic.

2.8 Lecture 12: Cosets

In this lecture, we will study an equivalence relation on a group induced by a subgroup. As a first step, we consider the equivalence relation induced on the domain of a function $\varphi : S \to T$.

Suppose that S and T are two sets and $\varphi : S \to T$ is a function from S to T. We define a relation on S by $a \sim b$ iff. $\varphi(a) = \varphi(b)$. That is, we identify two elements in S iff. they have the same image under φ.

Proposition 2.8.1 *If $\varphi : S \to T$ is a function, the relation $\sim$ defined on S by $a \sim b$ if and only if $\varphi(a) = \varphi(b)$ is an equivalence relation.*

Proof. Exercise. ∎

The equivalence relation in the last proposition is called the **equivalence relation induced by** φ. As for all equivalence relations, we use the notation $\overline{S}$ to denote the set of equivalence classes and we write $\overline{a}$ for the equivalence class containing $a \in S$. Therefore

$$\overline{a} = \{b \in S : \varphi(b) = \varphi(a)\}$$

and

$$\overline{S} = \{\overline{a} : a \in S\}.$$

Sometimes we write $\overline{S} = S/\varphi$ if we want to emphasize the role of the function φ. Note that if $t \in \operatorname{im}\varphi$, then the set

$$\varphi^{-1}(t) = \{a \in S : \varphi(a) = t\}$$

is precisely an equivalence class; i.e. an element of $\overline{S}$. The sets $\varphi^{-1}(t), t \in T$ are called the **fibers of the map** φ. Our remarks imply that the non-empty fibers are precisely the elements of $\overline{S} = S/\varphi$. Therefore we have a bijective map

$$\overline{\varphi} : \overline{S} \to \operatorname{im}\varphi$$

which is defined by $\overline{\varphi}(\overline{a}) \to \varphi(a)$. Note that this map is well defined since $\overline{a} = \overline{b}$ iff. $\varphi(a) = \varphi(b)$. In this sense, you can think of the equivalence classes in S, that is the elements of $\overline{S}$, simply as the elements in $\operatorname{im}\varphi$.

We now turn to the case where $\varphi : G \to G'$ is a group homomorphism. In this case, the equivalence relation on G induced by φ is referred to as **congruence** and is usually denoted by $\equiv$ rather than $\sim$. Therefore if $\varphi : G \to G'$ is a group homomorphism,

$$a \equiv b \Longleftrightarrow \varphi(a) = \varphi(b).$$

Example 2.8.2 *1.* If $\varphi : \mathbb{Z}^+ \to U_4$ is defined by $\varphi(n) = e^{2\pi i n/4}$, then $n \equiv m$ iff. $e^{2\pi i n/4} = e^{2\pi i m/4}$ iff $(n-m)/4 \in \mathbb{Z}$ iff. $n \sim_4 m$. Therefore we see that this map has four equivalence classes $\{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}$.

2. The map $\varphi : \mathbb{C}^\times \to \mathbb{R}^\times$ defined by $\varphi(z) = |z|$ is a group homomorphism. Here $\overline{z}$ consists of all $w \in \mathbb{C}$ such that $|w| = |z|$. Thus the fibers are concentric circles centered at the origin. Since $\operatorname{im}\varphi = \mathbb{R}^\times_{>0}$, we see that the fibers are in bijective correspondence with the set of positive real numbers. Geometrically, this is the (obvious) fact that each circle centered at the origin in $\mathbb{C}^\times$ intersects the positive real axis exactly once.

Our next goal is to seek a relationship between the relation induced by a group homomorphism $\varphi : G \to G'$ and the kernel of φ.

Proposition 2.8.3 *If $\varphi : G \to G'$ is a group homomorphism with $\ker\varphi = N$ and $a, b \in G$, then the following are equivalent:*

(1) $\varphi(a) = \varphi(b)$.

(2) $a^{-1}b \in N$.

(3) $b = an$ *for some* $n \in N$.

Proof.$((1) \Longrightarrow (2))$ If $\varphi(a) = \varphi(b)$, then

$$1 = \varphi(a)^{-1}\varphi(b) = \varphi(a^{-1}b)$$

so that $a^{-1}b = n \in N$ by definition.

$((2) \Longrightarrow (3))$ This is trivial: $a^{-1}b = n \Longrightarrow b = an$.

$((3) \Longrightarrow (1))$ If $b = an$ and $n \in N$, then

$$\varphi(b) = \varphi(an) = \varphi(a)\varphi(n) = \varphi(a).$$

■

The set of elements $\{an : n \in N\}$ is denoted by aN an is referred to as the **coset of N containing** a. For emphasis, we write

$$aN = \{g \in G : g = an \text{ for some } n \in N\}.$$

The previous proposition implies that the coset aN is the set of elements that are congruent to a. Therefore the set of cosets

$$\{aN : a \in G\}$$

partitions the group G. These cosets are the fibers of the map φ. In particular the circles centered about the origin are the cosets of the absolute value homomorphism.

Now suppose that G is a group and H is *any* subgroup of G (not necessarily the kernel of a homomorphism). We can still define the notion of coset.

Definition 2.8.4 *If $H \leq G$ and $a \in G$, then the* **left coset of H containing** *a is the subset*

$$aH = \{ah : h \in H\}.$$

Note that H itself is a coset: $H = 1H$. Note also that $a \in aH$ since $a = ea$ and $e \in H$ since H is a subgroup. In the proof of the following theorem, note how we constantly use the fact that H is a subgroup of G.

Proposition 2.8.5 *If $H \leq G$, then the relation $\equiv$ defined on G by $a \equiv b$ iff. $a = bh$ for some $h \in H$ is an equivalence relation on G. Moreover the equivalence class $\overline{a}$ containing $a \in G$ is precisely the left coset aH of H containing a.*

Proof. We first show that $\equiv$ is an equivalence relation.

Reflexive. Note $a = a1$ and $1 \in H$ since $H \leq G$ so that $a \equiv a$.

Symmetric. If $a \equiv b$, then $a = bh$ for some $h \in H$. But $H \leq G$ so that $h^{-1} \in H$ and clearly $b = ah^{-1}$ so that $b \equiv a$.

Transitivity. Suppose that $a \equiv b$ and $b \equiv c$ so that $a = bh_1$ and $b = ch_2$ for some $h_1, h_2 \in H$. Then $h_2h_1 \in H$ since $H \leq G$ and easily $a = bh_1 = (ch_2)h_1 = c(h_1h_2)$ so that $a \equiv c$.

To show the final statement, note that $b \in \overline{a}$ iff. $b \equiv a$ iff. $b = ah$ for some $h \in H$ iff. $b \in aH$. ■

Corollary 2.8.6 *If $H \leq G$, then the cosets of H partition G. Moreover, if $a, b \in G$, then $aH = bH$ if and only if $a^{-1}b \in H$.*

Proof. The first statement is immediate since the cosets are the equivalence classes of an equivalence relation. As for the second statement, $aH = bH$ iff. $b \equiv a$ iff. $b = ah$ for some $h \in H$ iff. $a^{-1}b = h \in H$. ■

Example 2.8.7 *1.* If $G = \mathbb{Z}^+$ and $H = 4\mathbb{Z}$, then in additive notation, the coset containing $n \in \mathbb{Z}$ is

$$n + 4\mathbb{Z} = \{n + 4k : k \in \mathbb{Z}\}.$$

In particular, there are four cosets: $4\mathbb{Z}, 1+4\mathbb{Z}, 2+4\mathbb{Z}$ and $3+4\mathbb{Z}$. Compare this to the set of equivalence classes in $\mathbb{Z}^+$ induced by the homomorphism $\varphi : \mathbb{Z}^+ \to U_4$ defined earlier.

2. Let $G = S_3 = \{1, x, x^2, y, xy, x^2y\}$ and let $H = \langle xy \rangle$ be the cyclic subgroup generated by the order 2 element xy so that $H = \{1, xy\}$. The left cosets of H in G are the three sets

$$\begin{array}{ccccc} H & = & \{1, xy\} & = & xyH \\ xH & = & \{x, x^2y\} & = & x^2yH \\ x^2H & = & \{x^2, y\} & = & yH. \end{array}$$

Note that the cosets do indeed partition the group. Note also that they each have exactly 2 elements and that 2 is the order of H. This is not an accident as the next proposition states.

Proposition 2.8.8 *Let G be a group and let H be a subgroup of G. Then for any $a \in G$, then number of elements in the left coset aH is the order of H.*

Proof. Our proof will use the common procedure for showing that two sets have the same number of elements: we exhibit a bijection between them. To this end, for any $a \in G$ we define a map $f : H \to aH$ by $f(h) = ah$. Clearly f is a surjection and $f(h_1) = f(h_2)$ implies $ah_1 = ah_2$ which in turn implies that $h_1 = h_2$ by the cancellation law in G. Therefore f is injective and hence $|H| = |aH|$ as claimed. ■

If H is a subgroup of a group G, the number of left cosets of H in G is called the **(left) index of H in G** and is denoted by $[G : H]$.

Example 2.8.9

1. By our previous example, $[\mathbb{Z} : 4\mathbb{Z}] = 4$. More generally, $[\mathbb{Z} : n\mathbb{Z}] = n$ for any $n \geq 1$.

2. If $H = \{1, xy\} \leq S_3$, then $[S_3 : H] = 3$.

Since the cosets aH together form a partition of G and each coset aH has the same number of elements as H, we have deduced the very important **counting formula for groups**:

$$|G| = |H|[G : H]$$

where the equality has the obvious meaning if $|G| = \infty$. Some very deep facts about the structure of groups follow from this innocent looking formula. We present some of them here. We leave some of the proofs to the reader.

Corollary 2.8.10 (Lagrange's Theorem) *If G is a finite group and H is a subgroup of G, then the order of H divides the order of G.*

Proof. This follows immediately from the counting formula: $|H|$ is a factor of $|G|$. ■

Corollary 2.8.11 *If G is a finite group and $a \in G$, then the order of a divides the order of G.* ■

Corollary 2.8.12 *Any group G with prime order p is cyclic.*

Proof. Since $|G| = p$ is prime, G is not the trivial group so that we can choose an element $a \in G$ with $a \neq 1$. Therefore the order of the cyclic subgroup $\langle a \rangle$ generated by a is greater than 1. But Lagrange's Theorem implies that the order of a divides $|G| = p$ so that we must have the order of a equal to p and hence G is cyclic. ■

Corollary 2.8.13 *If p is a prime integer, then there is only one group G of order p up to isomorphism.*

Proof. If $|G| = p$ is prime, then G is cyclic by Lagrange's Theorem. Therefore G is isomorphic to U_p since all cyclic groups of order p are isomorphic to this group. ■

Corollary 2.8.14 *If $\varphi : G \to G'$ is a homomorphism of finite groups, then*

$$|G| = |\ker \varphi| \cdot |\operatorname{im} \varphi|.$$

Proof. If we apply the counting formula to the subgroup $\ker \varphi$, we have

$$|G| = |\ker \varphi|[G : \ker \varphi].$$

Now the index $[G : \ker \varphi]$ is the number of left cosets of $\ker \varphi$ in G, and we have seen that this is precisely the number of fibers of the map φ which is in bijective correspondence with $\operatorname{im} \varphi$. In short we have $[G : \ker \varphi] = |\operatorname{im} \varphi|$ which proves the corollary. ■

We conclude this lecture with a remark about the "leftness" in our construction of left cosets. Suppose instead that we defined the **right cosets of H in G** by

$$Ha = \{ha : h \in H\}.$$

We leave it to the reader to show that the relation $\equiv$ on G defined by $a \equiv b$ iff. $a = hb$ for some $h \in H$ is an equivalence relation on G whose equivalence classes are precisely the right cosets of H

in G. We remark that in general, the left and right coset of H containing a need not be equal. That is, in general, $aH \neq Ha$. The reader should verify this for the example $H = \{1, xy\} \leq S_3$. However, the following is true.

Proposition 2.8.15 *If H is a subgroup of a group G, then the number of left cosets of H in G is equal to the number of right cosets of H in G. Therefore we may refer to the index $[G : H]$ as the number of cosets of H in G without mentioning left or right.*

Proof. We define a map f from the set of left cosets of H in G to the set of right cosets of H in G by

$$f(aH) = Ha^{-1}.$$

Since it is possible that $aH = bH$ with $a \neq b$, we must show that this map is well defined. That is we must show that if $aH = bH$, then $Ha^{-1} = Hb^{-1}$. But we have

$$aH = bH \Longleftrightarrow a = bh \Longleftrightarrow a^{-1} = h^{-1}b^{-1} \Longleftrightarrow Ha^{-1} = Hb^{-1}.$$

We leave the verification that this map is a bijection to the reader. ∎

An important fact in group theory is that the left and right cosets of a subgroup coincide precisely when that subgroup is normal. We end this lecture with this result.

Proposition 2.8.16 *A subgroup H of a group G is normal in G if and only if $aH = Ha$ for all $a \in G$.*

Proof. ($\Longrightarrow$) Suppose that H is normal in G and let $ah \in aH$. Note that $ah = (aha^{-1})a$ and the conjugate $h' = aha^{-1} \in H$ since H is normal in G. Therefore $ah = h'a \in Ha$ so that $aH \subset Ha$. Similarly we have $Ha \subset aH$ so that $aH = Ha$ as desired.

($\Longleftarrow$) Suppose that $aH = Ha$ for all $a \in G$ and let $h \in H$. Note that for all $a \in G$, $ah \in aH = Ha$ so that $ah = h'a$ for some $h' \in H$. Therefore $aha^{-1} = h' \in H$ so that H is a normal subgroup by definition. ∎

2.9 Lecture 13: Products of groups

We begin this lecture with a slight generalization of the definition of direct product given in the first lecture.

Definition 2.9.1 (Cartesian product) *The* **Cartesian product** *of the sets $S_1, S_2, \ldots, S_k$ is the set of all ordered k-tuples $(a_1, \ldots, a_k)$, where $a_i \in S_i$. The Cartesian product is denoted by*

$$S_1 \times S_2 \times \cdots \times S_k$$

or by

$$\prod_{i=1}^{k} S_i.$$

Of course we are interested in the case when each of the sets is a group. We want to make the Cartesian product into a group in a way that relates group structure to the group structure in the individual factors.

Theorem 2.9.2 *Let $G_1, \ldots, G_k$ be a collection of groups. For $(a_1, \ldots, a_k), (b_1, \ldots, b_k) \in \prod_{i=1}^{k} G_i$, define*

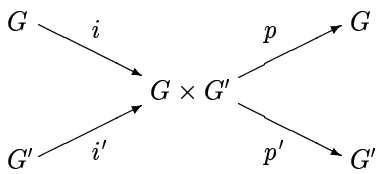
$$(a_1, \ldots, a_k)(b_1, \ldots, b_k) = (a_1 b_1, \ldots, a_k b_k).$$

Then $\prod_{i=1}^{k} G_i$ is a group, the **direct product of the groups** *G_i, under this product.*

Sketch of Proof. Note that since each G_i is a group $a_i b_i \in G_i$ whenever $a_i, b_i \in G_i$ so that the definition if the binary operation makes sense. The verification of the associative law follows from the associative law in each factor. The element $(e, \ldots, e)$ is the identity and the inverse of $(a_1, \ldots, a_k)$ is $(a_1^{-1}, \ldots, a_k^{-1})$. ∎

From now on, we will focus on the case when $k = 2$, but the reader is advised that everything we say is valid for any k. In mathematics, it is usually easier to multiply than it is to factor. Products of groups are no exception to this rule so that we begin by studying the relationship between the two factors G and G' and the product $G \times G'$. The situation is understood best in terms of four homomorphisms called the **canonical inclusions** and **projections**. Namely we have the following proposition.

Proposition 2.9.3 *If G and G' are groups, then the maps in the diagram*



defined by

$$i(a) = (a, 1), \quad i'(a') = (1, a')$$
$$p(a, a') = a, \quad p'(a, a') = a'$$

are group homomorphisms. Moreover, i and i' are injective so that we may identify G and G' with their images $G \times 1$ and $1 \times G'$ respectively. Finally, the maps p and p' are surjective, $\ker p = 1 \times G'$ *and* $\ker p' = G \times 1$.

Sketch of Proof. It is easy to show that the maps defined are group homomorphisms since we multiply in the product group by multiplying in each coordinate. Moreover, it is clear that $\operatorname{im} i = G \times 1$ and $i(a) = (1, 1)$ iff. $a = 1$ so that i is injective. The equation $p(a, 1) = a$ $(a \in G)$ shows that p is surjective and $p(a, a') = 1$ iff. $a = 1$ and $a' \in G'$ is arbitrary so that $\ker p = 1 \times G'$ as claimed. The statements about i' and p' are shown in the same way. ■

We remark that the proposition shows that G and G' are isomorphic to two *normal* subgroups of the product: $G \times 1$ and $1 \times G'$. They are normal since they are kernels of homomorphisms. We want to study groups by breaking them into product of smaller groups. Since our investigation of groups always involves homomorphisms, we need to study how homomorphisms into product groups behave with respect to the factors. The following theorem states that all such homomorphisms can be built by looking at homomorphisms into the factors one at at time.

Theorem 2.9.4 (Mapping property of products) *If H is any group, then the homomorphisms $\Phi : H \to G \times G'$ are in bijective correspondence with pairs (φ, φ') of homomorphisms $\varphi : H \to G$ and $\varphi' : H \to G'$. Moreover we have* $\ker \Phi = \ker \varphi \cap \ker \varphi'$.

Proof. First, if (φ, φ') is a pair of homomorphisms from H into G and G' respectively, the reader can check that the map $\Phi : H \to G \times G'$ defined by $\Phi(h) = (\varphi(h), \varphi'(h))$ is a group homomorphism. Conversely, if $\Phi : H \to G \times G'$ is a given group map, the compositions $\varphi = p\Phi$ and $\varphi' = p'\Phi$ are group homomorphisms from H into G and G' respectively and the pair (φ, φ') reassembles to give the map Φ. Finally, we note that $\Phi(h) = 1$ iff. $\varphi(h) = 1$ and $\varphi'(h) = 1$ so that $\ker \Phi = \ker \varphi \cap \ker \varphi'$.

■

As an application of the mapping property for products, we prove the following important fact about cyclic groups. Since there is only one cyclic group of order n up to isomorphism, we will let C_n denote "the" cyclic group of order n in what follows. Recall that two integers n and m are **relatively prime** if they have no common divisor greater than 1.

Theorem 2.9.5 *The cyclic group C_{nm} is isomorphic to the direct product $C_n \times C_m$ if and only if n and m are relatively prime.*

Proof. Let us denote the generators of C_{nm}, C_n and C_m by x, y and z respectively.
($\Longrightarrow$) If n and m are relatively prime, we define a map $\Phi : C_{nm} \to C_n \times C_m$ by $\Phi(x^i) = (y^i, z^i)$. Easily Φ is a group homomorphism and hence $\ker \Phi = \ker \varphi \cap \ker \varphi'$ where $\varphi = p\Phi$ and $\varphi' = p'\Phi$. Now $x^i \in \ker \varphi$ iff. $y^i = 1$ iff. $n|i$ iff. $i = nk$ for some integer k. Similarly $x^i \in \ker \varphi'$ iff. $i = mk', k' \in \mathbb{Z}$. It follows that i is a multiple of both n and m and hence nm since n is relatively prime to m. Therefore $x^i = 1$ so that $\ker \varphi \cap \ker \varphi' = \ker \Phi = \{1\}$ and hence Φ is injective. Moreover, this implies that the order of the image of Φ is nm and this is also the order of the direct product $C_n \times C_m$ so that Φ is surjective and hence an isomorphism.
($\Longleftarrow$) Now suppose that n is not relatively prime to m so that the least common multiple $[n, m]$ of n and m is less than nm. We claim that every element of the direct product $C_n \times C_m$ has order less than or equal to $[n, m]$ so that, in particular, $C_n \times C_m \ncong C_{nm}$. To see this, we recall from Lecture 8 that for $a \in C_n$, $a^k = 1$ iff. $n|k$. We note that by definition, $n|[n, m]$, $m|[n, m]$ and for any $(a, b) \in C_n \times C_m$,

$$(a, b)^{[n,m]} = (a^{[n,m]}, b^{[n,m]}) = (1, 1)$$

so that the order of (a, b) is at most $[n, m] < nm$. ∎

It is often much harder to determine if a given group G is isomorphic to a direct product of two groups. We end this lecture with a necessary and sufficient condition for a group to be isomorphic to a direct product of two subgroups. For notation, if A and B are two subsets of a group G, we define the product of A and B by

$$AB = \{g \in G : g = ab \text{ for some } a \in A \text{ and } b \in B.\}$$

Proposition 2.9.6 *Suppose that H and K are subgroups of a group G.*

1. *If $H \cap K = \{1\}$, then the product map $\pi : H \times K \to G$ defined by $\pi(h, k) = hk$ is injective. Moreover, $\operatorname{im} \pi = HK$.*

2. *If either H or K is a normal subgroup of G, then $HK = KH$ and $HK \leq G$ is a subgroup of G.*

3. *If H and K are normal, $H \cap K = \{1\}$ and $HK = G$, then G is isomorphic to the direct product $H \times K$.*

Proof. (1) Suppose that $(h_1, k_1), (h_2, k_2) \in H \times K$ and $h_1 k_1 = h_2 k_2$. Then it follows that $h_2^{-1} h_1 = k_2 k_1^{-1} \in H \cap K = \{1\}$. Therefore $h_1 = h_2$ and $k_1 = k_2$ so that π is injective.

(2) Suppose that H is a normal subgroup of G and let $h \in H$ and $k \in K$. Since H is normal, $khk^{-1} \in H$ and $(khk^{-1})k = kh$ so that $KH \subset HK$. Similarly one shows that $HK \subset KH$ and hence $HK = KH$. Now suppose that $hk, h'k' \in HK$. Then we have $(hk)(h'k') = h(kh')k'$ with $kh' \in KH = HK$. If we say $kh' = h''k''$, then we have

$$(hk)(h'k') = h(kh')k' = h(h''k'')k' = (hh'')(k''k') \in HK$$

so that HK is closed under the group product. Of course $1 = 1 \cdot 1 \in HK$. Finally, we have $(hk)^{-1} = k^{-1}h^{-1} \in KH = HK$ so that HK is closed under inverses and hence is a subgroup. The proof is similar if K is a normal subgroup.

(3) We claim the product map π is a group homomorphism in this case. Consider the product

$$(hkh^{-1})k^{-1} = h(kh^{-1}k^{-1}).$$

Since K is a normal subgroup, the left side is in K and since H is a normal subgroup, the right side is in H. Therefore $hkh^{-1}k^{-1} \in H \cap K = \{1\}$ so that $hk = kh$ for all $h \in H$ and $K \in K$. This immediately implies that the product map $\pi : H \times K \to G$ is a group homomorphism:

$$\pi((h,k),(h',k')) = \pi(hh', kk') = hh'kk' = hkh'k' = \pi(h,k)\pi(h',k').$$

By part (1), π is injective and by assumption, $\operatorname{im}\pi = HK = G$ so that π is an isomorphism. ∎

2.10 Lecture 14: Quotient groups

Our investigation of the kernel of a group homomorphism in Lecture 10 lead us to the notion of normal subgroups. In Lecture 11, we saw that these subgroups are precisely the subgroups whose left and right cosets coincide: N is normal iff. $aN = Na$ for all $a \in G$. In the present lecture, we want to show that this last property is equivalent to the set of cosets G/N having a group structure for which the canonical projection $G \to G/N$ is a group homomorphism. Constructions analogous to those contained in this lecture are found throughout all branches of mathematics and therefore this material should be throughly mastered. Recall if A and B are subsets of a group G, the product of A and B is the set

$$AB = \{ab : a \in A, b \in B\}.$$

In particular if $H \leq G$ is a subgroup, we can form the product of two cosets: $(aH)(bH)$. The following lemma is fundamental in what follows.

Lemma 2.10.1 *If N is a normal subgroup of a group G, then the operation*

$$(aN)(bN) = abN$$

is well defined on the set of left cosets of N in G.

Proof. We must show that the operation is well defined since its definition involves choices of representatives of the cosets of N in G. Therefore we must show that if an_1 and bn_2 represent aN and bN, then an_1bn_2 represents abN. Since N is normal, $bN = Nb$ so that $n_1b = bn_3$ for some $n_3 \in N$ and hence $an_1bn_2 = abn_3n_2$ so that $abN = an_1bn_2N$ as desired. ■

If $H \leq G$ is a subgroup of a group G, we will denote the set of (left) cosets of H in G by G/H. The last lemma says that if N is a normal subgroup, then the binary operation on G/N defined by $aNbN = abN$ is well defined. The following theorem confirms that this binary operation gives the set of cosets the structure of a group called the **quotient group of G by N.**

Theorem 2.10.2 *If N is a a normal subgroup of a group G, then the binary operation on G/N defined by $aNbN = abN$ makes G/N a group. Moreover the canonical map $\pi : G \to G/N$ defined by $\pi(a) = aN$ is a surjective group homomorphism with* $\ker \pi = N$.

Proof. The lemma implies the binary operation is well defined. The associative property follows directly from that in G:

$$(aNbN)cN = (abN)cN = (ab)cN = a(bc)N = aNbcN = aN(bNcN).$$

The element $eN = N$ is clearly the identity and $a^{-1}N$ is the inverse of aN. The map π is clearly surjective and

$$\pi(ab) = abN = aNbN = \pi(a)\pi(b)$$

so that π is a group homomorphism. Finally, we note that $aN = N$ iff. $a \in N$ so that $\ker \pi = N$. ■

The theorem has an important corollary.

Corollary 2.10.3 *A subgroup N is normal if and only if N is the kernel of a homomorphism.*

Proof. We have seen that kernels are normal subgroups and the theorem implies that if N is a normal subgroup of G, then N is the kernel of the canonical homomorphism $G \to G/N$. ∎

We have come to the central result of our lecture. It is a fundamental result in identifying quotient groups.

Theorem 2.10.4 (First isomorphism theorem) *If $\varphi : G \to G'$ is a group homomorphism with $N = \ker \varphi$, then the map $\overline{\varphi} : G/N \to \operatorname{im} \varphi$ defined by $\overline{\varphi}(aN) = \varphi(a)$ is a well defined isomorphism that satisfies $\varphi = \overline{\varphi}\pi$ where $\pi : G \to G/N$ is the canonical homomorphism.*

Proof. Here again we must show that if an also represents aN, then $\varphi(a) = \varphi(an)$ so that $\overline{\varphi}$ is well defined. But if $n \in N$, then $\varphi(an) = \varphi(a)\varphi(n) = \varphi(a)$. To show that $\overline{\varphi}$ is a homomorphism, we compute

$$\overline{\varphi}(aNbN) = \overline{\varphi}(abN) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(aN)\overline{\varphi}(bN).$$

Clearly $\overline{\varphi}$ is onto $\operatorname{im} \varphi$ and finally, $\overline{\varphi}(aN) = 1$ iff. $\varphi(a) = 1$ iff. $a \in \ker \varphi = N$ iff. $aN = N$ so that $\overline{\varphi}$ is injective and hence an isomorphism. To complete the proof, we note that for all $a \in G$,

$$\varphi(a) = \overline{\varphi}(aN) = \overline{\varphi}\pi(a).$$

∎

We end this lecture with some examples that show how the first isomorphism theorem is used to identify a quotient group.

Example 2.10.5

1. The absolute value homomorphism $\mathbb{C}^\times \to \mathbb{R}^\times$ maps the non-zero complex numbers onto the positive real numbers and its kernel is the subgroup S^1. Therefore $\mathbb{C}^\times/S^1 \simeq \mathbb{R}^\times_{>0}$.

2. The map $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is a surjective group homomorphism with kernel $\mathrm{SL}_n(\mathbb{R})$ so that $\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{R}) \simeq \mathbb{R}^\times$.

3. The map $\mathbb{Z} \to U_n$ defined by $k \mapsto e^{2\pi ik/n}$ is a surjective group homomorphism with kernel $n\mathbb{Z}$ so that $\mathbb{Z}/n\mathbb{Z} \simeq U_n$ is a cyclic group of order n. We often denote this quotient by $\mathbb{Z}_n$.

4. The map $\mathrm{sign} : S_n \to U_2$ is a surjective homomorphism with kernel A_n - the alternating group. Therefore $S_n/A_n \simeq U_2$ is cyclic.

We will discuss example (3) in detail in the next lecture.

2.11 Lecture 15: An example of quotient groups—modular arithmetic

In this lecture, we want to carefully study the quotient group $\mathbb{Z}/n\mathbb{Z}$ as it will be a prevalent example for the rest of the course.

Recall that since $\mathbb{Z}^+$ is an abelian group, every subgroup $n\mathbb{Z}$ is normal so that we can form the quotient group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. If we write

$$\overline{a} = a + n\mathbb{Z}$$

for the coset containing $a \in \mathbb{Z}$, then we see that $\overline{a} = \overline{b}$ if and only if $a - b \in n\mathbb{Z}$ if and only if $n|(a-b)$. Historically, two integers are called **congruent modulo** n if n divides their difference, or equivalently, if they have the same remainder when divided by n. We can use the division algorithm to show that the following n cosets together cover $\mathbb{Z}$:

$$n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}.$$

Therefore the quotient group $\mathbb{Z}_n$ has order n. We have seen in the last lecture that $\mathbb{Z}_n \simeq U_n$ so that $\mathbb{Z}_n$ is a cyclic group for all n. In fact, it is easy to see that $\overline{1}$ generates $\mathbb{Z}_n$. Our current goal is to find a complete list of generators for $\mathbb{Z}_n$. Our investigations will lead us to a discussion of the famous Euler φ-function. The reader may wish to review the material covered in Lecture 8, as it will play a heavy role here.

Recall that the greatest common divisor of two integers n and m is the positive generator of the cyclic subgroup $n\mathbb{Z}+m\mathbb{Z} = d\mathbb{Z}$ of $\mathbb{Z}^+$. We denote the greatest common divisor by the symbol (n, m), and we say that n and m are relatively prime if $(n, m) = 1$. We will need the following fact from number theory.

Lemma 2.11.1 *If $(n, m) = 1$ and $n|mk$, then $n|k$.*

Proof. Since $(n, m) = 1$, there exist integers $r, s \in \mathbb{Z}$ such that

$$nr + ms = 1.$$

Multiplying this equation by k on both sides yields

$$nkr + msk = k.$$

Now of course $n|nkr$ and $n|msk$ since $n|mk$ so that $n|k$. ■

The next theorem is the result we need to accomplish our goal.

Theorem 2.11.2 *Suppose that G is a cyclic group of order n generated by a. If $b \in G$ and $b = a^s$, then b generates a cyclic subgroup $\langle b \rangle$ of order $n/(n,s)$.*

Proof. Of course b generates a cyclic subgroup of G so that we need only verify that the order of b is $n/(n,s)$. Let m be the order of b and recall that m is the smallest positive integer such that $b^m = 1$. But $b^m = 1$ iff. $(a^s)^m = a^{ms} = 1$ so that $n|ms$. We want to find the smallest positive integer m such that $n|ms$. Let $d = (n,s)$ and note that we can find integers u and v such that

$$d = un + vs \Longleftrightarrow 1 = u(n/d) + v(s/d).$$

Note that both n/d and s/d are integers since $d = (n,s)$. It follows that n/d and s/d are relatively prime. We want to find the smallest positive integer m such that

$$\frac{ms}{n} = \frac{m(s/d)}{(n/d)} \text{ is an integer.}$$

From the lemma, we conclude that n/d must divide m so that the smallest such m is n/d. ■

Example 2.11.3 *1.* To find the order of $\overline{3}$ in $\mathbb{Z}_{12}$, we note that $\overline{3} = 3 \cdot \overline{1}$ and $(12,3) = 3$. Therefore the order of $\overline{3}$ is $12/3 = 4$.

Corollary 2.11.4 *If G is a cyclic group of order n generated by a, then $a^s \in G$ generates G if and only if n is relatively prime to s.* ■

If $n \in \mathbb{Z}$ is a positive integer, the **Euler φ-function** is defined by

$$\varphi(n) = \text{the number of integers less than } n \text{ that are relatively prime to } n.$$

For example, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$ and $\varphi(10) = 4$. The reader can show that $\varphi(p) = p - 1$ for all primes p. With this notation, we have the following corollary to the theorem above.

Corollary 2.11.5 *If G is a cyclic group of order n, the number of generators of G is $\varphi(n)$.* ■

We end this lecture by finding all subgroups of the cyclic group $\mathbb{Z}_{18}$.

Example 2.11.6 Recall that all subgroups of $\mathbb{Z}_{18}$ are cyclic so that we can simply list all the cyclic subgroups and be done. In what follows, we omit the bar notation so that 1 means $\overline{1}$ and so on.

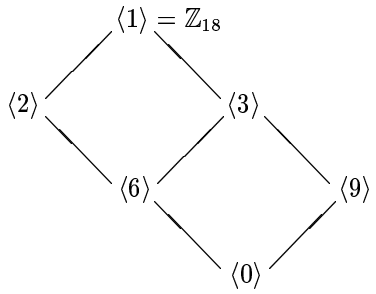
By the previous corollary, the elements $1, 5, 7, 11, 13$ and 17 are all generators of $\mathbb{Z}_{18}$. Now starting with 2,

$$\langle 2 \rangle = \{0, 2, 4, 6, 8, 10, 12, 14, 16\}$$

is of order 9 and has generators of the form $h2$ where h is relatively prime to 9. Namely $h = 1, 2, 4, 5, 7$ and 8 so that $h2 = 2, 4, 8, 10, 14$ and 16. The element 6 generates $\langle 6 \rangle = \{0, 6, 12\}$ and 12 also generates this group. We still have to check $3, 9$ and 15. We have

$$\langle 3 \rangle = \{0, 3, 6, 9, 12, 15\},$$

and 15 also generates this group since $15 = 5 \cdot 3$ and $(6, 5) = 1$. Finally, $\langle 9 \rangle = \{0, 9\}$. We end the example by drawing the **lattice diagram** for the subgroups of $\mathbb{Z}_{18}$.



Chapter 3

Vector Spaces

3.1 Lecture 16: Real and complex vector spaces

We assume the reader has some familiarity with real and complex vector spaces from an elementary linear algebra course. Our purpose here is to develop the same theory one sees in elementary linear algebra from the group theory point of view. To begin, let us denote the direct product of the additive group of real numbers with itself n times by $\mathbb{R}^n$ so that

$$\mathbb{R}^n = \underbrace{\mathbb{R}^+ \times \cdots \times \mathbb{R}^+}_{n}.$$

Historically, elements of this abelian group are called **vectors** and the group operation:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

is called **vector addition**. Notice that the set $\mathbb{R}^n$ is closed under another operation called **scalar multiplication**. Namely, if $\alpha \in \mathbb{R}$ is a real number, then we can define the product

$$\alpha(a_1, \ldots, a_n) = (\alpha a_1, \ldots, \alpha a_n)$$

for all $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. The reader can easily verify that this operation satisfies the following familiar properties for all $a, b \in \mathbb{R}^n$ and all $\alpha, \beta \in \mathbb{R}$.

1. $(\alpha\beta)a = \alpha(\beta a)$.
2. $(\alpha + \beta)a = \alpha a + \beta a$.

3. $\alpha(a+b) = \alpha a + \alpha b$.

4. $1a = a$.

All of this motivates the following definition.

Definition 3.1.1 (Real vector space) *A* **real vector space** *is an abelian group $V = (V, +)$ together with a function $\mathbb{R} \times V \to V$ written $(\alpha, v) \mapsto \alpha v$ called* **scalar multiplication** *such that for all $\alpha, \beta \in \mathbb{R}$ and all $u, v \in V$, we have*

1. $(\alpha\beta)v = \alpha(\beta v)$.

2. $(\alpha + \beta)v = \alpha v + \beta v$.

3. $\alpha(u + v) = \alpha u + \alpha v$.

4. $1v = v$.

The elements of the group V are called **vectors** *and the elements of the real numbers $\mathbb{R}$ are called scalars. In particular the additive identity $0 \in V$ of the group V is called the* **zero vector**.

Example 3.1.2

1. The direct product $\mathbb{R}^n$ is a real vector space with the component wise scalar multiplication defined above.

2. Let $V = C[a, b]$ be the abelian group of continuous functions on the interval $[a, b]$. We recall that if $f, g \in C[a, b]$, then by definition

$$(f + g)(x) = f(x) + g(x)$$

and if $\alpha \in \mathbb{R}$ is a scalar we define

$$(\alpha f)(x) = \alpha f(x).$$

It is an easy exercise to show that $V = C[a, b]$ is then a real vector space.

3. The set P_n of polynomials of degree less than or equal to n is a real vector space under the usual addition and scalar multiplication of polynomials.

Since a vector space V is also an abelian group, we can look at the subgroups of the group V. However, we should pay special attention to those subgroups that are compatible with the operation of scalar multiplication. We make the following definition.

Definition 3.1.3 (Subspace) *A subset W of a real vector space V is a* **subspace** *if*

1. *W is a subgroup of V.*
2. *$\alpha w \in W$ for all $\alpha \in \mathbb{R}$ and all $w \in W$.*

The second condition is often referred to by saying W is **stable** under scalar multiplication.

Example 3.1.4

1. If A is an $m \times n$ matrix, the set of solutions to the homogeneous system of linear equations $AX = 0$ is a subspace of $\mathbb{R}^n$.
2. The set $W = \{f \in C[0,1] : f(1/2) = 0\}$ is a subspace of the vector space of continuous functions on the unit interval.
3. P_{n-1} is a subspace of P_n.

If we go back to the very beginning of the lecture and replace the scalars $\mathbb{R}$ with complex numbers $\mathbb{C}$, then we have defined the notion of a complex vector space. We conclude this lecture by listing some elementary properties of real and complex vector spaces. The proofs are left as exercises.

Proposition 3.1.5 *If V is a real or complex vector space, then*

1. *$0_{\mathbb{R}} v = 0_V$ for all $v \in V$. (Here you may replace $\mathbb{R}$ with $\mathbb{C}$.)*
2. *$\alpha 0_V = 0_V$ for all $\alpha \in \mathbb{R}$. (Here again you may replace $\mathbb{R}$ with $\mathbb{C}$.)*
3. *$(-1)v = -v$ for all $v \in V$.*

■

3.2 Lecture 17: Abstract fields

You do not have to study real and complex vector spaces for very long before you realize that there is no difference between the two objects. That is, it is not important whether or not the scalars are real numbers or complex numbers (or even numbers at all!); all that matters is that the scalars have the same arithmetic properties that the real and complex numbers have. These properties are exactly what motivate the definition of an abstract field which we now give.

Definition 3.2.1 (Field) *A set F with two binary operations $+$ and $\cdot$ is called a* **field** *if*

1. *$(F, +)$ is an abelian group. The additive identity is written* 0.

2. *$(F^\times, \cdot)$ is an abelian group ($F^\times = F \setminus \{0\}$ is the set of non-zero elements of F). The multiplicative identity is written* 1.

3. *For all a, b and c in F, the following* **distributive law** *holds:*

$$a(b + c) = ab + ac.$$

Note that $\mathbb{R}$ and $\mathbb{C}$ are fields. The set of rational numbers $\mathbb{Q}$ is also a field. The integers $\mathbb{Z}$ do not form a field since the non-zero integers are not a group under multiplication. The fields $\mathbb{C}, \mathbb{R}$ and $\mathbb{Q}$ are familiar to the reader. In fact, $\mathbb{R}$ and $\mathbb{Q}$ are examples of subfields of $\mathbb{C}$. To be more precise, a **subfield of** $\mathbb{C}$ is a subset $F \subseteq \mathbb{C}$ that is a subgroup of $(\mathbb{C}, +)$ and $F^\times$ is a subgroup of $\mathbb{C}^\times$. It may come as a surprise that there are fields that are not subfields of $\mathbb{C}$. We will begin looking for them in the group $\mathbb{Z}_p$, where $p \in \mathbb{Z}$ is a prime number. We will need the following lemma.

Lemma 3.2.2 *Let $n \in \mathbb{Z}$ be a fixed positive integer and let $\overline{a}$ denote the coset of $n\mathbb{Z}$ in $\mathbb{Z}$ that contains the integer a. If $a, b \in Z$, then the operation*

$$\overline{a}\overline{b} = \overline{ab}$$

is well defined on $\mathbb{Z}_n$. Moreover, if $c \in \mathbb{Z}$, then the distributive law

$$\overline{a}(\overline{b} + \overline{c}) = \overline{a(b + c)}$$

holds.

Proof. Suppose that $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$ so that $a' = a + kn$ and $b' = b + jn$. Then we compute

$$a'b' = (a + kn)(b + jn) = ab + (kb + ja + jkn)n$$

so that $\overline{a'b'} = \overline{ab}$ and the operation is well defined. We have seen that addition of cosets is well defined in $\mathbb{Z}_n$ and consequently all operations in the distributive law are independent of the choice of representatives. ∎

The previous lemma is almost enough to show that $\mathbb{Z}_n$ is a field. That is, we know that $(\mathbb{Z}_n, +)$ is an abelian group and the lemma implies that $\mathbb{Z}_n$ has a multiplication that distributes over the addition in this group. The only thing that is possibly missing is the existence of multiplicative inverses. The following example gives a hint about what is going on here.

Example 3.2.3 Consider the group $(\mathbb{Z}_6, +)$. Note that $\overline{2}, \overline{3} \neq \overline{0}$ but $\overline{23} = \overline{0}$ so that $\overline{2}$ cannot be invertible. As we shall see, this behavior cannot happen if n is a prime number.

Theorem 3.2.4 *If $p \in \mathbb{Z}$ is a prime number and $\overline{a} \in \mathbb{Z}_p$ satisfies $\overline{a} \neq \overline{0}$ (i.e. p does not divide a), then there exists an element $\overline{b} \in \mathbb{Z}_p$ such that $\overline{ab} = \overline{1}$ and hence $\mathbb{Z}_p$ is a field.*

Proof. Since the order of $\mathbb{Z}_p$ is p, the set

$$\{\overline{a}^k : k \in \mathbb{Z}\}$$

must be finite so that $\overline{a}^k = \overline{a}^m$ for some $k < m$. Now,

$$\overline{a}^k = \overline{a}^m \iff p | a^m - a^k \iff p | a^k(a^{m-k} - 1).$$

By hypothesis, p does not divide a so that p does not divide a^k. It then follows that p divides $(a^{m-k} - 1)$ so that $\overline{a}^{m-k} = \overline{1}$ and $m - k > 0$. If we define $b = a^{m-k-1}$, then $\overline{a}\overline{b} = \overline{1}$. It follows that the non-zero elements of $\mathbb{Z}_p$ are all invertible and of course $\overline{1}$ is a multiplicative identity. The associative property follows directly from that in $\mathbb{Z}$ so that $\mathbb{Z}_p^\times$ is a group under multiplication of cosets. We have seen that the distributive law holds so that $\mathbb{Z}_p$ is a field. ∎

Fields have a rich enough algebraic structure to do most mathematical operations. For example, one can look at matrices with entries from a field F since the definitions of matrix addition and multiplication are still valid if the entries come from a field - just replace the addition and multiplication of real or complex numbers with the addition and multiplication in the field F.

Example 3.2.5 Let F be an arbitrary field and define

$$\mathrm{GL}_n(F) = \{n \times n \text{ matrices } A \text{ over } F \text{ such that } \det A \neq 0\}.$$

Then $\mathrm{GL}_n(F)$ is a group called the **general linear group over** F. In particular we can look at the very interesting groups $\mathrm{GL}_n(\mathbb{Z}_p)$. Note that $\mathrm{GL}_n(\mathbb{Z}_p)$ is a new example a family of finite groups. It is an interesting exercise to compute the order of $\mathrm{GL}_n(\mathbb{Z}_p)$.

Note that in the field $\mathbb{Z}_p$, we have $\overbrace{\overline{1} + \cdots + \overline{1}}^{p} = \overline{0}$ (this never happens in a subfield of $\mathbb{C}$!). Because of this, we say $\mathbb{Z}_p$ has characteristic p. To be more precise, we make the following definition.

Definition 3.2.6 (Characteristic) *If F is a field, the smallest positive integer n satisfying $n \cdot 1_F = 0_F$ is called the* **characteristic of** *F. If no such positive integer exists, we say F has* **characteristic zero**.

For example, the field $\mathbb{C}$ and all of its subfields have characteristic zero. The field $\mathbb{Z}_p$ has characteristic p. We end this lecture with the definitions of a vector space over an arbitrary field and a subspace. The reader should compare these definitions with the ones given for real and complex vector spaces in the previous lecture.

Definition 3.2.7 (Vector space) *Let F be a arbitrary field. An abelian group $V = (V, +)$ is called an F-***vector space** *if there exists a function $F \times V \to V$ written $(\alpha, v) \mapsto \alpha v$ called* **scalar multiplication** *such that for all $\alpha, \beta \in F$ and all $u, v \in V$, we have*

1. $(\alpha\beta)v = \alpha(\beta v)$.
2. $(\alpha + \beta)v = \alpha v + \beta v$.
3. $\alpha(u + v) = \alpha u + \alpha v$.
4. $1_F v = v$.

The elements of the group V are called **vectors** *and the elements of the field F are called scalars. In particular the additive identity $0 \in V$ of the group V is called the* **zero vector**.

Example 3.2.8

1. The direct product group F^n is a vector space over F with the obvious scalar multiplication.
2. The set of $m \times n$ matrices with entries in F is a vector space over F.
3. The set of functions $f : \mathbb{R} \to F$ is an F-vector space under pointwise addition and scalar multiplication.

Definition 3.2.9 (Subspace) *A subset W of a vector space V over a field F is a* **subspace** *if*

1. *W is a subgroup of V.*
2. *$\alpha w \in W$ for all $\alpha \in F$ and all $w \in W$.*

Therefore subspaces are just the subgroups of $(V, +)$ that are closed under scalar multiplication. Our previous examples of subspaces all carry over to the case of an arbitrary field.

Example 3.2.10

1. If A is an $m \times n$ matrix with entries in F, the set of solutions to the homogeneous system of linear equations $AX = 0$ is a subspace of F^n.

2. The set $W = \{f : [0,1] \to F : f(1/2) = 0\}$ is a subspace of the vector space of F-valued functions on the unit interval.

Our work in group theory already gives us an idea of what we should mean by an isomorphism of vector spaces over F. After all, each such vector space is an abelian group so that if we are to call two vector spaces V and V' isomorphic, they should at the very least be isomorphic as abelian groups. That is, there should be a bijective group homomorphism $\varphi : V \to V'$. However, the map φ should also know about the operation of scalar multiplication since V and V' are vector spaces. All of this motivates the following definition.

Definition 3.2.11 (Isomorphism) *Let V and V' be two vector spaces over the same field F. A bijective group homomorphism $\varphi : V \to V'$ is called an* **isomorphism** *if*

$$\varphi(\alpha v) = \alpha\varphi(v)$$

for all $\alpha \in F$ and all $v \in V$.

Example 3.2.12 The abelian group $\mathbb{C}^+$ is a real vector space since $\mathbb{R} \subset \mathbb{C}$ and hence we can define scalar multiplication αz ($\alpha \in \mathbb{R}, z \in \mathbb{C}$) as ordinary multiplication of complex numbers. We leave it as an exercise for the reader to show that the map $\varphi : \mathbb{R}^2 \to \mathbb{C}$ defined by $\varphi(x,y) = x + iy$ is an isomorphism of real vector spaces.

3.3 Lecture 18: Bases and dimension

It turns out that every element of a (usually infinite) vector space can be described in terms of a finite subset of vectors in V. The goal of the current lecture is to introduce the notions of span, linear independence and basis which, together, give this description. Before we begin, we remark that some of our calculations and notations will depend on the ordering of a set. Recall from elementary set theory that sets are unordered collections of elements. Therefore the two sets $\{a,b,c\}$ and $\{b,c,a\}$ are exactly the same. In linear algebra, we will often work with sets of vectors that are in some specific order, and we want our notation to reflect that this order matters. Therefore we will replace the curly set brackets { and } with round parentheses (and) when we want to fix the order of the sets. In this notation, the ordered sets (a,b,c) and (b,c,a) are **not** the same. We begin with the definition of linear combination.

Definition 3.3.1 *Let V be a vector space over a field F and let $(v_1, \ldots, v_n)$ be an ordered set of vectors in V. Then a* **linear combination of the** v_i *is a vector $v \in V$ of the form*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

where each $\alpha_i \in F$ is a scalar. The scalars α_i are called the **coefficients** *of the linear combination. If $S = (v_1, v_2, \ldots, v_n)$ is an ordered set of vectors in V, the set of all linear combinations of the vectors in S is called the* **span of** S *and is denoted* $\mathrm{Span}(S)$. *Therefore*

$$\mathrm{Span}(S) = \{v : v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, \alpha_i \in F\}.$$

If $S \subset V$ and $\mathrm{Span}(S) = V$, then we say S **spans** V.

Example 3.3.2 (Important!) Suppose that A is an $m \times n$ matrix with entries in F. The the matrix equation $AX = B$ exhibits the vector $B \in F^m$ as a linear combination of the columns of the matrix A where the coefficients are the entries of the vector $X \in F^n$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{n1} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

It is an easy exercise to show that $\mathrm{Span}(S)$ is a subspace of V (do it!). The following states that this is the smallest subspace of V containing the set S.

Proposition 3.3.3 *If W is a subspace of a vector space V and $S \subseteq W$, then* $\mathrm{Span}(S) \subseteq W$.

Proof. Suppose that $v \in \mathrm{Span}(S)$ so that $v = \sum_{i=1}^{n} \alpha_i v_i$ where $v_i \in S$ and $\alpha_i \in F$. Now, W is a subspace of V and hence W is closed under scalar multiplication and vector addition. But $v_i \in S$ so that $v_i \in W$ and hence $v \in W$. ∎

If $S \subset V$ spans V, then in some sense this means that there are "enough" vectors in S to describe all of V. We now turn our attention to the related notion of "no overlap". Specifically, we make the following definition.

Definition 3.3.4 *A subset $S = \{v_1, v_2, \ldots, v_n\}$ is called* **linearly independent** *if the equation*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

implies that $\alpha_i = 0$ for all $i = 1, 2, \ldots, n$. If S is not linearly independent, we say it is **dependent**.

In other words, a (finite) set S is linearly independent if the only linear combination of the $v_i \in S$ that gives the zero vector is the **trivial linear combination**. The next proposition explains the term "no overlap" used above.

Proposition 3.3.5 *If $S = \{v_1, v_2, \ldots, v_n\}$ is a linearly independent subset of a vector space V, then for every $v \in \text{Span}(S)$, there exist unique scalars $\alpha_i \in F$ such that*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.$$

Proof. The existence of the α_i follows immediately from the definition of the span of S. Therefore suppose that

$$v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

as well so that subtracting gives

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = 0.$$

Therefore $\alpha_i - \beta_i = 0$ for all i since S is linearly independent and hence $\alpha_i = \beta_i$ for all i. ■

It follows from this proposition and the remarks above that if S is a linearly independent spanning subset of a vector space V, then every vector $v \in V$ can be expressed uniquely as a linear combination of the vectors in S. The spanning part says that S is "big enough" to describe all vectors in V and the linearly independent part says that there is "no overlap" because each such representation is unique. This idea is important enough to have a name.

Definition 3.3.6 (Basis) *If V is a vector space over a field F, a subset $S \subseteq V$ is called a* **basis for** *V if S is a linearly independent spanning set.*

Example 3.3.7 Let $V = F^n$ be the vector space of column vectors over F and let e_i denote the column vector with a 1 in the i^{th} position and zeros elsewhere. Then the set $(e_1, e_2, \ldots, e_n)$ is a basis for F^n called the **standard basis**.

Our last goal of this lecture is to show that any two bases of a vector space V have the same cardinality. We will focus on the cases in which V has a finite basis so that we will show that any two bases of V have the same number of elements. It turns out that this number is the only invariant among vector spaces in the sense that any two vector spaces with bases of the same size are necessarily isomorphic. We will need three technical lemmas. The proofs are all very straightforward applications of the relevant definitions.

Lemma 3.3.8 *Let $L = (v_1, \ldots, v_r)$ be a linearly independent ordered set of vectors in a vector space V. If $v \in V$ and L' is the ordered set $L' = (L, v)$ obtained from L by inserting v at the end, then L' is linearly independent if and only if $v \notin \mathrm{Span}(L)$.*

Sketch of proof. ($\Longrightarrow$) If $v \in \mathrm{Span}(L)$, then

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r$$

so that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r + (-1)v = 0$$

and hence L' is dependent.
($\Longleftarrow$) If L' is dependent, then we must have

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r + \beta v = 0$$

with not all the coefficients zero. But $\beta \neq 0$, otherwise L would be dependent. Therefore you can solve this equation for v showing that $v \in \mathrm{Span}(S)$. ■

Lemma 3.3.9 *If S is an ordered set of vectors and $v \in V$ is any vector, let $S' = (S, v)$. Then $\mathrm{Span}(S) = \mathrm{Span}(S')$ if and only if $v \in \mathrm{Span}(S)$.* ■

Definition 3.3.10 *A vector space V is called* **finite dimensional** *if there is a finite subset $S \subset V$ such that $\mathrm{Span}(S) = V$.*

Lemma 3.3.11 *Any finite set S that spans V contains a basis. In particular, every finite dimensional vector space V has a basis.*

Proof. We induct on the number of elements in S. Suppose that $S = (v_1, \ldots, v_r)$ is dependent so that there is a nontrivial linear combination

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = 0.$$

Without loss of generality, we may assume that $\alpha_r \neq 0$ so that we can solve the last equation for v_r showing that $v_r \in \mathrm{Span}(v_1, \ldots, v_{r-1})$. The last lemma then implies that $\mathrm{Span}(v_1, \ldots, v_{r-1}) = V$ and by induction $(v_1, \ldots, v_{r-1})$ contains a basis. ■

We remark here that the above "induction" argument is slightly incomplete. Namely, we never verified the case $r = 1$, and we see that the proof breaks down in this case! To fix this nasty situation, we make the following conventions:

1. The empty set is a linearly independent set.

2. The span of the empty set is the zero subspace.

Lemma 3.3.12 *Every linearly independent subset L of a finite dimensional vector space V can be extended to a basis for V.*

Proof. If $\text{Span}(L) = V$, we're done. Otherwise there exists a vector $v \in V$ with $v \notin \text{Span}(L)$ and the set $L' = (L, v)$ is linearly independent by a lemma above. Since V is finite dimensional, continuing this process will eventually produce a spanning set and hence a basis. ∎

Lemma 3.3.13 *Suppose that S and L are finite subsets of a vector space V. If S spans V and L is linearly independent, then $|S| \geq |L|$ where $|S|$ denotes the number of elements in the set S.*

Proof. Suppose that $S = (v_1, \ldots, v_m)$ and $L = (w_1, \ldots, w_n)$. Since $\text{Span}(S) = V$, for each $1 \leq j \leq n$, we can find scalars $\alpha_{ij} \in F$ such that

$$\alpha_{1j}v_1 + \cdots + \alpha_{mj}v_m = w_j.$$

If we let $u = \beta_1 w_1 + \cdots + \beta_n w_n$, then substituting gives

$$u = \sum_{i,j} \beta_j \alpha_{ij} v_i.$$

The coefficient of v_i in this sum is $\sum_j \beta_j \alpha_{ij}$ and if this coefficient is zero for all i, then $u = 0$. Therefore to find a relation among the w_j, it suffices to solve the homogeneous system $\sum_j \alpha_{ij} x_j = 0$ of m equations in n unknowns. If $m < n$, such a system always has a solution and hence L is dependent if $m < n$. The result now follows by contraposition. ∎

We can now state and prove the final goal of this lecture.

Theorem 3.3.14 *Suppose that V is a finite dimensional vector space over a filed F. Then any two bases for V have the same number of elements.*

Proof. If B and B' are two basis for V, then the previous lemma implies that $|B| \geq |B'|$ since B spans V and B' is linearly independent. Interchanging the roles of B and B' shows that $|B'| \geq |B|$ so that $|B| = |B'|$. ∎

Definition 3.3.15 (Dimension) *If V is a finite dimensional vector space, the* **dimension of** *V is the number of vectors in a basis for V.*

The previous theorem implies that this number is independent of the particular basis used.

3.4 Lecture 19: Computations with bases

The main purpose for introducing bases in a vector space is to provide a method of performing computations. The primary goals of the current lecture are to illustrate these computations as well as investigate the changes in the computations when we change basis. An important consequence of our work will be a result that states exactly how many bases a given vector space has. We will begin with a discussion of coordinate vectors in the vector space F^n.

Suppose that we are given an ordered basis $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ for F^n so that every vector $v \in F^n$ can be expressed as a linear combination of the vectors v_i is a unique way:

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

The scalars α_j are called the **coordinates of the vector v with respect to the basis $\mathcal{B}$** and the vector $X = (\alpha_1, \ldots, \alpha_n)^T$ is called the **coordinate vector** of v with respect to $\mathcal{B}$. If we let $[\mathcal{B}]$ denote the $n \times n$ matrix whose j^{th} column is the vector v_j, then we can write the vector v as a matrix product $[\mathcal{B}]X$. In this notation, if we are given a vector $Y = (y_1, \ldots, y_n) \in F^n$ and we want to find the coordinates of Y with respect to $\mathcal{B}$, we must solve the matrix equation

$$[\mathcal{B}]X = Y$$

for X. The following proposition states that this equation is easy to solve.

Proposition 3.4.1 *If A is an $n \times n$ matrix over F, then A is invertible if and only if the columns of A form a basis for F^n.*

Proof. If we let v_j denote the j^{th} column of A, then for any vector $X = (x_1, \ldots, x_n) \in F^n$, the matrix product $AX = x_1 v_1 + \cdots + x_n v_n$ is a linear combination of the columns v_j. Therefore the columns of A are linearly independent if and only if the matrix equation $AX = 0$ has only the trivial solution $X = 0$ which is true if and only if A is invertible. Finally, since $\dim F^n = n$, a linearly independent set with n vectors is necessarily a basis. ∎

It follows immediately from this proposition that the coordinate vector of a vector $Y \in F^n$ with respect to the basis $\mathcal{B}$ is $X = [\mathcal{B}]^{-1}Y$. All of these computations are possible because the vectors v_i in $\mathcal{B}$ are given explicitly as elements of F^n. We now turn our attention to the case of an arbitrary vector space V over F. In this case, it is not possible to represent a basis $\mathcal{B} = (v_1, \ldots, v_n)$ as a matrix over F. However, we can formally manipulate this **hypervector** as if it were a matrix. To

this end, if $X = (x_1, \ldots, x_n)^T$ is a matrix of scalars, we define the product

$$(v_1, \ldots, v_n) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 v_1 + \cdots + x_n v_n.$$

We can abbreviate this notation and write the linear combination simply as $\mathcal{B}X$. Moreover, if A is an $n \times m$ matrix over F, we define the product

$$(v_1, \ldots, v_n)A = (w_1, \ldots, w_m)$$

where

$$w_j = a_{1j}v_1 + \cdots + a_{nj}v_n.$$

In this notation, each w_j is a linear combination of $(v_1, \ldots, v_n)$ and the scalars in the linear combination form the j^{th} column of the matrix A. We call this operation **multiplication by the hypervector $\mathcal{B}$.** Recall that our goal is to be able to find the coordinate vector of a given vector $v \in V$ for a given basis $\mathcal{B}$. The following theorem will allow us to achieve this goal.

Theorem 3.4.2 *If V is a finite dimensional vector space over a field F and $\mathcal{B} = (v_1, \ldots, v_n)$ is a basis for V, then the map $\psi : F^n \to V$ defined by*

$$\psi : X \mapsto \mathcal{B}X$$

is a vector space isomorphism.

Sketch of Proof. We will leave the verification that ψ is a group homomorphism to the reader. Since $\mathcal{B}$ is a basis for V, $\mathcal{B}$ spans V and hence the map ψ is surjective. Moreover, $\psi(X) = 0$ iff. $\mathcal{B}X = 0$ iff. $X = 0$ since $\mathcal{B}$ is linearly independent so that ψ is injective. ∎

Corollary 3.4.3 *If V and W are vector spaces over F with $\dim V = \dim W$, then V is isomorphic to W.* ∎

It follows that we can study all vector spaces by studying the spaces F^n.

We conclude this lecture with a discussion on changing basis. Suppose that V is a vector space over F and $\mathcal{B}$ and $\mathcal{B}'$ are two bases for V. We want to find out how the two bases are related to one another. In particular, we want to be able to find the coordinates of a vector v with respect to $\mathcal{B}'$ if

we know the coordinates with respect to $\mathcal{B}$. To keep the notations as simple as possible, we will refer to $\mathcal{B}$ as the "old" basis and $\mathcal{B}'$ as the "new" basis. We begin by noting that since the new basis spans V, every vector $v_i \in \mathcal{B}$ can be written as a linear combination of the new basis $\mathcal{B}' = (v_1', \ldots, v_n')$. Therefore we can find an $n \times n$ matrix of scalars P such that

$$\mathcal{B}'P = \mathcal{B}.$$

The matrix P is called the **change of basis matrix**. Note that the j^{th} column of P is the coordinate vector of $v_j \in \mathcal{B}$ with respect to the new basis $\mathcal{B}'$.

Lemma 3.4.4 *The change of basis matrix P is invertible.*

Proof. If we interchange the roles of $\mathcal{B}$ and $\mathcal{B}'$, we have an $n \times n$ matrix P' that satisfies $\mathcal{B}P' = \mathcal{B}'$ and hence

$$\mathcal{B} = \mathcal{B}'P = (\mathcal{B}P')P = \mathcal{B}(P'P).$$

This last expression gives each v_i as a linear combination of the vectors in $\mathcal{B}$, and the entries in the matrix $P'P$ are the coefficients. But $\mathcal{B}$ is a basis so that we can only write $v_i = v_i$ so that the matrix $P'P$ is the identity matrix so that P is invertible as claimed. ∎

Now, if $v \in V$ has the coordinate vector X with respect to the basis $\mathcal{B}$, then $v = \mathcal{B}X$. It follows immediately that $v = \mathcal{B}'PX$ so that PX is the coordinate vector of v with respect to $\mathcal{B}'$. This is why P is called the change of basis matrix: multiplication on the left by P changes from the old basis to the new basis. We remark again that the columns of P are the coordinates of the old basis vectors with respect to the new basis. Summarizing, we have an invertible matrix P that simultaneously satisfies the equations

$$\mathcal{B} = \mathcal{B}'P \qquad \text{and} \qquad PX = X'$$

where X and X' denote the coordinates of a vector $v \in V$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$ respectively. In particular, if $V = F^n$ and $\mathcal{B}$ is the standard basis, then we have $I_n = [\mathcal{B}']P$ so that $P = [\mathcal{B}']^{-1}$.

In the above discussion, we could have just as well started with a single basis $\mathcal{B}$ for V and an invertible matrix $P \in \mathrm{GL}_n(F)$ to form a new basis $\mathcal{B}' = \mathcal{B}P^{-1}$. This equation shows immediately that $\mathcal{B}'$ is a basis for V since it shows that each $v_i \in \mathcal{B}$ is in the span of $\mathcal{B}'$ and that $\mathcal{B}'$ has exactly $n = \dim V$ elements. All of this fits together to prove the following corollary whose precise proof we leave as an exercise for the reader.

Corollary 3.4.5 *If $\mathcal{B}$ is a basis for a finite dimensional vector space V, then any other basis has the form $\mathcal{B}' = \mathcal{B}P^{-1}$ where $P \in \mathrm{GL}_n(F)$ and therefore the number of distinct bases for V is the order of the group $\mathrm{GL}_n(F)$.* ■

We conclude this lecture with an application of the previous result. Namely, we will compute the order of the group $\mathrm{GL}_n(\mathbb{F}_p)$ where $\mathbb{F}_p$ denote the field of p elements.

Proposition 3.4.6 *If $\mathbb{F}_p$ is the field of p elements, then the order of the group $\mathrm{GL}_n(\mathbb{F}_p)$ is*

$$(p^n - 1)(p^n - p)(p^n - p^2) \cdots (p^n - p^{n-1}).$$

Proof. By the previous corollary, it suffices to count the number of bases for the vector space $\mathbb{F}_p^n$. Since the zero vector can never belong to a basis, we have $p^n - 1$ choices for the first basis vector v_1. Having made this choice, we can pick v_2 to be any vector that is not a scalar multiple of v_1, and there are exactly p such multiples. Therefore we have $p^n - p$ choices for v_2. Similarly, we can pick v_3 so that it is not a linear combination of v_1 and v_2. There are exactly p^2 such linear combinations so that we have $p^n - p^2$ choices for v_3. Continuing we see that we will have $p^n - p^{n-1}$ choices for v_n and since all of these choices are independent of one another, the result follows from the multiplication principle. ■

Chapter 4

Linear Transformations

4.1 Lecture 20: The rank-nullity theorem

We have already seen that an isomorphism between two vector spaces is an isomorphism of the underlying abelian groups that respects scalar multiplication. In this lecture we will begin to investigate such mappings that are not necessarily bijective. Here is the main definition.

Definition 4.1.1 (Linear transformation) *If V and W are vector spaces over a field F, then a* **linear transformation from V to W** *is a group homomorphism $T : V \to W$ that satisfies*

$$T(\alpha v) = \alpha T(v)$$

for all $v \in V$ and all $\alpha \in F$.

It would not be inappropriate to call linear transformations simply "vector space homomorphisms" since they are precisely the functions between two vector spaces that preserve the vector space structure. As we will see later in our course, there are many structures that algebraists can impose on a set (e.g. groups, vector spaces, rings, modules, algebras ...), and we will always refer to the functions that preserve these structures as homomorphisms. For vector spaces however, it is traditional to call such maps linear transformations and we will follow this tradition. Every linear transformation $T : V \to W$ is, at the very least, a group homomorphism so that we can speak of the the two subgroups $\ker T$ and $\operatorname{im} T$. The following proposition states that these subgroups are in fact subspaces.

Proposition 4.1.2 *If $T : V \to W$ is a linear transformation, then* $\ker T$ *is a subspace of V and* $\mathrm{im}\, T$ *is a subspace of W.*

Proof. Exercise. ∎

Here are some examples of linear transformations. As we will soon see, the first example is really the only one!

Example 4.1.3 *1.* Let A be an $m \times n$ matrix with entries in F. Then A defines a linear transformation $T_A : F^n \to F^m$ by

$$T_A(X) = AX.$$

The linear transformation properties follow directly from the familiar properties of matrix multiplication. We will call the linear transformation induced by a matrix A "left multiplication by A".

2. Let P_n denote the vector space of polynomials over $\mathbb{C}$ with degree less than or equal to n. The map $\frac{d}{dx} : P_n \to P_{n-1}$ defined by

$$\frac{d}{dx}(f) = \frac{df}{dx}$$

is a linear transformation. Indeed, the linearity follows immediately from two familiar properties of the derivative.

3. If $V = C[0,1]$ is the vector space of continuous real valued functions on the unit interval, the map $T : V \to \mathbb{R}$ defined by

$$T(f) = f(1/2)$$

is a linear transformation.

Let us expand on the first example slightly. If A is an $m \times n$ matrix with entries in F, then the kernel of the linear transformation $T_A : F^n \to F^m$ is the set of all $X \in F^n$ such that $AX = 0$. In elementary linear algebra courses, this kernel is usually referred to as the solution space of the homogeneous system of equations $AX = 0$. Of course it is a subspace of F^n. If $B \in F^m$ is a non-zero vector, the system of equations $AX = B$ has a solution if and only if $B \in \mathrm{im}\, T_A$. If X_p is a solution to $AX = B$, the reader can verify that every element of the coset $X_p + \ker T_A$ is also a solution to

$AX = B$. This is simply the familiar fact that if you add a solution to the homogeneous system $AX = 0$ to any solution of $AX = B$, you get another solution of $AX = B$. Therefore we see that the set of all solutions to $AX = B$ is simply the coset of $\ker T_A$ containing any particular solution. Our final goal of this lecture is to investigate the relationship between the dimensions of the subspaces $\ker T$ and $\operatorname{im} T$ for a given linear map $T : V \to W$. The reader should note carefully how the facts about bases and linear independence are used in the proof; the methods are more than typical. Before we state the result, we remark that the dimension of $\operatorname{im} T$ is usually referred to as the **rank of** T and the dimension of $\ker T$ is usually called the **nullity of** T.

Theorem 4.1.4 (Rank-Nullity Theorem) *Let V and W be vector spaces over a field F with V finite dimensional. If $T : V \to W$ is a linear transformation, then*

$$\dim V = \dim \ker T + \dim \operatorname{im} T.$$

Proof. Let $\dim V = n$ and choose a basis $(v_1, \ldots, v_k)$ for $\ker T$. We can extend this set to a basis $\mathcal{B} = (v_1, \ldots, v_k, u_1, \ldots, u_{n-k})$ for V. We need to show that $\dim \operatorname{im} T = n - k$. To do this, it suffices to exhibit a basis for $\operatorname{im} T$ with $n - k$ vectors in it. Let us define $w_j = T(u_j)$ for $j = 1, 2, \ldots, n - k$. Clearly each $w_j \in \operatorname{im} T$. We claim that the set $(w_1, \ldots, w_{n-k})$ is a basis for $\operatorname{im} T$. If $w \in \operatorname{im} T$, then $w = T(v)$ for some $v \in V$ by definition. Since $\mathcal{B}$ is a basis for V, we can find scalars α_i and β_j such that

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 u_1 + \cdots + \beta_{n-k} u_{n-k}$$

and since $T(v_i) = 0$ for all $i = 1, \ldots, k$ we have

$$\begin{aligned} w = T(v) &= 0 + \cdots + 0 + \beta_1 T(u_1) + \cdots + \beta_{n-k} T(u_{n-k}) \\ &= \beta_1 w_1 + \cdots + \beta_{n-k} w_{n-k} \end{aligned}$$

so that the w_j span $\operatorname{im} T$. Now, if

$$\beta_1 w_1 + \cdots + \beta_{n-k} w_{n-k} = 0,$$

then we note that the element

$$v = \beta_1 u_1 + \cdots + \beta_{n-k} u_{n-k} \in V$$

belongs to the kernel of T (why?). Therefore we can find scalars α_i such that

$$\beta_1 u_1 + \cdots + \beta_{n-k} u_{n-k} = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

or

$$-\alpha_1 v_1 - \cdots - \alpha_k v_k + \beta_1 u_1 + \cdots + \beta_{n-k} u_{n-k} = 0.$$

But $\mathcal{B}$ is a basis for V so that we must have $\alpha_i = 0$ and $\beta_j = 0$ for all i and j so that $(w_1, \ldots, w_{n-k})$ is linearly independent and the proof is complete. ■

We conclude this lecture with a remark on the similarity of the result in the Rank-Nullity Theorem and the counting formula for groups. Namely, we recall if $\varphi : G \to H$ is a homomorphism between two finite groups, then

$$|G| = |\ker \varphi||\operatorname{im} \varphi|.$$

4.2 Lecture 21: The matrix of a linear transformation

The goal of the current lecture is to show that the example of left multiplication by a matrix actually describes all linear transformations on finite dimensional vector spaces once you choose bases for the spaces. We will begin with the spaces F^n.

Suppose that $T : F^n \to F^m$ is a linear map and consider the images of the standard basis vectors $T(e_j) \in F^m$. For each j, we can find scalars $a_{ij} \in F$ such that

$$T(e_j) = a_{1j} e_1 + \cdots + a_{mj} e_m.$$

Consequently, we can define an $m \times n$ matrix $A = (a_{ij})$ whose j^{th} column is the coordinate vector of $T(e_j)$. Now, if $X \in F^n$ is any vector and we write $X = e_1 x_1 + \cdots + e_n x_n$, then we have

$$T(X) = \sum_{j=1}^{n} T(e_j) x_j = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = AX$$

so that $T = T_A$.

Example 4.2.1 If $T : \mathbb{R}^2 \to \mathbb{R}^3$ satisfies $T(e_1) = (1, 2, 3)^T$ and $T(e_2) = (-1, 0, 4)^T$, then T is given by left multiplication by the matrix

$$\begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 4 \end{bmatrix}.$$

If $T : V \to W$ is a linear map on arbitrary vector spaces, then we can still describe T as multiplication by a matrix using the isomorphism $\psi : V \to F^n$. Recall that this isomorphism depends on choosing a basis for the space V, so that we expect the matrix to also depend on the choice of basis. The reader may wish to take a moment to review the notation established in Lecture 19.

Let V and W be finite dimensional vector spaces over F and let $\mathcal{B} = (v_1, \ldots, v_n)$ and $\mathcal{C} = (w_1, \ldots, w_m)$ be bases for V and W respectively. If $T : V \to W$ is a linear map, we define the hypervector $T(\mathcal{B})$ by

$$T(\mathcal{B}) = (T(v_1), \ldots, T(v_n)).$$

Since $\mathcal{C}$ is a basis for W, there is an $m \times n$ matrix A such that $T(\mathcal{B}) = \mathcal{C}A$. Recall that this is just notation for the existence of scalars $a_{ij} \in F$ such that

$$T(v_j) = w_1 a_{1j} + w_2 a_{2j} + \cdots + w_m a_{mj}.$$

Therefore A is the matrix whose j^{th} column is the coordinate vector of $T(v_j)$ with respect to the $\mathcal{C}$ basis. This matrix is called the **matrix of the linear transformation with respect to the bases $\mathcal{B}$ and $\mathcal{C}$**. Different choices of bases will, in general, lead to different matrices.

Our next goal is to show that left multiplication by the matrix of a linear map acts like the linear map provided we write all coordinates in the appropriate basis. To begin, let $v \in V$ and write v in terms of the $\mathcal{B}$ basis:

$$v = \mathcal{B}X = v_1 x_1 + \cdots + v_n x_n.$$

If we apply T to both sides of this equation, we have

$$T(v) = T(v_1)x_1 + \cdots + T(v_n)x_n = T(\mathcal{B})X = \mathcal{C}AX.$$

Therefore we see that $Y = AX$ is the coordinate vector of $T(v)$ with respect to the $\mathcal{C}$ basis. In summary then, given $T : V \to W$, $\mathcal{B}$ and $\mathcal{C}$, we can construct an $m \times n$ matrix A over F such that

$$T(\mathcal{B}) = \mathcal{C}A \quad \text{and} \quad AX = Y$$

where $v = \mathcal{B}X$ and $T(v) = \mathcal{C}Y$.

We conclude this lecture by determining what happens to the matrix A when we change basis. To this end, suppose that $\mathcal{B}' = (v_1', \ldots, v_n')$ and $\mathcal{C}' = (w_1', \ldots, w_m')$ are also bases for V and W respectively. Then there are invertible matrices $P \in \mathrm{GL}_n(F)$ and $Q \in \mathrm{GL}_m(F)$ that satisfy

$$PX = X' \qquad \text{and} \qquad QY = Y'$$

where X and X' denote the coordinate vectors of a vector $v \in V$ with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$ respectively and similarly for Y and Y'. Let A' denote the matrix of $T : V \to W$ with respect to the new bases $\mathcal{B}'$ and $\mathcal{C}'$ so that $A'X' = Y'$. It follows that

$$QAP^{-1}X' = QAX = QY = Y'$$

and hence $A' = QAP^{-1}$. As before, we can take P and Q as arbitrary invertible matrices (of the appropriate sizes). Specifically, we have the following theorem.

Theorem 4.2.2 *Let A be the matrix of a linear transformation $T : V \to W$ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. Then the matrices A' that represent T with respect to other bases are of the form*

$$A' = QAP^{-1}$$

where $Q \in \mathrm{GL}_m(F)$ and $P \in \mathrm{GL}_n(F)$ are arbitrary invertible matrices.

Proof. Our comments above show that every other matrix A' that represents T has this form. Conversely, given A, Q and P, the matrix QAP^{-1} represents T with respect to the bases $\mathcal{B}'$ and $\mathcal{C}'$ obtained from $\mathcal{B}$ and $\mathcal{C}$ by applying P and Q respectively. ■

4.3 Lecture 22: Eigenvectors

Our goal in the current lecture is to apply what we know about linear transformations and the matrices that represent them to the special case $V = W$. A linear map $T : V \to V$ from a vector space to itself is always called a **linear operator**. Moreover, since the domain and codomain coincide, we will always use a single basis $\mathcal{B}$ in V and hence we will refer to the matrix of the operator $T : V \to V$ with respect to $\mathcal{B}$. Since there is only one basis involved, there is only one invertible matrix $P \in \mathrm{GL}_n(F)$ in the change of basis formula. The details of the proof of the following theorem are left to the reader.

Theorem 4.3.1 *Let A be the matrix of a linear operator $T : V \to V$ with respect to the basis $\mathcal{B}$. Then the matrices A' that represent T with respect to other bases are of the form*

$$A' = PAP^{-1}$$

where $P \in \mathrm{GL}_n(F)$ is an arbitrary invertible matrix. ■

In general, we will say that two $n \times n$ matrices A and B are **similar** if $B = PAP^{-1}$ for some $P \in \mathrm{GL}_n(F)$. One of the goals of the current lecture is as follows: given a linear operator $T : V \to V$, is there a basis for with the matrix of the transformation is particularly simple? In the language of matrices, given a square matrix A over F, can we find an invertible matrix P such that PAP^{-1} is particularly simple? To explain what we mean by "particularly simple", we will need the following definition.

Definition 4.3.2 (Invariant subspace) *Let V be a vector space over F and $T : V \to V$ be a linear operator. A subspace $W \leq V$ is* **invariant under** *T if*

$$T(W) \subset W.$$

In other words, if W is invariant under T, then $T(w) \in W$ for all $w \in W$. Notice that in this case T induces a linear map $T_W : W \to W$ called the **restriction of T to W**. If W is invariant under T and $(w_1, \ldots, w_k)$ is a basis for W, then if we extend this to a basis $\mathcal{B}$ for V, the matrix of T with respect to $\mathcal{B}$ has the **block form**

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

where A is a $k \times k$ matrix. In fact, A is the matrix of the restriction of T to W with respect to the basis $(w_1, \ldots, w_k)$. Among the most important invariant subspaces are the 1-dimensional ones.

Definition 4.3.3 (Eigenvector) *If $T : V \to V$ is a linear operator on a vector space V, then a non-zero vector $v \in V$ is called an* **eigenvector** *for T if*

$$T(v) = \lambda v$$

for some scalar $\lambda \in F$. In this case the scalar λ is called an **eigenvalue associated to the eigenvector** *v.*

We remark here that the scalar $0 \in F$ may be an eigenvalue, but the zero vector $0 \in V$ is never an eigenvector. If v is an eigenvector for an operator $T : V \to V$, then the 1-dimensional subspace $W = Fv = \{\alpha v : \alpha \in F\}$ is invariant under T. Conversely, if Fv is T-invariant, then v is an eigenvector. Therefore the eigenvectors for T are precisely the vectors that form bases for the 1-dimensional T-invariant subspaces. Recall that invariant subspaces of dimension k give $k \times k$ blocks in the matrix of the operator T. Therefore eigenvectors will give 1×1 blocks in the matrix of T.

So far we have defined the notion of eigenvector and eigenvalue for linear operators $T : V \to V$. It is easy to guess how we might make corresponding definitions for $n \times n$ matrices over F since each such matrix gives rise to an operator $F^n \to F^n$. Specifically, we will call a non-zero column vector $X \in F^n$ an eigenvector for the matrix A if $AX = \lambda X$ for some $\lambda \in F$. As before λ is called an eigenvalue associated with X. A natural question at this point is: what is the relationship between the eigenvectors of an operator T and the eigenvectors of a matrix that represents T in some basis.

Proposition 4.3.4 *Let $T : V \to V$ be a linear operator and let A be the matrix of T with respect to a basis $\mathcal{B}$. Denote the coordinate of a vector $v \in V$ with respect to $\mathcal{B}$ by X. Then v is an eigenvector for T if and only if X is an eigenvector for A.*

Proof. We know that the coordinate vector of $T(v)$ is AX and of course the coordinate vector of λv is λX for all $\lambda \in F$. Therefore $T(v) = \lambda v$ iff. $AX = \lambda X$. ∎

Corollary 4.3.5 *Similar matrices have the same eigenvalues.*

Proof. If A is similar to B, then A and B represent the same linear operator T on F^n with respect to two different bases. If $\lambda \in F$ is an eigenvalue for A, then there is a non-zero vector X such that $AX = \lambda X$. The previous proposition implies that $T(v) = \lambda v$ where X is the coordinate vector for $v \in V$. If we let Y denote the coordinate vector of v with respect to the other basis, another application of the proposition shows that $BY = \lambda Y$ so that λ is an eigenvalue for B as well. Of course the above argument is reversible and the proof is complete. ∎

Definition 4.3.6 (Diagonalizable) *A $n \times n$ matrix A over F is* **diagonalizable** *if it is similar to a diagonal matrix. That is, A is diagonalizable iff. $PAP^{-1} = D$ for some diagonal matrix D.*

The following proposition is left as an exercise.

Proposition 4.3.7 *A $n \times n$ matrix A over F is diagonalizable if and only if F^n has a basis of eigenvectors for A.* ∎

4.4 Lecture 23: The characteristic polynomial

In this lecture, our goal is to be able to find all eigenvectors for a given linear operator $T : V \to V$. That is, given $T : V \to V$, we want to find all non-zero vectors $v \in V$ such that $T(v) = \lambda v$ for

some scalar $\lambda \in F$. Our method will be to first find all possible eigenvalues of T, for once we know that $\lambda \in F$ is a eigenvalue for T, then finding associated eigenvectors is equivalent to solving a linear system of equations $T(v) = \lambda v$. Recall from your homework problem 4.2.9 that the set of all linear operators on V form a vector space over F under pointwise addition and scalar multiplication. Therefore if we let $1_V : V \to V$ denote the identity transformation on V, for each $\lambda \in F$, the map $T - \lambda 1_V : V \to V$ is a linear operator on V. The reader should stop and write a careful proof of the following proposition.

Proposition 4.4.1 *If $T : V \to V$ is a linear operator on a F-vector space V, then $v \in V$ is an eigenvector for T if and only if v is a non-zero element of* $\ker(T - \lambda 1_V)$ *for some $\lambda \in F$.* ■

We have made a small step toward our goal. Namely, we have replaced the problem of finding the eigenvectors of an operator T with finding the non-zero elements of the kernel of another operator: $T - \lambda 1_V$. Our next step will be to replace the operators T and $T - \lambda 1_V$ with matrices that represent them in some basis. We will need the following lemma.

Lemma 4.4.2 *If $T : V \to V$ is a linear operator on a finite dimensional F-vector space V, then the following are equivalent:*

1. $\ker T \neq 0$.
2. $\operatorname{im} T \neq V$
3. *If A is the matrix of T with respect to any basis, then* $\det A = 0$.
4. 0 *is an eigenvalue for T.*

Proof. $(1 \implies 2)$ If $\ker T \neq 0$, then $\dim_F(\ker T) \neq 0$ so that the rank-nullity theorem implies that $\dim_F(\operatorname{im} T) \neq \dim_F V$. Therefore $\operatorname{im} T \neq V$.
$(2 \implies 3)$ If $\operatorname{im} T \neq V$, then T is not onto and hence T is not invertible. Therefore if A is any matrix that represents T, A is not invertible and hence $\det A = 0$.
$(3 \implies 4)$ Suppose that A is the matrix of T with respect to a basis $\mathcal{B}$ and that $\det A \neq 0$. Then the homogeneous system of equations $AX = 0$ has a non-trivial solution $X \in F^n$ and hence $v = \mathcal{B}X \in V$ is a non-zero element of $\ker T$. Therefore $T(v) = 0 \cdot v = 0$ and hence 0 is an eigenvalue for the eigenvector v of T.
$(4 \implies 1)$ If 0 is an eigenvalue for T, then by definition there is a non-zero vector $v \in V$ such that $T(v) = 0 \cdot v = 0$ so that $\ker T \neq 0$. ■

Definition 4.4.3 *A linear operator $T : V \to V$ is called* **singular** *if it satisfies one (and hence all) of the conditions in the previous lemma.*

In this terminology, a scalar $\lambda \in F$ is an eigenvalue for the operator T if and only if the operator $T - \lambda 1_V$ is singular. Furthermore, if we note that the matrix of the identity map $1_V : V \to V$ is the $n \times n$ identity matrix $I = I_n$ for any basis $\mathcal{B}$ of V, we have the following corollary.

Corollary 4.4.4 *If $T : V \to V$ is a linear operator on a finite dimensional F-vector space V, then $\lambda \in F$ is an eigenvalue for T if and only if $\det(A - \lambda I) = 0$ where A is the matrix of T in an arbitrary basis for V.*

Proof. It follows from homework problem 4.2.9 that the matrix of $T - \lambda 1_V$ is $A - \lambda I$. Therefore $\lambda \in F$ is an eigenvalue for T iff. $T - \lambda 1_V$ is singular iff. $\det(A - \lambda I) = 0$. ∎

We remark here that if A is an $n \times n$ matrix over F, then

$$\det(A - \lambda I) = 0 \Longleftrightarrow \det(\lambda I - A) = 0.$$

Definition 4.4.5 (Characteristic polynomial) *If $T : V \to V$ is a linear operator on a finite dimensional F-vector space V, the* **characteristic polynomial** *of T is*

$$p(\lambda) = \det(\lambda I - A)$$

where A is the matrix of T with respect to some basis.

We need to settle something right away. Namely, we must show that the characteristic polynomial $p(\lambda)$ does not depend on the choice of basis.

Proposition 4.4.6 *The characteristic polynomial of T does not depend on the choice of basis.*

Proof. Suppose that A and A' represent T with respect to the bases $\mathcal{B}$ and $\mathcal{B}'$ respectively. Then there exists $P \in \mathrm{GL}_n(F)$ such that $A' = PAP^{-1}$. We then compute

$$\det(\lambda I - A') = \det(\lambda I - PAP^{-1}) = \det(P(\lambda I - A)P^{-1}) = \det(\lambda I - A).$$

∎

If we roll all of this information up into one result, we have shown the following.

Corollary 4.4.7 *If $T : V \to V$ is a linear operator on a finite dimensional F-vector space V, then the eigenvalues of T are the roots of the characteristic polynomial $p(\lambda)$. In particular, the eigenvalues of an upper or lower triangular matrix A are its diagonal entries.*

Proof. The fist statement follows immediately from the above remarks. The reader should organize a careful proof. As for the second statement, we simply notice that $\lambda I - A$ is upper or lower triangular with diagonal entries $\lambda - a_{ii}$ so that

$$p(\lambda) = (\lambda - a_{11}) \cdots (\lambda - a_{nn}).$$

■

We end this lecture with a useful proposition.

Proposition 4.4.8 *Let $T : V \to V$ be a linear operator on a finite dimensional F-vector space V.*

1. *If $\dim_F V = n$, then T has at most n eigenvalues.*
2. *If $F = \mathbb{C}$ and $V \neq 0$, then T has at least one eigenvalue.*

Proof. Although we have not proved it in this course, a polynomial of degree n over *any* field can have at most n distinct roots. This shows (1). As for (2), a little fact called the *Fundamental Theorem of Algebra* states that every non-constant polynomial over $\mathbb{C}$ has a root. If $V \neq 0$, then $\dim_{\mathbb{C}} V \geq 1$ so that the degree of the characteristic polynomial has degree at least one and hence has a root in $\mathbb{C}$. ■

4.5 Lecture 24: Diagonalization

Our goal in this lecture is twofold: first we want to explain what we mean by particularly nice matrices and then we want to show that linear operators over complex vector spaces admit such matrix representations. For purposes of computation, we will say that a matrix is particularly nice if it is triangular. We begin with the following theorem.

Theorem 4.5.1 *Let $T : V \to V$ be a linear operator on a finite dimensional complex vector space V. Then there exist a basis $\mathcal{B}$ of V such that the matrix of T is upper-triangular with respect to $\mathcal{B}$.*

Proof. We proceed by induction on $n = \dim_{\mathbb{C}} V$. If $n = 1$, then every 1×1 matrix A is triangular so that we may take any basis (non-zero vector) for V. Suppose then that $n > 1$ and recall from the previous lecture that T has at least one eigenvalue $\lambda_1 \in \mathbb{C}$, and hence at least one eigenvector $v_1 \in V$. We can extend this eigenvector to a basis $\mathcal{B}' = (v_1, v_2', \ldots, v_n')$ for V. If A' is the matrix of

T with respect to $\mathcal{B}'$, then it follows that A' has the form

$$A' = \begin{bmatrix} \lambda_1 & * \\ 0 & B \end{bmatrix}$$

where B is an $(n-1) \times (n-1)$ matrix. If A is the matrix of T in an arbitrary basis for V, then we have shown that there is an element $P \in \mathrm{GL}_n(\mathbb{C})$ such that $A' = PAP^{-1}$. By induction, we may assume there exists an element $Q \in \mathrm{GL}_{n-1}(\mathbb{C})$ such that QBQ^{-1} is triangular. If Q_1 is the $n \times n$ matrix given in block form by

$$\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix},$$

then it follows that

$$(Q_1 P)A(Q_1 P)^{-1} = Q_1(PAP^{-1})Q_1^{-1} = Q_1 A' Q_1^{-1}$$

has the form

$$\begin{bmatrix} \lambda_1 & * \\ 0 & QBQ^{-1} \end{bmatrix},$$

which is triangular. ■

We remark here that the only property of the complex field $\mathbb{C}$ that is used in the above result is that every non-constant polynomial over $\mathbb{C}$ has a root in $\mathbb{C}$. If F is an arbitrary field with the property that every non-constant polynomial over F has a root in F, then F is called **algebraically closed**. Note that the fields $\mathbb{R}$ and $\mathbb{Q}$ are not algebraically closed. The reader is encouraged to modify the above proof to show the following theorem.

Theorem 4.5.2 *Let V be a finite dimensional vector space over an algebraically closed field F, and Let $T : V \to V$ be a linear operator on V. Then there exist a basis $\mathcal{B}$ of V such that the matrix of T is upper-triangular with respect to $\mathcal{B}$.* ■

Now, the best triangular matrices are the diagonal ones. There are many reasons for this. Perhaps one of the most important is that if D is a diagonal matrix with diagonal entries d_{ii}, then for any $k \in \mathbb{N}$, D^k is diagonal with diagonal entries d_{ii}^k. We therefore turn to the question: which linear operators over F-vector spaces admit a basis in which the matrix of the operator is diagonal? In some sense, we answered this question in Lecture 22. Namely we saw that an operator on V is

diagonalizable if and only if there is a basis of eigenvectors. Our goal here then is to discover a sufficient condition for the existence of such a basis. We begin with a lemma.

Lemma 4.5.3 *Let $v_1, \ldots, v_r$ be eigenvectors for a linear operator $T : V \to V$ with distinct eigenvalues $c_1, \ldots, c_r$. Then the (ordered) set $(v_1, \ldots, v_r)$ is linearly independent.*

Proof. We induct on r, the case $r = 1$ being trivial (a single non-zero vector is always an independent set). Suppose that

$$0 = a_1 v_1 + \cdots + a_r v_r$$

so that applying T gives

$$0 = a_1 c_1 v_1 + \cdots + a_r c_r v_r.$$

If we multiply the first relation by c_r and subtract the second, we have

$$0 = a_1(c_r - c_1)v_1 + \cdots + a_{r-1}(c_r - c_{r-1})v_{r-1}.$$

By induction, $a_j(c_r - c_j) = 0$ for all $j < r$. But $c_j \neq c_r$ if $j < r$ so that we must have $a_1 = \cdots = a_{r-1} = 0$. But then the original relation reduces to $a_r v_r = 0$ and hence $a_r = 0$ since $v_r \neq 0$. ∎

The next theorem is the sufficient condition we are after.

Theorem 4.5.4 *Let T be a linear operator on a vector space V of dimnesion n over a field F. If the characteristic polynomial $p(\lambda)$ for T has n distinct roots in F, then V has a basis $\mathcal{B}$ for which the matrix of T is diagonal.*

Proof. If $p(\lambda)$ has n distinct roots, then the above lemma immediately implies that the n corresponding eigenvectors are linearly independent and hence form a basis since $\dim_F V = n$. But we have already noted that the matrix of T with respect to such a basis is diagonal. ∎

We end this lecture with a remark about other nice matrices. If the characteristic polynomial $p(\lambda)$ has multiple roots, then T will not be diagonalizable in general. The study of this case leads one to something called the **Jordan canonical form of an operator**. We will return to this problem in MAT 150C as an application of the *Fundamental Theorem of finitely generated modules over a principal ideal domain.*

Index