Algebra I: Groups and Rings
Lectures Notes (Math 122)

Won I. Lee

These are a set of notes loosely based on the Math 122 course at Harvard, taught by Prof. Benedict Gross, and on M. Artin's text.[1] It also contains additional material based on chapters 7, 8, and 9 of Artin's text (on bilinear forms, linear groups, and group representations).

1 Group Theory

1.1 Linear Algebra Redux

$n \times n$ Matrices. Let us define the space of all $n \times n$ matrices over the reals as $M_n(\mathbb{R})$; this forms a real vector space of dimension n^2. We have the operations:

- Addition: $A + B = (a_{ij} + b_{ij})$
- Scalar multiplication: $\alpha A = (\alpha a_{ij})$
- Matrix multiplication: $A \cdot B = (c_{ij}) = (\sum_{k=1}^{n} a_{ik} b_{kj})$

We consider the last operation as *composition of linear transformations* after considering square matrices as representations of *linear operators* on a vector space, $T : \mathbb{R}^n \to \mathbb{R}^n$. **Note**: Matrix multiplication is not commutative for $n \geq 2$.

Other properties and laws of interest include:

- Identity: there exists $I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ such that $AI = IA = A$.
- Distributivity: $A(B + C) = AB + AC$
- Associativity: $A(BC) = (AB)C = ABC$
- Invertibility: A is *invertible* iff there exists $B \in M_n(\mathbb{R})$ such that $AB = BA = I$
- Determinant: There exists a unique mapping $\det : M_n(\mathbb{R}) \to \mathbb{R}$ such that it has the following properties: 1) $\det(I) = 1$; 2) det is linear in the rows of the matrix; 3) If two adjacent rows of A are equal, then $\det(A) = 0$.

The major theorem of invertibility is that:

$$A \text{ is invertible} \Leftrightarrow \det(A) \neq 0$$

General Linear Group, $GL_n(\mathbb{R})$. We consider all *invertible* matrices in $M_n(\mathbb{R})$ to form the *general linear group*, $GL_n(\mathbb{R})$, that is:

$$GL_n(\mathbb{R}) = \{A : \det(A) \neq 0\} = \{A : \exists A^{-1}, AA^{-1} = I\}$$

We note the following lemma:

Lemma 1.1 *If A^{-1} exists for matrix A, then it is unique.*

[1] M. Artin. *Algebra*. Prentice-Hall, 1991.

Proof Suppose there exists B, C such that $AB = BA = AC = CA = I$. Then we must have: $B = B(AC) = (BA)C = C$.

It is reasonable to ask what we gain and lose by imposing invertibility on the set of all $n \times n$ matrices:

- **Lose**: $GL_n(\mathbb{R})$ is no longer closed under addition or scalar multiplication!
  (i.e. $A + (-A) = 0 \cdot A = \mathbf{0}$)
- **Gain**: Closed under matrix multiplication.
  (follows from multiplicative property of determinants)

To close this description, we note the following properties of $GL_n(\mathbb{R})$ that we will precisely abstract to form the general definition of a group:

1. Product is associative: $(AB)C = A(BC)$
2. Identity exists: I
3. Every element A has an inverse A^{-1}

1.2 Groups + Subgroups

Definition (Group). A *group* G is a set with a product operation $\cdot : G \times G \to G$ such that:

1. Associative
2. Identity exists (e or $1 \in G$)
3. Inverses exist ($g^{-1} \in G$ for every $g \in G$)

The product is in general not commutative; if it is, we call G a *commutative* or *Abelian group*.

Examples (Groups). 1) The integers $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$ form an Abelian group under $+$.

2) $G = V$ for a vector space V forms an Abelian group under $+$ (ignoring scalar multiplication).

3) For any set T, we can construct the group:

$$G = \{\text{All bijections } g : T \to T\} = \text{Sym}(T)$$

is the *symmetric group of* T, where the product operation is the composition of mappings.

- e is the identity map
- g^{-1} is the inverse map, which exists since all g are bijections
- Composition of maps is associative by definition

If T is a *finite* set, $T \simeq \{1, 2, \dots, n\}$, then we say $\text{Sym}(T) = S_n$ is the *symmetric group on n letters*, and is a finite group of order $n!$ (i.e. number of elements). It is non-Abelian for $n \geq 3$.

Finally, note that:

$$GL_n(\mathbb{R}) \subset \text{Sym}(\mathbb{R}^n)$$

that is, the general linear group is a subset of the symmetric group on $\mathbb{R}^n$, since $GL_n(\mathbb{R})$ is the set of all bijections $\mathbb{R}^n \to \mathbb{R}^n$ that *preserve linear structure*. In fact, $GL_n(\mathbb{R})$ is a *subgroup*!

Definition (Subgroup). A *subgroup* $H \subset G$ is a subset such that H is:

1. Closed under $\cdot$
2. Contains e (identity)
3. Closed under inverses ($g \in H \Rightarrow g^{-1} \in H$)

Permutation Group. We now focus on the symmetric group on n letters, S_n, also called the *permutation group*. As we noted earlier, it is finite group of order $n!$, and the simplest examples are:

1. $S_1 = \{e\}$ is the simplest possible group, since any group G must contain the identity.

2. $S_2 = \{e, \tau\}$: there is only one possible choice for τ on 2 letters, $1, 2$, since it must be bijective and not equal to the identity; that is, $\tau = (12)$.

   Then, the multiplication table follows: $ee = e, e\tau = \tau e = \tau, \tau\tau = e$. The group is Abelian.

3. $S_3 = \{e, \tau, \tau', \tau'', \sigma, \sigma'\}$: we have $3! = 6$ elements, with τ being the *transpositions* (exchanging two letters, keeping one fixed), and σ being the *cyclic permutations* (keeping no element fixed). We note in particular that $\tau\sigma \neq \sigma\tau$, and therefore S_3 is not Abelian.

   Corollary 1.2 *S_n is non-Abelian for any $n \geq 3$.*

   Proof We can consider S_3 as a subgroup of S_n fixing all letters $\{4, 5, \dots, n\}$, and so τ, σ can be appropriately extended to S_n by doing the exact operations on $\{1, 2, 3\}$ and fixing the rest of the letters. Thus, $\tau\sigma \neq \sigma\tau$ holds in every S_n.

   As shown in the proof above, any S_k for $k \leq n$ is a subgroup of S_n.

Examples (Subgroups). 1) The general linear group: $GL_n(\mathbb{R}) = \{\text{Linear bijections of } \mathbb{R}^n\} \subset \text{Sym}(\mathbb{R}^n)$

- Composition of linear maps is linear, so closed under composition
- The identity map I is linear
- If a map is linear, so is the inverse

2) There is a further subgroup of $GL_2(\mathbb{R})$ on the plane that *stabilizes* the line $y = 0$:

$$H = \left\{ A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : ad \neq 0 \right\} \subset GL_2(\mathbb{R})$$

which follows since the first column represents Te_1, and so the mapping of basis vector e_1 under these matrices still lies on the line $y = 0$.

3) Subgroups of $(\mathbb{Z}, +)$ are particularly simple:

Proposition 1.3 *Subgroups of $(\mathbb{Z}, +)$ are exactly $(b\mathbb{Z}, +)$ for some fixed $b \in \mathbb{Z}$.*

Proof First, we show that all sets of the form $b\mathbb{Z}$ are subgroups. They are closed since $bm + bn = b(m + n) \in b\mathbb{Z}$; identity is simply $0 = b \cdot 0$; and inverses are $-bm = b(-m) \in b\mathbb{Z}$.

Next, we show that $b\mathbb{Z}$ exhaust the possible subgroups. Suppose $H \subset \mathbb{Z}$ is a subgroup. If $H = \{0\}$, then $H = 0\mathbb{Z}$ so we are done. If $H \neq \{0\}$, then there exists some $m \neq 0$, and so $-m \in H$ as well. One of them is positive, so we consider $m > 0$. Let $b > 0$ be the smallest positive integer in H; then $b\mathbb{Z} \subset H$, since H must be closed. Now suppose $h \in H$, then $h = mb + r$ for $0 \leq r < b$ by the Euclidean algorithm. However, if $r \neq 0$, then $r = h - mb \in H$ is positive yet smaller than b, which contradicts the fact that b is the smallest positive integer in H. Thus, $r = 0$ and $H = b\mathbb{Z}$.

4) For any element $g \in G$, we can generate the cyclic subgroup, or smallest subgroup, generated by that element:

Definition For any group G, the *cyclic subgroup generated by g* for some $g \in G$ is the smallest subgroup containing g, that is:

$$\langle g \rangle = \{e, g, g^{-1}, g^2, g^{-2}, \dots\} = \{g^m : m \in \mathbb{Z}\}$$

Note, however, that not all powers may be distinct. For example, in S_3, the element τ was such that $\tau^2 = e$, and so $\langle \tau \rangle = \{e, \tau\}$.

If m is the smallest power such that $g^m = e$, then we say that m is the *order of $g \in G$*. If $g^n \neq e$ for every n, then we say that g has infinite order.

An example of an infinite order element is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in GL_2(\mathbb{R})$:

$$\left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z}_+ \right\}$$

1.3 Homomorphisms + Isomorphisms

Motivating Example. Consider the following groups:

1. $G_1 = \{\pm 1, \pm i\} = \{i^k : k = 0, 1, 2, 3\} \subset \mathbb{C}^\times$
2. $G_2 = \{e, \rho, \rho^2, \rho^3\} \subset S_4$ where $\rho = (1234)$ is the cyclic permutation through the 4 letters.

We note that we can simply relabel $\rho \to i$ and the multiplication and operations would work out exactly; for example, $\rho^4 = i^4 = 1$, or the identity in the group. This motivates the idea of an *isomorphism*, which essentially tells us two groups are the "same" up to relabeling.

Definition (Isomorphism). An *isomorphism* is a mapping $f : G_1 \to G_2$ between groups such that:

1. f is bijective
2. $f(x \cdot y) = f(x) \cdot f(y)$

As an example, in G_1, G_2 in the motivating example above, $f(\rho^k) = i^k$ would be an isomorphism between the two groups.

Proposition 1.4 *Every cyclic group of order n is isomorphic to each other.*

Proof Let G_1, G_2 be two cyclic groups of order n, i.e. $G_1 = \langle x_1 \rangle$ and $G_2 = \langle x_2 \rangle$. Then let $f : G_1 \to G_2$ be defined by $f(x_1^k) = x_2^k$. It is clearly bijective, since there exists an inverse $f^{-1}(x_2^k) = x_1^k$, and preserves the multiplicative structure: $f(x_1^m x_1^n) = f(x_1^{m+n}) = x_2^{m+n} = x_2^m x_2^n = f(x_1^m) f(x_1^n)$.

We also note that there exists a cyclic group of order n for every $n \in \mathbb{Z}_+$, since we can simply take the cyclic subgroup generated by the cyclic permutation $\sigma_n = (123 \cdots n) \in S_n$.

Examples (Isomorphisms). 1) One particularly interesting example of isomorphic groups is:

$$(\mathbb{R}, +) \simeq (\mathbb{R}_+, \times)$$

where the mapping is given by $f : \mathbb{R} \to \mathbb{R}_+$ such that $f(x) = e^x$. Then:

$$f(x + y) = e^{x+y} = e^x e^y = f(x) f(y)$$

2) *Klein 4-group.* We have two representations of the Klein 4-group:

$$G_1 = \{e, \tau_1, \tau_2, \tau_1\tau_2\}$$

where $\tau_1 = (12)(34)$, $\tau_2 = (13)(24)$, and $\tau_1\tau_2 = (14)(23)$;

$$G_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

then the isomorphism maps:

$$f : e \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tau_1 \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \dots$$

Note: This group is not isomorphic to the earlier group of order 4, namely $G_3 = \{\pm 1, \pm i\}$. This is because G_3 has an element of order 4, i, whereas no element in G_1, G_2 has order 4.

Testing for Isomorphisms. To quickly rule out cases in which groups G_1, G_2 are not isomorphic, we can check that:

1. $|G_1| = |G_2|$
2. G_1 Abelian $\Leftrightarrow$ G_2 Abelian
3. G_1, G_2 have same number of elements of every order

If any of these properties do not hold, G_1, G_2 are not isomorphic.

Definition (Automorphism). An *automorphism* is an isomorphism from a group to itself.

Note that this is more restrictive than $\mathrm{Sym}(G)$, since if, say, $f(x) = x', f(y) = y'$ and $x \cdot y = z$, then we must have $f(z) = z'$ where $x' \cdot y' = z'$ in G. Thus, the possible permutations are restricted by the multiplication table.

The primary need for this definition is to construct the *automorphism group of* G, namely:

$$\mathrm{Aut}(G) = \{\text{Isomorphisms } G \to G\} = \{\text{“Symmetries” of } G\}$$

which can be verified as a group (mainly inverses exist).

Definition (Homomorphism). A *homomorphism* is a map $f : G_1 \to G_2$ such that $f(x \cdot y) = f(x) \cdot f(y)$.

This is a natural generalization of the idea of an isomorphism, removing the necessity of being a bijection. There exist many groups, some in the examples to follow, that are not isomorphic yet share a multiplicative structure in some way (the determinant mapping from $GL_n(\mathbb{R})$ to $\mathbb{R}^\times$). It turns out that the idea of a homomorphism is much more fundamental to further group and ring theory than is isomorphism.

Examples (Homomorphisms). 1) $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times = (\mathbb{R} \setminus \{0\}, \times)$

We have $\det(AB) = \det(A)\det(B)$, but the groups are not isomorphic since $\mathbb{R}^\times$ is Abelian while $GL_n(\mathbb{R})$ is not. Moreover, two different matrices can in fact have the same determinant.

2) $f : S_3 \to S_n$ which permutes the letters $\{1, 2, 3\}$ and keeps $\{4, 5, \ldots, n\}$ fixed.

3) $f : \mathbb{Z} \to S_2$ or the *sign mapping*, where:

$$f(\text{even}) = e, f(\text{odd}) = \tau$$

Note that $e \cdot e = e$, just as even $\times$ even = even, and similarly $\tau \cdot \tau = e$, just as odd $\times$ odd = even.

Properties of Homomorphisms. If f, h are homomorphisms, the following properties hold:

- $f(e) = e'$
- $f(g^{-1}) = f(g)^{-1}$
- $f \circ h$ is also a homomorphism

1.4 Kernels, Images, and Normal Subgroups

Definition (Kernel + Image). Let f be a homomorphism from G to G'. Then the *image of* f is:

$$\mathrm{im}(f) = \{g' : f(g) = g' \text{ for some } g \in G\} \subset G'$$

and the *kernel of* f is:

$$\ker(f) = \{g : f(g) = e'\} \subset G$$

One observation we make is that the kernel happens to be *closed under conjugation.* That is, if $h \in \ker(f)$, then $ghg^{-1} \in \ker(f)$, since $f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)e'f(g)^{-1} = e'$. This motivates the idea of a *normal subgroup.*

Definition (Normal Subgroup). For a group G, a subgroup H is *normal* in G, denoted $H \lhd G$ if:

$$\forall g \in G, gHg^{-1} = H$$

Examples (Normal Subgroups). 1) As noted above, the kernel of any homomorphism is normal.

2) All Abelian groups are normal (since $ghg^{-1} = gg^{-1}h = h \in H$)

3) As a non-example, $G = S_3$ has a non-normal subgroup. This is because we can consider $\tau = (12)$ and $\tau' = (23)$; then conjugating τ by τ' yields τ''. However, this implies that $H = \{e, \tau\}$ is not normal.

Examples (Kernels + Images). 1) $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$. We know that det is a homomorphism. Then:

$$\text{im}(\det) = \mathbb{R}^\times$$

$$\ker(\det) = \{A : \det(A) = 1\} = SL_n(\mathbb{R})$$

where $SL_n(\mathbb{R})$ is the *special linear group*, and arises as the kernel of the determinant homomorphism. This subgroup is normal, since:

$$\det(BAB^{-1}) = \det(B)\det(A)\det(B^{-1}) = \det(B)\det(B)^{-1} = 1$$

so that we have $SL_n(\mathbb{R}) \lhd GL_n(\mathbb{R})$.

2) $f : S_n \to GL_n(\mathbb{R})$, where the mapping associates a permutation with its matrix, $f(\sigma) = A_\sigma$. For example, in S_3 with the cyclic permutation $\sigma = (123)$, we have:

$$f(\sigma) = A_\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Note that composition of permutations is exactly mapped to matrix multiplication of permutation matrices, which is almost tautological when matrices are considered as representations of linear maps. Thus, f is a homomorphism, and we can consider the image and kernel:

$$\text{im}(f) = \{\text{Permutation matrices in } GL_n(\mathbb{R})\}$$

$$\ker(f) = \{e\}$$

since the kernel of f consists of permutations mapping to $I \in GL_n(\mathbb{R})$.

3) *Sign homomorphism.* Composition of (1) and (2), i.e. $\text{sign} = \det \circ f : S_n \to \mathbb{R}^\times$. Recall that composition of homomorphisms is itself a homomorphism. Moreover, we have: $\det(f(\sigma)) = \pm 1$ for every $\sigma \in S_n$, since $f(\sigma)$ is a permutation matrix. Thus:

$$\text{im}(\text{sign}) = \{\pm 1\} \in \mathbb{R}^\times$$

$$\ker(\text{sign}) = \{\sigma : \text{sign}(\sigma) = \det(f(\sigma)) = 1\} = \{\text{Even permutations}\} = A_n$$

where A_n is the *alternating group on n letters*, and arises as the kernel of the sign homomorphism. Thus, $A_n \lhd S_n$ is a normal subgroup, and $|A_n| = \frac{n!}{2}$.

Definition (Center). For a group G, the *center of* G is defined as the commutative elements:

$$Z(G) = \{z \in G : \forall g \in G, zg = gz\}$$

Examples (Centers). 1) $G = Z(G) \Leftrightarrow G$ is Abelian.

2) $G = S_n \Rightarrow Z(G) = \{e\}$ for $n \geq 3$.

3) $G = GL_n(\mathbb{R}) \Rightarrow Z(G) = \{\lambda I : \lambda \in \mathbb{R}^\times\}$

Conjugation + Inner Automorphisms. Consider attempting to construct a homomorphism from G to the automorphism group of G; we want a homomorphism $f : G \to \text{Aut}(G)$. The first suggestion to come to mind may be *left translation by* g, that is, $f(g)(h) = g \cdot h$. However, this element $f(g)$ is not in fact a homomorphism:

$$f(g)(hh') = ghh' \neq (gh)(gh') = f(g)(h) \cdot f(g)(h')$$

The solution is to consider instead *conjugation by* g, namely:

$$f(g)(h) = ghg^{-1} \in \text{Aut}(G)$$

which is in fact a homomorphism because $f(g)(hh') = ghh'g^{-1} = ghg^{-1} \cdot gh'g^{-1} = f(g)(h) \cdot f(g)(h')$. The mapping f is bijective since we can exhibit an inverse, namely $f(g)^{-1} = f(g^{-1})$. Finally, f is itself a homomorphism because:

$$f(gg')(h) = (gg')h(gg')^{-1} = g(g'h(g')^{-1})g^{-1} = g(f(g')(h))g^{-1} = (f(g) \circ f(g'))(h)$$

In this case, the kernel turns out to be exactly the center of G:

$$\ker(f) = Z(G)$$

since $f(g) = e \Leftrightarrow ghg^{-1} = h \Leftrightarrow gh = hg \Leftrightarrow g \in Z(G)$.

Now it makes sense to ask what the *image of* f is; that is, which automorphisms of G are hit by the conjugation operation?

Definition If $f : G \to \mathrm{Aut}(G)$ is the conjugation map, then the *inner automorphisms of* G are:

$$\mathrm{Inn}(G) = \mathrm{im}(f) = \{a(h) = ghg^{-1} \text{ for some } g \in G\} \subset \mathrm{Aut}(G)$$

1.5 Equivalent Relations + Cosets

Definition (Equivalence Relation). For set S, an *equivalent relation* on S is a relation $\sim$ that is:

1. Reflexive: $a \sim a$
2. Symmetric: $a \sim b \Leftrightarrow b \sim a$
3. Transitive: $a \sim b, b \sim c \Rightarrow a \sim c$

More precisely, $\sim$ defines a subset of the Cartesian product $S \times S$, namely: $\{(a, b) : a \sim b\}$.

Proposition 1.5 *Every equivalence relation* $\sim$ *defines a* partition *of the set* S*, and vice versa.*

Proof Given an equivalence relation $\sim$, consider any element $a \in S$, and define the subset $C_a = \{b : a \sim b\} \subset S$. If $C_a = S$, we are done. Otherwise, select an element in $S \setminus C_a$, and repeat until $\cup_a C_a = S$. This is a partition since every element $a \in S$ is included in some C_a, and if $c \in C_a \cap C_b$, then $a \sim c, c \sim b \Rightarrow a \sim b$ and $C_a = C_b$.

Given a partition $C_1, \ldots, C_k$ of S, let $a \sim b \Leftrightarrow a, b \in C_i$ for some i. This defines an equivalence relation since clearly $a \sim a$; $a \sim b \Rightarrow a, b \in C_i \Rightarrow b \sim a$; and $a \sim b, b \sim c \Rightarrow a, b, c \in C_i \Rightarrow a \sim c$.

Definition The *equivalence classes* of S for an equivalence relation $\sim$ are the subsets forming the partition defined by $\sim$ and are denoted $\bar{S}$.

The *canonical map* $S \to \bar{S}$ is given by $a \mapsto \bar{a} = \{x \in S : a \sim x\}$. The map is surjective but not injective.

Proposition 1.6 *The map* $f : S \to T$ *generates an equivalence relation and partition on* S*, defined by* $a \sim b \Leftrightarrow f(a) = f(b) \in T$.

Note that we can then consider $\bar{S} = \mathrm{im}(f) \subset T$, since different $t \in \mathrm{im}(f)$ index the equivalence classes of S. Namely, $\bar{S}$ is exactly the *fibers of* f:

$$\bar{S} = \{f^{-1}(t) : t \in T\}$$

Example (Fibers, $\bar{S}$). Consider the group $S = \mathbb{R}^+$, and define the homomorphism $f(x) = e^{2\pi i x}$, where T is then the unit circle in $\mathbb{C}$. Then we have:

$$f^{-1}(1) = \mathbb{Z} \subset \mathbb{R}$$

Definition (Coset). Let G be a group and $H \subset G$ be a subgroup. Then the *left coset of* H is given for $a \in G$:

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

Now suppose that $f : G \to G'$ is a homomorphism, and $H = \ker(f)$ be the kernel. Then $H \lhd G$, and we can construct an equivalence relation on G using $a \sim b \Leftrightarrow f(a) = f(b)$. In this case, H is one of the equivalence classes, namely:

$$H = \ker(f) = f^{-1}(e')$$

Proposition 1.7 *All equivalence classes for the equivalence relation generated by a homomorphism f have the form aH, where $H = \ker(f)$.*

Proof Consider the equivalence class of an arbitrary element a, namely C_a. Then if $b \in C_a$, $f(a) = f(b) \Rightarrow f(a)^{-1}f(b) = e' \Rightarrow f(a^{-1}b) = e'$. Thus, $a^{-1}b \in H$, i.e. $a^{-1}b = h$ for some $h \in H$. But then $b = aa^{-1}b = ah \in aH$, so $C_a \subset aH$. Now suppose $b \in aH$. Then $b = ah \Rightarrow f(b) = f(ah) = f(a)f(h) = f(a)$, so that $b \in C_a$. Thus, $aH \subset C_a$, so $C_a = aH$.

Proposition 1.8 *For any subgroup $H \subset G$, the mapping $h \mapsto ah$ yields a bijection of sets $H \simeq aH$.*

Proof This map is injective since if $ah = ah' \Rightarrow h = h'$. It is clearly onto, since every element $z \in aH$ is of the form ah for some $h \in H$.

Corollary 1.9 *If $|H|$ is finite, then $|H| = |aH|$ for every $a \in G$.*

Theorem 1.10 *If $f : G \to G'$ is a homomorphism of groups G, G', then $H = \ker(f) \lhd G$ and the left cosets aH partition the group G with the same order.*

Corollary 1.11 *If G is a finite group and $f : G \to G'$ is a homomorphism with $h = \ker(f)$, then:*

$$|G| = |H| \cdot |im(f)|$$

Proof As noted earlier, the set $\text{im}(f)$ indexes the equivalence classes (or fibers) generated by the homomorphism f; that is, $\bar{S} \simeq \text{im}(f)$ so that $|\bar{S}| = |\text{im}(f)|$. Moreover, $|G| = \sum_{\bar{a} \in \bar{S}} |\bar{a}|$. We know from Proposition 1.7 and Corollary 1.9 that every equivalence class $\bar{a} = aH$ and $|aH| = |H|$, so we have: $|G| = \sum_{\bar{a} \in \bar{S}} |\bar{a}| = |\bar{S}| \cdot |H| = |H| \cdot |\text{im}(f)|$.

In particular, we define the *index of* H as the number of distinct cosets/equivalence classes, denoted by: $[G : H]$. Thus, the formula becomes:

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

More generally, we have:

Proposition 1.12 *Let $H \subset G$ be any subgroup of the group G. Then the left cosets aH partition G, and are in bijection with H.*

Proof Let $H \subset G$ be a subgroup. Note that since $e \in H$ for every subgroup, $a = a \cdot e \in aH$, so that every $a \in G$ is in some coset aH. Moreover, consider any two cosets aH, bH such that $aH \neq bH$. Now suppose there exists some $c \in aH \cap bH$; then $c = ah = bh'$ for some $h, h' \in H$. But this implies that $a^{-1}b = h(h')^{-1} \in H$, so that $a^{-1}b = h'' \in H \Rightarrow b = ah'' \in aH \Rightarrow bH \subset aH$. But we can flip this exactly to get $b^{-1}a \in H$, so $aH \subset bH$ and $aH = bH$. This is a contradiction, so the cosets are disjoint if they are not equal, and thus form a partition of G. Moreover, by the same argument as earlier, every coset aH is in bijection with H, by the mapping $h \mapsto ah$, which is clearly onto and injective since $ah = ah' \Rightarrow h = h'$.

In particular, the formula $|G| = |H| \cdot [G : H]$ holds for every subgroup H in G.

Lagrange's Theorem. If $|G|$ is finite and $g \in G$, then the order of g divides $|G|$.

Proof Let $H = \langle g \rangle$ be the cyclic subgroup generated by $g \in G$. Then $H = \{e, g, \ldots, g^{m-1}\}$ for some m, since G is a finite group. Thus, the order of g is $m = |H|$. Note that since $|G| = |H| \cdot [G : H]$, in particular $|H|$ divides $|G|$.

Proposition 1.13 *Let G be a finite group with $|G| = p$ for prime p. Then G is cyclic and is generated by any $g \neq e$; namely $G = \langle g \rangle$. Moreover, the only subgroups of G are $\{e\}$ and G itself.*

Proof Let $g \in G$ such that $g \neq e$. Then the order of g divides p but is not 1, since $g \neq e$. Thus, the order of g is p, namely $|\langle g \rangle| = p$. Since G itself has order p, we must have $\langle g \rangle = G$.

Definition (Simple Group). A group G is *simple* if the only normal subgroups of G are $\{e\}$ and G.

Examples (Simple Groups). 1) Any groups of prime order p.

2) A_n is simple for every $n \geq 5$.

3) Any finite, non-Abelian simple group has even order.

1.6 Congruence mod n

We take a brief excursion to discuss congruence modulo n for positive integer n, which generalizes the concept of evens/odds. In addition, it provides concrete examples of cosets and equivalence relations, as well as the first major example of a quotient group.

Definition (Congruence mod n). For fixed $n \in \mathbb{N}$, *congruence mod n* is the equivalence relation with:

$$a \sim b \Leftrightarrow n|(a-b) \Leftrightarrow a - b \in n\mathbb{Z}$$

We can show this is an equivalence relation by noting that clearly $a \sim a$ since $n|0$; $a \sim b \Rightarrow n|(a-b) \Rightarrow n|(b-a) \Rightarrow b \sim a$; and $a \sim b, b \sim c \Rightarrow n|(a-b), n|(b-c) \Rightarrow n|[(a-b)+(b-c)] \Rightarrow n|(a-c) \Rightarrow a \sim c$.

For notation, we will say that $a \sim b \bmod n \Leftrightarrow a \equiv b \mod n$, and state this as a is congruent to b mod n.

Equivalence Classes mod n. Since $a \equiv b \mod n$ is an equivalence relation on $\mathbb{Z}$, there exists a set of equivalence classes that partition $\mathbb{Z}$. We see that:

$$\bar{a} = \{b : a \equiv b \mod n\} = a + n\mathbb{Z} = \{a + nk : k \in \mathbb{Z}\}$$

Thus, the equivalence classes are precisely the *left cosets of $n\mathbb{Z}$*. We denote the set of equivalence classes as:

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

where we know that this set of n elements exhausts the equivalance classes, since if $a \in \mathbb{Z}$, then $a = nq + r$ for $r \in 0, 1, \ldots, n-1$, and so $\overline{a} = \overline{r}$.

Canonical Homomorphism. There exists a canonical homomorphism mapping $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, such that $a \mapsto \overline{a}$. This follows from the fact that:

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

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

Group of Equivalence Classes. In fact, the equivalence classes $\mathbb{Z}/n\mathbb{Z}$ form a *group* under $+$; that is, $(\mathbb{Z}/n\mathbb{Z}, +)$ is an Abelian group.

1. Associative: Since $(\mathbb{Z}, +)$ was associative, so is $\mathbb{Z}/n\mathbb{Z}$.
2. Identity: $\overline{0}$ is the additive identity on $\mathbb{Z}/n\mathbb{Z}$.

3. Inverses: $-\overline{a} \equiv (\overline{a})^{-1} = \overline{n-a} \equiv \overline{-a}$ is the inverse of $\overline{a}$.

Moreover, $\mathbb{Z}/n\mathbb{Z}$ is a *cyclic group of order n*!, generated by $\overline{1}$, that is:

$$\mathbb{Z}/n\mathbb{Z} = \langle \overline{1} \rangle$$

since we can define $\overline{k} = \sum_{i=1}^{k} \overline{1}$.

Group of Units. Note that $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ is not itself a group, since $\overline{0}$ has no inverse (that is, no element $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$ that yields $\overline{k} \cdot \overline{0} = \overline{1}$). However, we can construct a subset of $\mathbb{Z}/n\mathbb{Z}$ that is in fact a group under $\cdot$, called the *group of units of* $\mathbb{Z}/n\mathbb{Z}$ and denoted $(\mathbb{Z}/n\mathbb{Z})^\times$:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{\overline{a} \in \mathbb{Z}/n\mathbb{Z} : \exists \overline{c} \in \mathbb{Z}/n\mathbb{Z}, \overline{a} \cdot \overline{c} = \overline{1}\}$$

namely, all the elements of $\mathbb{Z}/n\mathbb{Z}$ that have inverses. We want a more explicit formulation of the precise elements of $\mathbb{Z}/n\mathbb{Z}$ that make up the group of units, and we start with the following lemma:

Lemma 1.14 *For $m, n \in \mathbb{Z}$, $m\mathbb{Z} + n\mathbb{Z} = \gcd(m, n)\mathbb{Z}$.*

Proof We can easily check that $m\mathbb{Z} + n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$ by checking the axioms (0 is in the subgroup; inverses exist; and closure easily follows). Thus, since every subgroup of $\mathbb{Z}$ is of the form $k\mathbb{Z}$, we must have $m\mathbb{Z} + n\mathbb{Z} = k\mathbb{Z}$ for some $k \in \mathbb{Z}$. In particular, $m \in m\mathbb{Z} + n\mathbb{Z}$, so $m = kz \Rightarrow k|m$. Similarly, $n \in m\mathbb{Z} + n\mathbb{Z}$ so $k|n$. Thus, $k \leq \gcd(m, n)$. Now suppose $l|m, l|n$. Then since $k \in k\mathbb{Z} \Rightarrow k = mr + ns \Rightarrow l|k$. Thus, $k = \gcd(m, n)$.

Proposition 1.15 $(\mathbb{Z}/n\mathbb{Z})^\times = \{\overline{a} \in \mathbb{Z}/n\mathbb{Z} : \gcd(a, n) = 1\}$

That is, $(\mathbb{Z}/n\mathbb{Z})^\times$ consists precisely of the elements relatively prime to n.

Proof Suppose $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ and $\gcd(a, n) = 1$. Then $a\mathbb{Z} + n\mathbb{Z} = \gcd(a, n)\mathbb{Z} = \mathbb{Z}$ by Lemma 1.14. Thus, there exist $r, s \in \mathbb{Z}$ such that $ar + ns = 1 \Rightarrow ar - 1 \in n\mathbb{Z} \Rightarrow \overline{ar} = \overline{1} \Rightarrow \mathbf{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$.

Now suppose $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$. Then there exists some $\overline{c}$ such that $\overline{ac} = \overline{1}$. Thus, $ac - 1 \in n\mathbb{Z} \Rightarrow ac - 1 = nb \Rightarrow 1 = ac + nb' \in a\mathbb{Z} + n\mathbb{Z} = \gcd(a, n)\mathbb{Z}$. But 1 is only in the cosets $k\mathbb{Z}$ for $k = \pm 1$, so $\gcd(a, n) = 1$.

1.7 Quotient Groups

Given our discussion regarding $\mathbb{Z}/n\mathbb{Z}$, one natural question may be to ask whether this idea generalizes. That is, $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$, and we formed a new group of equivalence classes, $\mathbb{Z}/n\mathbb{Z}$, by "quotienting out" the subgroup $n\mathbb{Z}$, which turned out to be the set of cosets of $n\mathbb{Z}$. We then established a group structure on $\mathbb{Z}/n\mathbb{Z}$ using either $+$ or $\cdot$, the latter of which required additional restrictions on the allowable elements.

We know that if $H \subset G$ is a subgroup, then the left cosets aH always *partition* the group G, as proved earlier. However, when is it possible to construct a group structure on the set of cosets?

First, let us use the notation G/H (later reserved for quotient groups) to denote the set of cosets of H in G. We attempt to discover when we can transport the group structure of G onto G/H and still retain a group.

Examples. 1) Suppose $f : G \to G'$ is a homomorphism and $H = \ker(f) \lhd G$. Then the set of cosets $G/H = \{aH\} \simeq$ fibers of map $f \simeq \text{im}(f) \subset G'$. However, $\text{im}(f)$ is a subgroup of G', so that we can import the structure of G' onto G/H and yield a group. Namely:

$$aH \cdot bH = abH$$

This works since:

$$\overline{a} \cdot \overline{b} = f(a) \cdot f(b) = f(ab) = \overline{a \cdot b}$$

and every coset $aH \leftrightarrow \overline{a}$ bijectively. This makes:

$$F : G \to G/H$$

mapping $a \mapsto aH$ a surjective group homomorphism.

2) More generally, suppose that we let $H \subset G$ be any subgroup, and G/H be the left cosets of H. If we again attempt to set $aH \cdot bH = abH$, we find that this is not well-defined!

This is because the representatives a, b are not unique; namely, there are other elements in the cosets such that $aH = a'H, bH = b'H$. Then, we must have $abH = a'b'H$ in order for this multiplication to be well-defined on G/H. However, this is not true in general. Suppose that H is not normal, so that $aHa^{-1} \neq H$. Then:

$$(aH)(a^{-1}H) = (aa^{-1})H = eH = H$$

under the above definition of multiplication on cosets. However, since H is not normal, there exists $h \in H$ such that $aha^{-1} \notin H$. Thus, since we must have $ah \in aH$, $a^{-1} \in a^{-1}H$, we then find that:

$$(ah)a^{-1} = aha^{-1} \notin H$$

and so we cannot possibly have $(aH)(a^{-1}H) = H$!

3) From example (2), we see that a crucial condition is for the subgroup H to be *normal*. If $H \lhd G$, we can prove that the multiplication law $aH \cdot bH = abH$ actually yields the correct result.

First, note that $H \lhd G$ implies that $aHa^{-1} = H$, so that $aH = Ha$. In other words, $ah = h'a$ for some $h' \in H$. To show that the above multiplication defines a group structure on G/H, we first note that $eH = H$ is the identity element, and $(aH)^{-1} = a^{-1}H$ is the inverse. Now we must check that the operation is well-defined under change of representatives.

We want to show that $(ah) \cdot (bh') = (ab)h''$ for some $h'' \in H$. Now by normality, $(ah)(bh') = (ha)(bh') = h(ab)h' = (ab)hh' = (ab)h'' \in abH$. Thus, the multiplication is well-defined!

Theorem 1.16 *Let G be a group and $H \lhd G$ be a normal subgroup. Then the* quotient group G/H *is defined by:*

$$G/H = \{aH : a \in G\}$$

with the group operation:

$$aH \cdot bH = (ab)H$$

and yields the canonical homomorphism*:*

$$f : G \to G/H$$

defined by: $f(a) = aH$ with $\ker(f) = H$.

In particular, note that we have proven that *every normal subgroup H is the kernel of some homomorphism*! (Namely, the canonical homomorphism mapping G to G/H.)

First Isomorphism Theorem. If $f : G \to G'$ is a surjective group homomorphism with $\ker(f) = H$, then f induces an *isomorphism* of groups:

$$\overline{f} : G/H \to G'$$

such that $\overline{f}(aH) = f(a)$.

Proof Note that if $aH = a'H$, then $\overline{f}(aH) = f(a) = f(a') = \overline{f}(a'H)$ since $a' \in aH$, so that $a' = ah$ and $f(a') = f(a)f(h) = f(a)$. Thus, $\overline{f}$ is well-defined. It is also a homomorphism since $\overline{f}(aH \cdot bH) = \overline{f}(abH) = f(ab) = f(a)f(b) = \overline{f}(aH)\overline{f}(bH)$. Finally, the map is surjective since f is surjective, and it is injective since $\overline{f}(aH) = e' \Leftrightarrow aH = eH = H$ so that $\ker(\overline{f}) = \{H\}$ is trivial.

Subgroups in G and G/H. Suppose that $H \lhd G$, so that G/H is a quotient group. Now suppose there exists some subgroup K such that:

$$H \lhd K \subset G$$

Note that H is normal in K, since $gHg^{-1} = H$ for every $g \in G$ implies $kHk^{-1} = H$ for every $k \in K \subset G$. Thus, K/H is also a quotient group, and in fact is a subgroup of G/H. The converse, that a subgroup of G/H corresponds to a subgroup of G containing H, also holds.

Proposition 1.17 *Let G be a group and $H \lhd G$. Then there exists a 1-1 correspondence:*

$$\{\textit{Subgroups } K \subset G \textit{ containing } H\} \leftrightarrow \{\textit{Subgroups of } G/H\}$$

Proof Suppose $H \lhd K \subset G$, then $K/H \subset G/H$ is a subgroup of G/H. This is because K is a subgroup, and therefore if $aH, bH \in K/H$, then $ab \in K$ and so $abH \in K/H$. Conversely, if $\{aH\}$ is a subgroup of G/H, then $K = \cup(aH)$ is a subgroup of G containing H, since if $a, b \in K$, then aH, bH are cosets in the subgroup, so abH is in the subgroup and $ab \in K$. Moreover, the subgroup of cosets must include H, so $H \in \cup(aH) = K$.

Corollary 1.18 *Let $\mathbb{Z}$ be the group of integers under $+$ and p be a prime integer. If $p\mathbb{Z} \subset K \subset \mathbb{Z}$ for subgroup K, then either:*

1. $K = \mathbb{Z}$
2. $K = p\mathbb{Z}$

That is, pZ is the maximal subgroup *of $\mathbb{Z}$.*

Proof Let $G = \mathbb{Z}$ and $H = p\mathbb{Z}$. Then the quotient group $G/H = \mathbb{Z}/p\mathbb{Z}$ is a cyclic group, and $K/p\mathbb{Z}$ is a subgroup of this cyclic group. Since $|K/p\mathbb{Z}|$ divides $|\mathbb{Z}/p\mathbb{Z}| = p$, either $|K/p\mathbb{Z}| = 1$ or p. If $|K/p\mathbb{Z}| = 1$, then $K = p\mathbb{Z}$; if $|K/p\mathbb{Z}| = p$, then $K = \mathbb{Z}$. This is because if $K/p\mathbb{Z} = \mathbb{Z}/p\mathbb{Z}$, then $K = \cup_{C \in K/p\mathbb{Z}} C = \cup_{C \in \mathbb{Z}/p\mathbb{Z}} C = \mathbb{Z}$, and similar for $K = p\mathbb{Z}$.

2 Vector Spaces

2.1 Abstract Vector Spaces

We are familiar with the elementary vector spaces over the reals or complex numbers, for which the canonical example is $\mathbb{R}^n$. These allow us to abstract the desired properties of vector spaces in axiomatic form to construct vector spaces over any given field F. We always fix some field F for the remainder.

Definition (Field). A *field* F is a set with the following properties:

1. Abelian group under $+$, with identity 0 and inverse $-a$
2. $F \setminus \{0\} = F^\times$ is Abelian group under $\times$, with identity 1 and inverse $a^{-1} = 1/a$

Note that the rational numbers $\mathbb{Q} = \{a/b : a, b \in \mathbb{Z}, b \neq 0\}$ forms a field, whereas $\mathbb{Z}$ does not.

Definition (Subfield). A *subfield* $F' \subset F$ is a subset of F that is closed under $+$, $\times$, and inverses.

The canonical example is that $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ are subfields of the complex numbers.

Given this observations, it is natural to ask whether there exist fields that are not subfields of the complex numbers. One is given by the *trivial field*, i.e. $\{0, 1\}$. Note that this is the simplest possible field, since all fields must contain at least two elements, $0, 1$.

Proposition 2.1 *If p is a prime number, then $\mathbb{Z}/p\mathbb{Z}$ with $+, \times$ forms a field.*

Proof We know that $\mathbb{Z}/p\mathbb{Z}$ is an Abelian group under $+$. Thus, we must show that $(\mathbb{Z}/p\mathbb{Z})^\times$ is an Abelian group under $\times$, namely that it is closed under inverses. We recall that $p\mathbb{Z} \subset \mathbb{Z}$ is the maxmial subgroup. Thus, if $a \not\equiv 0 \mod p$, then $a \notin p\mathbb{Z}$, and so $p\mathbb{Z} + a\mathbb{Z} = \gcd(a, p)\mathbb{Z} = Z$. Thus, $1 \in Z = p\mathbb{Z} + a\mathbb{Z} \Rightarrow 1 = pn + ab \Rightarrow 1 \equiv ab \mod p$. Thus, b is the inverse of a in $\mathbb{Z}/p\mathbb{Z}$.

Lastly, we note that $\mathbb{Z}/p\mathbb{Z}$ is not a subfield of $\mathbb{C}$, since $\mathbb{Z}/p\mathbb{Z}$ is a cyclic group generated by 1, whereas $\mathbb{C}$ does not have any elements of order p.

Definition (Vector Space). A *vector space* V over the field F is the set such that:

1. V is an Abelian group under $+$
2. There exists a scalar product operation $V \times F \to V$ that is distributive, associative, has identity 1, $0 \cdot v = 0_V$, and so forth

Examples (Vector Spaces). 1) $\{0_V\}$ is the simplest possible vector space.

2) $V = F$, or the field itself, is a vector space

3) $V = F^n = \{(a_1, \ldots, a_n) : a_i \in F\}$

4) $V = F[x] = \{\text{All polynomials } p(x) \text{ with coefficients in } F\}$

Definition (Subspace). W is a *subspace* of V if:

1. $W \subset V$
2. W forms a subgroup under $+$
3. W is closed under scalar multiplication in F

Now a point to note is that in our discussion of groups, we defined a "structure-preserving" map, called a homomorphism, that preserved the multiplication table of groups upon mapping. In the case of vector spaces, we would like to define a similar "structure-preserving" map, except that we have more structure to preserve: namely, the map must not only be a group homomorphism under $+$ so that $f(u + v) = f(u) + f(v)$, but also a mapping that preserves scalar multiplication: $f(cv) = c \cdot f(v)$ for $c \in F$. But this is precisely our idea of a *linear transformation*; that is, "homomorphisms" between vector spaces are the linear transformations between them.

Definition (Linear Transformation). A mapping $T : V \to W$ between vector spaces is a *linear transformation* if:

1. $T(v + w) = Tv + Tw$
2. $T(cv) = cTv$ for all $c \in F$

Identically to the case of group theory, we have:

$$\ker(T) = \{v \in V : Tv = 0_W\} \subset V$$

$$\text{im}(T) = \{w \in W : w = Tv \text{ for some } v \in V\} \subset W$$

which are both subgroups of V, W respectively.

Definition (Quotient Space). Given vector space V and subspace $W \subset V$, the *quotient space* V/W is defined as the cosets of W under $+$, namely:

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

The quotient space itself forms a vector space over F, with the canonical linear transformation:

$$f : V \to V/W$$

given by $f(v) = v + W$ with $\ker(f) = W$.

Then, $f(v+u) = (v+u)+W = (v+W)+(u+W) = f(v)+f(u)$ and $f(cv) = cv+W = cv+cW = c(v + W) = cf(v)$.

Definition (Linear Combination). Given a set of vectors $\{v_1, \ldots, v_n\}$, a *linear combination* of the vectors is some:

$$w = a_1v_1 + \cdots + a_nv_n$$

for $a_i \in F$.

We can then define the *span of* $(v_1, \ldots v_n)$ as the set of all linear combinations of $v_1, \ldots, v_n$:

$$\text{span}(v_1, \ldots, v_n) = \{w \in V : w = a_1v_1 + \cdots + a_nv_n, a_i \in F\} \subset V$$

and the span is a subspace of V.

Moreover, V is a *finite-dimensional* vector space if there exists a finite set S of vectors in V such that $\text{span}(S) = V$.

Examples (Finite-Dimensional, Span). 1) $V = F^n$: finite-dimensional, since:

$$S = \{v_1 = (1, 0, \ldots, 0), \ldots, v_n = (0, 0, \ldots, 1)\}$$

spans V with $(a_1, \ldots, a_n) = a_1v_1 + \cdots + a_nv_n$.

2) $V = F[x]$: *not* finite-dimensional, since any finite set S of polynomials must have some maximal degree N. However, $F[x]$ contains polynomials of any finite degree, namely of $N+1$. Thus, S does not span $F[x]$.

Finally, we consider the idea of a *linear relation*. A linear relation on $\{v_1, \ldots, v_n\}$ is the equation $a_1v_1 + \cdots + a_nv_n = 0_V$ for some $a_1, \ldots, a_n \in F$, namely a linear combination that yields the null vector in V.

Definition (Linear Independence). The set $\{v_1, \ldots, v_n\} \subset V$ is *linearly independent* in V if:

$$a_1v_1 + \cdots + a_nv_n = 0_V \Leftrightarrow a_1 = \cdots = a_n = 0 \in F$$

That is, if there exists a linear relation on $\{v_1, \ldots, v_n\}$, then the linear relation must be the trivial one, with $a_i = 0$.

Example (Linear Independence). Let $V = \mathbb{R}^3$. Then consider the vectors:

$$S = \{v_1 = (1,0,0), v_2 = (1,1,0), v_3 = (1,2,3)\}$$

Then S is linearly independent, since $a_1v_1 + a_2v_2 + a_3v_3 = 0$ implies that $3a_3 = 0 \Rightarrow a_3 = 0$, so $a_1v_1 + a_2v_2 = 0$. But then this implies that $a_2 = 0$, so that $a_1v_1 = 0$ and thus $a_1 = 0$ since $v_1 \neq 0$.

Definition (Basis). The ordered set $(v_1, \ldots, v_n)$ is a *basis* of V if it:

1. Spans V ($\text{span}(v_1, \ldots, v_n) = V$)
2. Is linearly independent

Lemma 2.2 *If $(v_1, \ldots, v_n)$ is a basis, then any vector $w \in V$ can be uniquely expressed as a linear combination of $v_1, \ldots, v_n$.*

Proof Suppose otherwise, namely that there exist $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ such that $w = a_1v_1 + \cdots + a_nv_n = b_1v_1 + \cdots + b_nc_n$. Then $(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = w - w = 0$, and the vectors are linearly independent, so $a_i - b_i = 0 \Rightarrow a_i = b_i$ for every i.

Proposition 2.3 *Any basis $(v_1, \ldots, v_n)$ for V yields an isomorphism of vector spaces:*

$$f : V \to F^n$$

where if $v = a_1v_1 + \cdots + a_nv_n$ uniquely, then $f(v) = (a_1, \ldots, a_n) \in F^n$.

Proof Note that this mapping is clearly linear, since $f(v+w) = f((a_1+b_1)v_1+\cdots+(a_n+b_n)v_n) = (a_1 + b_1, \ldots, a_n + b_n) = (a_1, \ldots, a_n) + (b_1, \ldots, b_n) = f(v) + f(w)$, and $f(cv) = (ca_1, \ldots, ca_n) = c(a_1, \ldots, a_n) = cf(v)$. It is also surjective, since for any arbitrary $(a_1, \ldots, a_n) \in F^n$, we can construct $v = a_1v_1 + \cdots + a_nv_n \in V$ such that $f(v) = (a_1, \ldots, a_n)$. Finally, it is injective since the kernel of the mapping is given by $f(v) = f(a_1v_1 + \cdots + a_nv_n) = a_1f(v_1) + \cdots + a_nf(v_n) = a_1(1, \ldots, 0) + \cdots + a_n(0, \ldots, 1) = a_1e_1 + \cdots + a_ne_n = (0, \ldots, 0)$ but the vectors e_i form a basis of F^n and are linearly independent, so $a_i = 0$ and $\ker(f) = \{0_V\}$.

Theorem 2.4 *1) If S is a finite set spanning V, there is a subset of S that forms a basis of V.*
2) If L is a linearly independent set of vectors, it can be extended to form a basis of V.

Proof 1) Suppose S spans V and is finite, so $S = \{v_1, \ldots, v_n\}$. If S is linearly independent, then it is a basis. If not, then there exist a_i such that $a_1v_n + \cdots + a_nv_n = 0$ with some $a_i \neq 0$. We can reorder the vectors if necessary such that $a_n \neq 0$. Then, $v_n = -\frac{1}{a_n}(a_1v_1 + \cdots + a_{n-1}v_{n-1})$, and so $v_n \in \text{span}(v_1, \ldots, v_{n-1})$. Thus, $V = \text{span}(S) = \text{span}(v_1, \ldots, v_{n-1})$. We can repeat this procedure until the remaining set is linearly independent while preserving spanning, since the base case $S = \{v_1\}$ is linearly independent.

2) If L spans V then L is a basis. Otherwise, let S be a finite set spanning V. Let $v \in S$ be such that $v \notin L$; such an element must exist, because otherwise the span of L is the same as that of S, so L would span V. Then $L' = L \cup \{v\}$ is linearly independent, since if $L = \{w_1, \ldots, w_m\}$ and $a_1w_1 + \cdots + a_mw_m + bv = 0$, then $b = 0$ since otherwise $v = -\frac{1}{b}(a_1w_1 + \cdots + a_mw_m)$ would be in the span of L. Thus, $a_1w_1 + \cdots + a_mw_m = 0$, but these are linearly independent, so all $a_i = 0$. Thus, L' is indeed linearly independent. If L' spans V, then L' is a basis; otherwise, we can find another element of the spanning set S and repeat. We must eventually reach a basis, since S is finite, while preserving linear independence.

Theorem 2.5 *If $S = \{v_1, \ldots, v_n\}$ spans V and $L = \{w_1, \ldots, w_m\}$ is linearly independent, then $n \geq m$.*

Proof Since S is a spanning set, we can write $w_j = \sum_{i=1}^n a_{ij}v_i$ for any $w_j \in L$. Suppose there exist c_j such that $0_V = \sum_{j=1}^m c_jw_j = \sum_{j=1}^m c_j \left(\sum_{i=1}^n a_{ij}v_i\right) = \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}c_j\right) v_i$. Then a nontrivial linear relation would hold on L if we can find $\sum_{j=1}^m a_{ij}c_j$ for all $i = 1, \ldots, n$. But these are n equations in m unknowns c_j, which always has a nontrivial solution if $m > n$. But if so, then this implies that there are nontrivial c_j that yield a linear relation on L, which implies L is linearly dependent, a contradiction. Thus, $n \geq m$.

Corollary 2.6
1) All bases of V have the same number of elements, defined as the dimension of V.

2) All spanning sets S have $|S| \geq \dim V$.

3) All linearly independent sets L have $|L| \leq \dim V$.

Proof Clearly, (2) and (3) follow from Theorem 2.5. To prove (1), suppose B and B' are two bases of V. Then B spans while B' is linearly independent, so $|B| \geq |B'|$. But similarly, B' spans while B is linearly independent, so $|B'| \geq |B|$ and $|B| = |B'|$.

Corollary 2.7 *Suppose $W \subset V$ is a subspace for finite-dimensional V, and let $(w_1, \ldots, w_m)$ be a basis for W. Then this set can be extended to a basis for V, i.e. $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$.*

Proof Since $(w_1, \ldots, w_m)$ is a basis for W, it is linearly independent. Thus, by Theorem 2.4(2), it can be extended to a basis for V.

Now recall that if $W \subset V$ is a subspace, then this yields a quotient space V/W and a canonical linear map $f : V \to V/W$ given by $f(v) = v + W$.

Proposition 2.8 *If $(w_1, \ldots, w_m)$ is a basis for $W \subset V$ and $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$ is a basis for V, then $(f(v_{m+1}), \ldots, f(v_n))$ is a basis for V/W.*

Proof Consider any $v+W \in V/W$. Then $v \notin W$, so $v = \sum_{i=m+1}^{n} a_i v_i$, so $v+W = \sum_{i=m+1}^{n} a_i f(v_i) \in \operatorname{span}(f(v_{m+1}), \ldots, f(v_n))$. If $v + W = W$, then we simply map $f(0) = W$. Now suppose $W = \sum_{i=m+1}^{n} a_i f(v_i) = f\left(\sum_{i=m+1}^{n} a_i v_i\right) \Rightarrow \sum_{i=m+1}^{n} a_i v_i = 0_V$. But since the elements are part of a basis, they are linearly independent, so all $a_i = 0$ and thus $f(v_i)$ are linearly independent. Thus, $(f(v_{m+1}), \ldots, f(v_n))$ form a basis of V/W.

Corollary 2.9 $\dim V = \dim W + \dim V/W$

Putting everything together, we have an "isomorphism theorem" in terms of vector spaces:

Isomorphism Theorem on Vector Spaces. Suppose $(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ is a basis of V. Then:

1) If $W = \operatorname{span}(v_1, \ldots, v_m)$ and $W' = \operatorname{span}(v_{m+1}, \ldots, v_n)$, then $W \cap W' = \{0\}$ and there exists a linear isomorphism:

$$W \times W' \to V$$

given by $(w, w') \mapsto w + w' \in V$.

2) If $W \subset V$ is any subspace, then there exists another subspace W' such that the composite map given by:

$$W' \hookrightarrow V \twoheadrightarrow V/W$$

where $w' \mapsto w' \mapsto w' + W$ is an isomorphism. In particular,

$$W' \simeq V/W$$

3) For a subspace $W \subset V$:

$$V \simeq W \times V/W$$

4) If $f : V \to U$ is a linear transformation, then:

$$V \simeq \ker(f) \times \operatorname{im}(f)$$

and $\dim V = \dim(\ker(f)) + \dim(\operatorname{im}(f))$.

Proof 1) The map $f : W \times W' \to V$ is clearly a homomorphism since $f(w, w') + f(v, v') = (w + w') + (v + v') = (w + v) + (w' + v') = f(w + v, w' + v')$. It is surjective, since every $v = \sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{m} a_i v_i + \sum_{j=m+1}^{n} a_j v_j = w + w'$ for some $w \in W$ and $w' \in W'$. Finally, it is injective, since if $0 = w + w' = \sum_{i=1}^{n} a_i v_i$ then $a_i = 0$ so that $w = w' = 0$ and $\ker(f) = \{0\}$.

2) If $(v_1, \ldots, v_m)$ is a basis for W, then it can be extended to a basis $(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ for V, and $W' = \text{span}(v_{m+1}, \ldots, v_n)$ forms a subspace. Let $f : W' \hookrightarrow V$ be the inclusion map and $g : V \twoheadrightarrow V/W$ be the surjection, such that $h(w') = g \circ f(w') = w' + W \in V/W$. It is a homomorphism since $h(w_1' + w_2') = g \circ f(w_1' + w_2') = g(w_1' + w_2') = w_1' + w_2' + W = (w_1' + W) + (w_2' + W) = h(w_1') + h(w_2')$. It is surjective since by Proposition 2.8, $(h(v_{m+1}), \ldots, h(v_n))$ forms a basis for V/W. It is injective since if $W = h(w') = w' + W \Rightarrow w' \in W$. However, $w' \in W'$ and $W \cap W' = \{0\}$, so $w' = 0$ and the kernel is trivial.

3) From (1) and (2), $V \simeq W \times W' \simeq W \times V/W$ since $W' \simeq V/W$.

4) $W = \ker(f)$ is a subspace of V, so $V \simeq \ker(f) \times V/\ker(f)$. But by the First Isomorphism Theorem for groups, there exists an isomorphism $\overline{f} : V/\ker(f) \to \text{im}(f)$ such that $\overline{f}(v + \ker(f)) = f(v) \in \text{im}(f)$. Thus, $V/\ker(f) \simeq \text{im}(f)$ and $V \simeq \ker(f) \times \text{im}(f)$.

Change of Basis. Establishing a *basis* on the vector space V yields a number of 1-1 correspondences that are particularly useful tools for discerning the kernels and images of linear transformations in terms of matrices. We have the following correspondences:

1. If V is an n-dimensional vector space over field F:

$$\{\text{Bases of } V\} \leftrightarrow \{\text{Linear isomorphisms } F^n \to V\}$$

where if $B = (v_1, \ldots, v_n)$ is a basis of V, then the corresponding $\rho_B : F^n \to V$ is given by $(a_1, \ldots, a_n) \in F^n \mapsto a_1 v_1 + \cdots + a_n v_n \in V$.
If ρ is a linear isomorphism of $F^n \to V$, then the associated basis is given by $(\rho(e_1), \ldots, \rho(e_n))$.

2. In addition:

$$\{\text{Linear transformations } F^n \to F^m\} \leftrightarrow M_{m \times n}(F)$$

where we denote $f \mapsto [f]$ as the matrix associated to the linear transformation f, which is given by:

$$[f] = \begin{pmatrix} f(e_1) & f(e_2) & \cdots & f(e_n) \end{pmatrix}$$

If $[f] \in M_{m \times n}(F)$ is a $m \times n$ matrix, then the associated linear transformation f is given by $f : v \mapsto [f] \cdot v \in F^m$.

3. Putting correspondences (1) and (2) together: if we have a linear transformation $f : V \to V'$ with associated bases B and B' of dimensions n and m respectively, then the mapping $\rho_{B'}^{-1} f \rho_B : F^n \to F^m$ can be represented as:

$$F^n \underset{\rho_B}{\longrightarrow} V \underset{f}{\to} V' \underset{\rho_{B'}^{-1}}{\longrightarrow} F^m$$

Thus, this is a linear transformation on $F^n \to F^m$, which has the associated matrix:

$$[\rho_{B'}^{-1} f \rho_B] \equiv [f]_B^{B'}$$

In particular, we have proven the following:

Theorem 2.10 *For vector spaces V, V' over a field F of dimensions n, m respectively:*

$$Hom(V, V') \simeq M_{m \times n}(F)$$

Proof We have constructed a bijective mapping from $\text{Hom}(V, V')$ to $M_{m \times n}(F)$, but we have yet to show that this is a homomorphism. To do so, we observe that:

$$[c_1 f_1 + c_2 f_2]_B^{B'} = c_1 [f_1]_B^{B'} + c_2 [f_2]_B^{B'}$$

from the properties of matrices. Moreover, if $f : V \to V'$ and $g : V' \to V''$ with associated bases B, B', B'' respectively:

$$[g \circ f]_B^{B''} = [g]_{B'}^{B''} [f]_B^{B'}$$

from the composition of mappings, so the multiplication holds as well.

The matrix is given by the coefficients of the basis vectors in V mapped to the basis in V'; i.e. column j in $[f]_B^{B'}$ represents the coefficients of $f(v_j)$ in terms of the basis B'. That is, if $B = (v_1, \ldots, v_n)$ and $B' = (w_1, \ldots, w_m)$, then:

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

and so the j^{th} column of $[f]$ would be precisely $(a_{1j}, \ldots, a_{mj})$. Thus, we have:

$$\begin{array}{ccc} V & \underset{\sim}{\longrightarrow} & F^n \\ \downarrow & & \downarrow \\ W & \underset{\sim}{\longrightarrow} & F^m \end{array}$$

4. Finally, we can consider the *change of basis matrix* by supposing that V has two bases B_1, B_2 and V' has two bases B_1', B_2'. In this case, we see that:

$$\begin{aligned} [f]_{B_2}^{B_2'} &= [\rho_{B_2'}^{-1} f \rho_{B_2}] \\ &= [\rho_{B_2'}^{-1} \rho_{B_1'} \rho_{B_1'}^{-1} f \rho_{B_1} \rho_{B_1}^{-1} \rho_{B_2}] \\ &= [\rho_{B_2'}^{-1} \rho_{B_1'}] \cdot [\rho_{B_1'}^{-1} f \rho_{B_1}] \cdot [\rho_{B_1}^{-1} \rho_{B_2}] \\ &= [\rho_{B_2'}^{-1} \rho_{B_1'}] \cdot [f]_{B_1}^{B_1'} \cdot [\rho_{B_1}^{-1} \rho_{B_2}] \end{aligned}$$

Thus, we can define $[\rho_{B_1}^{-1} \rho_{B_2}]$ as the *change of basis matrix*. Then $[\rho_{B_1}^{-1} \rho_{B_2}]^{-1} = [\rho_{B_2}^{-1} \rho_{B_1}]$. In particular, we can consider the case when f is an operator, that is $V = V'$.

Definition If V is a vector space with bases B_1, B_2, then the *change of basis matrix* is:

$$P \equiv [\rho_{B_1}^{-1} \rho_{B_2}]$$

In this case, a linear operator $f : V \to V$ can be represented as a matrix in either basis by the relation:

$$[f]_{B_2}^{B_2} = P^{-1} [f]_{B_1}^{B_1} P$$

Back to $GL_n(F)$. Using the correspondence between $\mathrm{Hom}(V, V)$ and $M_{n \times n}(F)$, for n-dimensional vector space V, we can now construct an isomorphism between $GL(V)$, the linear isomorphisms from V to itself, and $GL_n(F)$, the general linear group.

$$GL(V) = \{\text{Linear isomorphisms } V \to V\} \subset \mathrm{Aut}(V)$$

Theorem 2.11 *If V be an n-dimensional vector space, then:*

$$GL(V) \simeq GL_n(F)$$

Proof Using previous results we have:

$$\begin{aligned} GL(V) &= \{\text{Linear isomorphisms } V \to V\} \\ &= \{\text{Invertible linear operators on } V\} \\ &= \mathrm{Hom}(V, V)^{\times} \\ &\simeq \{\text{Invertible matrices in } M_{n \times n}(F)\} \\ &= GL_n(F) \end{aligned}$$

Why Deal with Vector Spaces? Given that any finite-dimensional vector space is isomorphic to F^n, it seems that we can forget about abstract vector spaces and simply deal with concrete ones of the form F^n with transformations represented as matrices. However, there is an advantage to working in $GL(V)$ rather than $GL_n(F)$; that is, working with linear operators on vector spaces, which are more abstract, rather than concrete matrices in $GL_n(F)$. This is because when we deal with

abstract linear operators, without a choice of basis, that allows us to choose a more convenient basis later on, which yields a simpler form for the operator in terms of its matrix representation.

As an example, consider $T : V \to W$ where $(v_1, \ldots, v_k)$ is a basis for $\ker(T)$. Then we can extend this to $(v_{k+1}, \ldots, v_n, v_1, \ldots, v_k)$, a basis for V. Then as shown earlier, $(Tv_{k+1}, \ldots, Tv_n)$ is a basis for $\text{im}(T) \subset W$, which can be extended to a basis for W. In terms of these bases for V and W, the matrix of T becomes:

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $r = \text{rank}(T)$, since T maps the first r basis vectors $\{v_{k+1}, \ldots, v_n\}$ to $\{Tv_{k+1}, \ldots, Tv_n\}$ and sends the rest to 0, since they are in the kernel.

This naturally leads to the question of when simple representations such as the above exist for *linear operators* (not linear transformations), since we no longer have the freedom to choose *two* bases; we can only choose a single basis B for V, and $T : V \to V$. This problem is equivalent to that of finding *invariant subspaces* of V.

Definition (Invariant Subspace). For a vector space V, the subspace $W \subset V$ is an *invariant subspace* for the operator $T : V \to V$ if $T(W) \subset W$.

Suppose that $(w_1, \ldots, w_m)$ is a basis for W, an invariant subspace, which we can extend to $(w_1, \ldots, w_m, w'_{m+1}, \ldots, w'_n)$ as a basis for V and where $W' = \text{span}(w'_{m+1}, \ldots, w'_n)$ is the complement subspace. We can consider the following cases:

1. If W is invariant but W' not invariant under T, then:
$$[T] = \begin{pmatrix} A & B \\ \mathbf{0} & D \end{pmatrix}$$
2. If both W, W' are invariant subspaces, then $V = W \oplus W'$ (that is, $v = w + w'$ uniquely for every $v \in V$), and:
$$[T] = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$$
3. Extreme case: if there exist n one-dimensional subspaces, i.e. $W_i = F \cdot w_i$, that are invariant under T (namely, $Tw_i = c_i w_i$) then:
$$[T] = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}$$
this decomposes $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$, which are "lines" (one-dimensional subspaces) that are stable under T. In this case, we say that w_i are eigenvectors of T, and c_i are the associated eigenvalues.

Definition (Eigenvector + Eigenvalue). Let V be a vector space over F and T be a linear operator. Then v is an *eigenvector* of T if:

$$Tv = cv$$

for some $c \in F$. Then c is called the *eigenvalue* associated with the eigenvector v.

Non-Examples (Eigenvectors). 1) Consider the rotation of $\mathbb{R}^2$ by an angle θ. The associated matrix is given by:

$$A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

but this has no eigenvectors, since evidently no line is preserved during a rotation.

2) Consider $T : F^2 \to F^2$ given by $Te_1 = e_1$ and $Te_2 = e_1 + e_2$; then the associated matrix is:

$$[T] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Clearly e_1 is an eigenvector with eigenvalue 1. However, there does not exist any other eigenvectors. In particular, we cannot form a basis of eigenvectors.

Consider an arbitrary vector $v = ae_1 + be_2$; for v to be an eigenvector, we need: $Tv = aTe_1 + bTe_2 = ae_1 + b(e_1 + e_2) = (a+b)e_1 + be_2 = cv = cae_1 + cae_2$. However, this implies that $ca = a + b$ and $cb = b$, so since $b \neq 0$ (otherwise v is just a multiple of e_1), we must have $c = 1$. But in this case, we must have $a = a + b$, which is a contradiction if $b \neq 0$. Thus, e_1 is the only eigenvector.

The previous examples lead us to ask when eigenvectors exist for a given operator; a related question is to ask what the possible eigenvalues are. We note that if w is an eigenvector, $w \neq 0$ and $Tw = cw \Rightarrow (T - cI)w = 0$, so that $w \in \ker(T - cI)$. Since this is a nontrivial kernel ($w \neq 0$), $T - cI$ is not bijective, and therefore not invertible. Conversely, if $T - cI$ is not invertible for some c, then c is an eigenvalue for T! This leads us to conclude:

$$\{\text{Eigenvalues of } T\} = \{c \in F : T - cI \text{ is not invertible}\} = \{c \in F : \det([T] - cI) = 0\}$$

Definition (Characteristic Polynomial). For a linear operator T with associated matrix $[T]$, we define the *characteristic polynomial of* T as:

$$f(t) = \det(tI - [T])$$

Lemma 2.12 *The characteristic polynomial of T is well-defined; that is, it is independent of the basis used to obtain the matrix $[T]$.*

Proof Suppose there are two bases of V, for which the matrix representations of T are $[T], [T]'$, and P is the change of basis matrix between the bases. Then $[T]' = P^{-1}[T]P$. Thus:

$$f'(t) = \det(tI - [T]') = \det(tI - P^{-1}[T]P) = \det(P^{-1}(tI - [T])P) = \det(P^{-1}) \cdot f(t) \cdot \det(P) = f(t)$$

Proposition 2.13 *The roots of the characteristic polynomial $f(t)$ for a linear operator T are the eigenvalues of the operator.*

This Proposition allows us to limit the number of possible eigenvalues of any given operator, through the following Lemma:

Lemma 2.14 *A polynomial of degree n over a field F has at most n distinct roots in F.*

Proof Use induction. For $n = 1$, namely $f(t) = ax + b$, we clearly have one root: $x = -b/a$. Suppose every polynomial of degree $n-1$ has at most $n-1$ distinct roots. Consider a polynomial $f(t)$ of degree n. Then if c is a root of $f(t)$, we have $f(t) = (t-c)g(t) + d$ where $g(t)$ is a polynomial of degree $n-1$ and $d \in F$. Since c is a root, $f(c) = 0 \Rightarrow d = 0$ so that $f(t) = (t-c)g(t)$. Now if c' is any other root of $f(t)$, then $c' \neq c$, so $(c' - c) \neq 0$. Thus, $f(c') = 0 \Rightarrow g(c') = 0$, so c' must be a root of $g(t)$ of degree $n-1$. But by assumption, $g(t)$ must have at most $n-1$ roots. Thus, $f(t)$ has at most n roots.

Corollary 2.15 *There exist at most $n = \dim V$ eigenvalues of a linear operator T.*

Examples (Characteristic Polynomials + Eigenvalues).
1) Consider again the rotation in $\mathbb{R}^2$:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

then the characteristic polynomial is:

$$f(t) = \det(tI - A) = \det \begin{pmatrix} t - \cos\theta & \sin\theta \\ -\sin\theta & t - \cos\theta \end{pmatrix} = t^2 - (2\cos\theta)t + 1$$

Note that if $\theta \neq 0, \pi$, then $|2\cos\theta| < 2$, and therefore the discriminant of the polynomial is $b^2 - 4ac = 4\cos^2\theta - 4 < 0$. Thus, there are no real roots and so no eigenvalues over $\mathbb{R}$.

2) For $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ as earlier, we have: $f(t) = (t-1)^2$ and so $t = 1$ is the only eigenvalue, as shown.

3) For general $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have:

$$f(t) = t^2 - (a+d)t + (ad - bc)$$

As a concrete example, consider $A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$; then $f(t) = t^2 - 7t + 10 = (t-5)(t-2)$. Thus, the eigenvalues are 5 and 2. To find the associated eigenvectors, we must solve:

$$Av_1 = 5v_1, Av_2 = 2v_2$$

Now note that whatever they are, v_1, v_2 must be linearly independent if $5 \neq 2$ in F. Otherwise, $v_1 = cv_2$ and therefore $Tv_1 = T(cv_2) = cTv_2 = 2cv_2$. But also: $Tv_1 = 5v_1 = 5cv_2$, which is a contradiction. Thus, more generally:

Theorem 2.16 *Eigenvectors corresponding to distinct eigenvalues are linearly independent.*

Cayley-Hamilton Theorem. Consider the linear operator $T : V \to V$. Then $T \in \mathrm{Hom}(V, V)$, which is a vector space of dimension n^2 (considered in terms of matrices). Then the set:

$$\{I, T, T^2, \ldots, T^{n^2}\}$$

must be linearly dependent, since it consists of $n^2 + 1$ vectors in an n^2-dimensional space. Thus, there exists a nontrivial linear relation:

$$a_0 I + a_1 T + a_2 T^2 + \cdot + a_{n^2} T^{n^2} = 0$$

This defines a polynomial of degree $\leq n^2$ satisfied by T; that is, we can define:

$$p(t) = a_{n^2} t^{n^2} + \cdots + a_1 t + a_0$$

for which $p(T) = 0$.

One may ask whether there is a "smallest degree" polynomial satisfied by the operator T.

Theorem 2.17 (Cayley-Hamilton). *An operator T satisfies its own characteristic polynomial.*

Proof We consider the case of n distinct eigenvectors, such that the characteristic polynomial becomes:

$$f(t) = (t - c_1)(t - c_2) \cdots (t - c_n)$$

Then we can choose a basis of eigenvectors such that:

$$[T] = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}$$

Then $f(T) \equiv f([T]) = ([T] - c_1 I) \cdots ([T] - c_n I)$, and the i^{th} term has $c_j - c_i$ in the diagonal entries and 0 in the $(i,i)^{th}$ entry. Thus, since the product of diagonal matrices is simply the product of the diagonal entries, every diagonal entry in the final product is 0 (since there is a 0 term for every product on the diagonal), and thus T satisfies the polynomial $f(t)$.

3 Symmetries + Group Actions

What ties together concepts from the seemingly disparate domains of group theory and vector space theory is the orthogonal group acting on $\mathbb{R}^n$, and more generally group actions in linear spaces.

Definition (Inner Product). For vector space F^n over a field F, with vectors $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in F^n$, the *inner product* of v, w is defined as

$$\langle v, w \rangle = v_1 w_1 + \cdots + v_n w_n$$

An inner product puts additional structure on the vector space $V = F^n$. Recall that we define the general linear group on F as $GL_n(F) \simeq \{$ Isomorphisms $F^n \to F^n\}$. Then, just as isomorphisms and homomorphisms preserve the *multiplicative structure* of the vector space, we can consider mappings that preserve the *inner product structure*.

Definition (Orthogonal Group). For vector space F^n over a field F, the *orthogonal group* is the subgroup of $GL_n(F)$ that preserves the inner product, denoted $O_n(F)$:

$$O_n(F) = \{A \in GL_n(F) : \langle Av, Aw \rangle = \langle v, w \rangle\}$$

It is easy to check that $O_n(F) \subset GL_n(F)$ is a subgroup. It trivially contains I. Moreover, it is closed under multiplication, since $\langle ABv, ABw \rangle = \langle A(Bv), A(Bw) \rangle = \langle Bv, Bw \rangle = \langle v, w \rangle$. Finally, it contains inverses, since if $A \in O_n(F)$: $\langle v, w \rangle = \langle Iv, Iw \rangle = \langle A(A^{-1}v), A(A^{-1}w) \rangle = \langle A^{-1}v, A^{-1}w \rangle$.

In fact, we know that $\langle e_i, e_j \rangle = \delta_{ij}$ for the standard orthonormal basis on F^n, which allows us to derive a much simpler condition for checking whether a matrix is in the orthogonal group:

Proposition 3.1 $A \in O_n(F) \Leftrightarrow A^t = A^{-1}$.

Proof Suppose that $A \in O_n(F)$. Then $\langle Ae_i, Ae_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$. But the first inner product is the matrix product of the i^{th} and j^{th} columns of A, since Ae_i is the vector representing the i^{th} column of the matrix A. Thus, this implies that $A^t A = I$, and therefore $A^t = A^{-1}$.

Conversely, suppose that $A^t = A^{-1}$. Then $\langle Av, Aw \rangle = (Av)^t (Aw) = v^t (A^t A) w = v^t w = \langle v, w \rangle$. Thus, $A \in O_n(F)$.

Corollary 3.2 *For matrix $A \in O_n(F)$, we have* $\det(A) = \pm 1$.

Proof $\det(A^t A) = \det(I) = 1$, but $\det(A^t A) = \det(A^t)\det(A) = \det(A)^2$. Thus, $\det(A) = \pm 1$.

Definition (Special Orthogonal Group). We define the subgroup of $O_n(F)$ with determinant $+1$ as the *special orthogonal group*, denoted $SO_n(F)$:

$$SO_n(F) = \{A \in O_n(F) : \det(A) = 1\}$$

Note that $SO_n(F)$ has index 2 in $O_n(F)$, since the determinant is a homomorphism; that is, $\det : O_n(F) \to \{\pm 1\}$ is a homomorphism, since $\det(AB) = \det(A)\det(B)$. But this implies that:

$$SO_n(F) = \ker(\det) \lhd O_n(F)$$

and so we can construct a quotient group:

$$O_n(F)/SO_n(F) = \{ \text{ Cosets of } SO_n(F)\} \simeq \{\pm 1\}$$

Proposition 3.3 *Permutation matrices are in $O_n(F)$, which yields an injective homomorphism:*

$$\begin{array}{ccc} A_n & \lhd & S_n \\ \downarrow & & \downarrow \\ SO_n(F) & \lhd & O_n(F) \end{array}$$

Proof We note that if P is the permutation matrix corresponding to a permutation ρ, then $\langle Pe_i, Pe_j \rangle = \langle e_{\rho(i)}, e_{\rho(j)} \rangle = \delta_{ij}$ as desired.

We now move on to the real case, in which $F = \mathbb{R}$. Then, $\langle v, v\rangle = \sum_{i=1}^{n} v_i^2 \geq 0$; that is, we have an *order structure* on $\mathbb{R}$. Moreover, $\langle v, v\rangle = 0 \leftrightarrow v = 0$. Thus, we can define the notion of a "length" on $\mathbb{R}^n$.

Definition (Norm). For $v \in \mathbb{R}^n$, we define the *norm of* v as: $|v| = \sqrt{\langle v, v\rangle}$

This yields the famous Cauchy-Schwarz inequality, which in turn allows us to define *angles* in an arbitrary $\mathbb{R}^n$ space:

Lemma 3.4 ***(Cauchy-Schwarz).*** *For any $v, w \in \mathbb{R}^n$, we have:*

$$-1 \leq \frac{\langle v, w\rangle}{|v| \cdot |w|} \leq 1$$

Thus, we can define the *angle between* v, w as:

$$\theta = \cos^{-1}\left(\frac{\langle v, w\rangle}{|v| \cdot |w|}\right)$$

We finally note that the concepts of length and angle both depended entirely on the inner product. Thus, if a transformation preserves the inner product, it automatically preserves lengths and angles in $\mathbb{R}^n$.

Corollary 3.5 *$O_n(\mathbb{R})$ acts linearly on $\mathbb{R}^n$ and preserves lengths and angles (is an isometry).*

We also state an important result about the eigenvalues of an orthogonal matrix over an arbitrary field F:

Proposition 3.6 *If $A \in O_n(F)$ and v is an eigenvector of A, the associated eigenvalue is ± 1.*

Proof Let λ be the eigenvalue of eigenvector v; then $Av = \lambda v$. Then $\langle v, v\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = \lambda^2\langle v, v\rangle$ so that $\lambda^1 = 1 \Rightarrow \lambda = \pm 1$.

Geometry of SO_2. We now explore the *geometry* of $SO_2(\mathbb{R})$ in more depth to gain intuition for the special orthogonal group. Consider arbitrary $A \in SO_2(\mathbb{R})$, and standard basis (e_1, e_2) for $\mathbb{R}^2$. We note the following:

1. Ae_1, Ae_2 must lie on the unit circle, since A preserves norms.
2. $Ae_1 \perp Ae_2$, since $\langle Ae_1, Ae_2\rangle = \langle e_1, e_2\rangle = 0$.
3. If we let $Ae_1 = (\cos\theta, \sin\theta)$ since it must lie on the unit circle from (1), we note that by (2), we must have $Ae_2 = (-sin\theta, \cos\theta)$ or $Ae_2 = (\sin\theta, -\cos\theta)$.
4. However, we rule out the latter because this would yield a matrix with determinant -1.

Thus, we have found that an arbitrary matrix $A \in SO_2(\mathbb{R})$ must have the form:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

That is, an element of $SO_2(\mathbb{R})$ rotates the plane by an angle θ! Moreover, if we let $\rho_\theta = \text{rot}(\theta)$ denote the transformation of the plane rotating by θ, we note that $\rho_\theta \circ \rho_\psi = \rho_{\theta+\psi} = \rho_\psi \circ \rho_\theta$, so $SO_2(\mathbb{R})$ is in fact *Abelian.*

Proposition 3.7 *There exists an isomorphism of Abelian groups*

$$f : SO_2(\mathbb{R}) \to \{z \in \mathbb{C}^\times : |z| = 1\}$$

such that if A rotates the plane by θ, then $f : A \mapsto e^{i\theta}$.

Geometrically speaking, $SO_2(\mathbb{R})$ is exactly the unit circle; there is a one-to-one correspondence (isomorphism) between elements of $SO_2(\mathbb{R})$ and a point on the unit circle, given by θ.

However, the surprising but true fact is that $O_2(\mathbb{R})$ is *not Abelian*. Thus, what are the elements in $O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$? We begin by noting the following:

Proposition 3.8 *Every element $A \in O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$ has 2 orthogonal eigenvectors, with associated eigenvalues $+1$ and -1.*

Proof The characteristic polynomial of A is given by $x^2 - \text{trace}(A)x - 1 = 0$, since $\det(A) = -1$. We note that the discriminant of this quadratic polynomial is $\text{trace}(A)^2 + 4 > 0$, and so it has two real roots; that is, A is two eigenvalues. Since in terms of the eigenbasis A is a diagonal matrix with eigenvalues on the diagonal, we have $\det(A) = \lambda_1\lambda_2 = -1$, and $\lambda_i = \pm 1$ by Proposition 3.6. Thus, we have $\lambda_1 = +1$, $\lambda_2 = -1$, and there exist eigenvectors v_1, v_2 such that $Av_1 = v_1, Av_2 = -v_2$. Moreover, $v_1 \perp v_2$, since $\langle v_1, v_2\rangle = \langle Av_1, Av_2\rangle = \langle v_1, -v_2\rangle = -\langle v_1, v_2\rangle \Rightarrow \langle v_1, v_2\rangle = 0$.

Geometrically, we can consider $A \in O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$ as reflecting v_2 about the line along v_1; thus, every element in $O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$ has order 2.

Geometry of SO_3. Now let us consider the special orthogonal group in $\mathbb{R}^3$, denoted $SO_3(\mathbb{R})$. We first note the following:

Theorem 3.9 ***(Euler).*** *Any $A \in SO_3(\mathbb{R})$ has an eigenvalue of $+1$. Thus, there exists $v \in \mathbb{R}^3$ such that $Av = v$.*

Proof The characteristic polynomial of any $A \in SO_3(\mathbb{R})$ has degree 3, so it has 3 complex roots. There are two possibilities: 1) $\{\lambda_1, \lambda_2, \lambda_3\}$ are all real; 2) $\{\lambda, z, \bar{z}\}$ such that $z, \bar{z}$ are complex conjugates and $\lambda \in \mathbb{R}$. In case (1), since we must have $\lambda_i = \pm 1$ and $\det(A) = \prod_{i=1}^3 \lambda_i = +1$, we must have at least one $\lambda_i = +1$. In case (2), $z\bar{z} \geq 0$, so $\det(A) = \lambda z\bar{z} = +1$ implies that $\lambda = +1$.

Now to determine the geometric interpretation of an element in $SO_3(\mathbb{R})$, we make the following observation:

Proposition 3.10 *If v is the eigenvector of $A \in SO_3(\mathbb{R})$ with eigenvalue 1, then A preserves the plane perpendicular to v.*

Proof If w is in the plane perpendicular to v, then $w \perp v$, so that $\langle v, w\rangle = 0$. But then $0 = \langle Av, Aw\rangle = \langle v, Aw\rangle$, so $Aw \perp v$ as well.

This implies that A must take the form:

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

where the basis is given by $\{v, e_1, e_2\}$ for e_1, e_2 spanning the plane orthogonal to v. Now if we denote the 2×2 submatrix by A', then we must have that $\det(A') = 1$, since $1 = \det(A) = 1 \cdot \det(A')$. Moreover, A' must be an orthogonal transformation of the plane, since the transformation A restricted to the plane must preserve the inner product. Thus, we find that $A' \in SO_2(\mathbb{R})$, and so:

Proposition 3.11 *For given $A \in SO_3(\mathbb{R})$, there exists a basis (v, e_1, e_2) of $\mathbb{R}^3$ such that the matrix representation of A takes the form:*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

That is, there exists a basis with respect to which A preserves the first basis vector and rotates the plane orthogonal to the vector by some angle θ.

Definition (Isometry). An *isometry* of $\mathbb{R}^n$ is a map $\mathbb{R}^n \to \mathbb{R}^n$ that preserves distance $d(v,w) = |v-w|$ between any two points $v, w \in \mathbb{R}^n$.

Proposition 3.12 *If $m : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry and $m(0) = 0$, then $m = A$ is a linear transformation in $O_n(\mathbb{R})$.*

Proof We note that:

$$|v-w|^2 = \langle v-w, v-w\rangle = |v|^2 + |w|^2 - 2\langle v, w\rangle$$

Thus, since m preserves $|v-w|, |v|, |w|$, it must preserve $\langle v, w\rangle$.

The set of all isometries of $\mathbb{R}^n$ is denoted the *isometry group* of $\mathbb{R}^n$. Let us consider the group G of isometries of $\mathbb{R}^n$. Clearly one subgroup of G is the group of translations, consisting of transformations m such that $m(v) = v + b$. This subgroup is isomorphic to $(\mathbb{R}^n, +)$, or the additive group of $\mathbb{R}^n$, since the map $m(v) = v + b \mapsto b \in \mathbb{R}^n$ is an isomorphism.

Proposition 3.13 *If G is the isometry group on $\mathbb{R}^n$, then $G \approx \mathbb{R}^n \times G_0$, where G_0 is the group of isometries preserving the origin, in the sense that every element $m \in G$ can be written as $t_b \circ m_0$.*

Proof Given an arbitrary motion $m \in G$ such that $m(0) = b$, then we note that $(t_{-b} \circ m)(0) = 0 \Rightarrow t_{-b} \circ m = m_0 \in G_0$ so $m = t_b \circ m_0$.

Proposition 3.14 $G_0 \simeq O_n(\mathbb{R})$

Proof If $m \in G_0$, then m preserves the inner product, since $d(v,w)^2 = d(v,0)^2 + d(w,0)^2 - 2\langle v, w\rangle$ and distances are preserved (as is the origin); thus, $G_0 \subset O_n$. Now to show that $O_n \subset G_0$, suppose that $(e_1, \ldots, e_n)$ is the standard orthonormal basis of $\mathbb{R}^n$. Then for $m \in G_0$, $(m(e_1), \ldots, m(e_n))$ is another orthonormal basis, since it preserves the inner product. Now let $A \in O_n$ with column vectors $\begin{pmatrix} m(e_1) & m(e_2) & \cdots & m(e_n)\end{pmatrix}$. Then, $A^tA = I$ since every element $(A^tA)_{ij} = \langle m(e_i), m(e_j)\rangle = \delta_{ij}$.

4 Bilinear Forms

We build on the concept of an inner product by introducing their generalization, *bilinear forms*. Bilinear forms help to solidify the link between groups, their representations, and vector spaces, and hint at the classical groups.

5 Linear Groups + Group Representations

The classical groups (orthogonal, unitary, and symplectic) precisely tie together ideas from abstract group actions (i.e. action of $GL_n(F)$, stabilizers) and bilinear forms/vector spaces.

6 Ring Theory

To be continued!