MATH456 - Algebra 3

Groups; ring theory; fields.

Based on lectures from Fall 2024 by Prof. Henri Darmon.
Notes by Louis Meunier

Contents

§1 Groups

§1.1 Definitions

↪**Definition 1.1** (Group): A **group** is a set G endowed with a binary composition rule $G \times G \to G, (a,b) \mapsto a \star b$, satisfying

1. $\exists\, e \in G$ s.t. $a \star e = e \star a = a \;\forall\, a \in G$
2. $\forall\, a \in G, \exists\, a' \in G$ s.t. $a \star a' = a' \star a = e$
3. $\forall\, a,b,c \in G, (a \star b) \star c = a \star (b \star c)$.

If the operation on G also commutative for all elements in G, we say that G is *abelian* or *commutative*, in which case we typically adopt additive notation (i.e. $a+b$, $a^{-1} = -a$, etc).

⊛ **Example 1.1**: An easy way to "generate" groups is consider some "object" X (be it a set, a vector space, a geometric object, etc.) and consider the set of symmetries of X, denoted $\mathrm{Aut}(X)$, i.e. the set of bijections of X that preserve some desired quality of X.

1. If X just a set with no additional structure, $\mathrm{Aut}(X)$ is just the group of permutations of X. In particular, if X finite, then $\mathrm{Aut}(X) \cong S_{\#X}$.
2. If X a vector space over some field $\mathbb{F}$, $\mathrm{Aut}(X) = \{T : X \to X \mid \text{linear, invertible}\}$. If $\dim(X) = n < \infty$, $X \cong \mathbb{F}^n$ as a vector space, hence $\mathrm{Aut}(X) = \mathrm{GL}_n(\mathbb{F})$, the "general linear group" consisting of invertible $n \times n$ matrices with entires in $\mathbb{F}$.
3. If X a ring, we can always derive two groups from it; $(R, +, 0)$, which is always commutative, using the addition in the ring, and $(R^\times, \times, 1)$, the units under multiplication (need to consider the units such that inverses exist in the group).
4. If X a regular n-gon, $\mathrm{Aut}(X)$ can be considered the group of symmetries of the polygon that leave it globally invariant. We typically denote this group by D_{2n}.
5. If X a vector space over $\mathbb{R}$ endowed with an inner product $(\cdot,\cdot) : V \times V \to \mathbb{R}$, with $\dim V < \infty$, we have $\mathrm{Aut}(V) = O(V) = \{T : V \to V \mid T(v \cdot w) = v \cdot w \;\forall\, v,w \in V\}$, the "orthogonal group".

↪**Definition 1.2** (Group Homomorphism): Given two groups G_1, G_2, a *group homomorphism* $\varphi : G_1 \to G_2$ is a function satisfying $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in G_1$.

If φ is bijective, we call it an *isomorphism* and say G_1, G_2 are *isomorphic*.

↪**Proposition 1.1**:

- $\varphi\left(1_{G_1}\right) = 1_{G_2}$
- $\varphi\left(a^{-1}\right) = \varphi(a)^{-1}$

⊛ **Example 1.2**: Let $G = \mathbb{Z}/n\mathbb{Z} = \{0, ..., n-1\}$ be the cyclic group of order n. Let $\varphi \in \mathrm{Aut}(G)$; it is completely determined by $\varphi(1)$, as $\varphi(k) = k \cdot \varphi(1)$ for any k. Moreover, it must be then that $\varphi(1)$ is a generate of G, hence $\varphi(1) \in (\mathbb{Z}/n\mathbb{Z})^\times$ (ie the units of the group considered as a ring), and thus

$$\mathrm{Aut}(G) \cong \left((\mathbb{Z}/n\mathbb{Z})^\times, *\right).$$

§1.2 Actions of Groups

↪**Definition 1.3** (Group Action): An *action* of G on an object X is a function $G \times X \to X, (g, x) \mapsto g \cdot$ such that

- $1 \cdot x = x$
- $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$
- $m_g : x \mapsto g \cdot x$ an *automorphism* of X.

↪**Proposition 1.2**: The map $m : G \to \mathrm{Aut}(X), g \mapsto m_g$ a group homomorphism.

Proof. One need show $m_{g_1 g_2} = m_{g_1} \circ m_{g_2}$. ■

↪**Definition 1.4** (G-set): A *G-set* is a set X endowed with an action of G.

↪**Definition 1.5** (Transitive): We say a G-set X is *transitive* if $\forall\, x, y \in X$, there is a $g \in G$ such that $g \cdot x = y$.

A transitive G-subset of X is called on *orbit* of G on X.

↪**Proposition 1.3**: Every G-set is a disjoint union of orbits.

Proof. Define a relation on X by $x \sim y$ if there exists a $g \in G$ such that $g \cdot x = y$. One can prove this is an equivalence relation on X. Equivalence relations partition sets into equivalence classes, which we denote in this case by X/G. The proof is done by remarking that an equivalence class is precisely an orbit. ■

Remark 1.1: As with most abstract objects, we are more interested in classifying them up to isomorphism. The same follows for G-sets.

↪**Definition 1.6**: An *isomorphism of G-sets* is a map between G-sets that respects the group actions. Specifically, if G a group and X_1, X_2 are G-sets, with the action G on X_1 denoted $\star$ and G on X_2 denoted $*$, then an isomorphism is a bijection

$$f : X_1 \to X_2,$$

such that

$$f(g \star x) = g * f(x)$$

for all $g \in G$, $x \in X_1$.

↪**Definition 1.7** (Cosets): Let $H \subseteq G$ be a subgroup of a group G. Then G carries a natural structure as an H set; namely we can define

$$H \times G \to g, \quad (h, g) \mapsto g \cdot h,$$

which can readily be seen to be a well-defined group action. We call, in this case, the set of orbits of the action of H on G *left cosets* of H in G, denoted

$$\begin{aligned} G/H &= \{\text{orbits of } H \text{ acting on } G\} \\ &= \{aH : a \in G\} = \{\{ah : h \in H\} : a \in G\} \subseteq 2^G. \end{aligned}$$

Symmetric definitions give rise to the set of *right cosets* of H in G, denoted $H \setminus G$, of orbits of H acting by left multiplication on G.

Remark 1.2: In general, $G/H \neq H \setminus G$. Further, note that at face value these are nothing more than sets; in general they will not have any natural group structure. They do, however, have a natural structure as G-sets, as a theorem to follow will elucidate.

↪**Theorem 1.1**: Let $H \subseteq G$ be a finite subgroup of a group G. Then every coset of H in G has the same cardinality.

Proof. Define the map $H \mapsto aH$ by $h \mapsto ah$. This is a bijection. ■

Remark 1.3: In general, if one considers the general action of G on some set X, then the orbits X/G need not all have the same size, though they do partition the set. It is in the special case where X a group and G a subgroup of X that we can guarantee equal-sized partitions.

↪**Theorem 1.2** (Lagrange's): Let G be a finite group and H a subgroup. Then

$$\#G = \#H \cdot \#(G/H).$$

In particular, $\#H \mid \#G$ for any subgroup H.

Proof. We know that G/H is a partition of G, so eg $G = H \sqcup H_1 \sqcup \cdots \sqcup H_n$. By the previous theorem, each of these partitions are the same size, hence

$$\begin{aligned}\#G &= \#(H \sqcup H_1 \sqcup \cdots \sqcup H_n) \\ &= \#H + \#H_1 + \cdots + \#H_{n-1} \quad \text{since } H_i\text{'s disjoint} \\ &= n \cdot \#H \quad \text{since each } H \text{ same cardinality} \\ &= \#(G/H) \cdot \#H.\end{aligned}$$

■

↪**Proposition 1.4**: G/H has a natural left-action of G given by

$$G \times G/H \to G/H, \quad (g, aH) \mapsto (ga)H.$$

Further, this action is always transitive.

↪**Proposition 1.5**: If X is a transitive G-set, there exists a subgroup $H \subseteq G$ such that $X \cong G/H$ as a G-set.

In short, then, it suffices to consider coset spaces G/H to characterize G-sets.

Proof. Fix $x_0 \in X$, and define the *stabilizer* of x_0 by

$$H := \mathrm{Stab}_G(x_0) := \{g \in G : gx_0 = x_0\}.$$

One can verify H indeed a subgroup of G. Define now a function

$$f : G/H \to X, \quad gH \mapsto g \cdot x_0,$$

which we aim to show is an isomorphism of G-sets.

First, note that this is well-defined, i.e. independent of choice of coset representative. Let $gH = g'H$, that is $\exists\, h \in H$ s.t. $g = g'h$. Then,

$$f(gH) = gx_0 = (g'h)x_0 = g'(hx_0) = g'x_0 = f(g'H),$$

since h is in the stabilizer of x_0.

For surjectivity, we have that for any $y \in X$, there exists some $g \in G$ such that $gx_0 = y$, by transitivity of the group action on X. Hence,

$$f(gH) = gx_0 = y$$

and so f surjective.

For injectivity, we have that

$$\begin{aligned} g_1x_0 = g_2x_0 &\Rightarrow g_2^{-1}g_1x_0 = x_0 \\ &\Rightarrow g_2^{-1}g_1 \in H \\ &\Rightarrow g_2h = g_1 \text{ for some } h \in H \\ &\Rightarrow g_2H = g_1H, \end{aligned}$$

as required.

Finally, we have that for any coset aH and $g \in G$, that

$$f(g(aH)) = f((ga)H) = (ga)x_0,$$

and on the other hand

$$gf(aH) = g(ax_0) = (ga)x_0.$$

Note that we were very casual with the notation in these final two lines; make sure it is clear what each "multiplication" refers to, be it group action on X or actual group multiplication. ■

↪**Corollary 1.1**: If X is a transitive G set with G finite, then $\#X \mid \#G$. More precisely,

$$X \cong G/\mathrm{Stab}_G(x_0)$$

for any $x_0 \in X$. In particular, the *orbit-stabilizer formula* holds:

$$\#G = \#X \cdot \#\mathrm{Stab}_G(x_0).$$

The assignment $X \to H$ for subgroups H of G is not well-defined in general; given $x_1, x_2 \in X$, we ask how $\mathrm{Stab}_G(x_1)$, $\mathrm{Stab}_G(x_2)$ are related?

Since X transitive, then there must exist some $g \in G$ such that $x_2 = gx_1$. Let $h \in \mathrm{Stab}(x_2)$. Then,

$$hx_1 = x_2 \Rightarrow (hg)x_1 = gx_1 \Rightarrow g^{-1}hgx_1 = x_1,$$

hence $g^{-1}hg \in \mathrm{Stab}\ (x_1)$ for all $g \in G, h \in \mathrm{Stab}(x_2)$. So, putting $H_i = \mathrm{Stab}\ (x_i)$, we have that

$$H_2 = gH_1g^{-1}.$$

This induces natural bijections

$$\begin{aligned} \{\text{pointed transitive } G-\text{sets}\} &\leftrightarrow \{\text{subgroups of } G\} \\ (X, x_0) &\rightsquigarrow H = \mathrm{Stab}(x_0) \\ (G/H, H) &\rightsquigarrow H, \end{aligned}$$

and

$$\begin{aligned} \{\text{transitive } G-\text{sets}\} &\leftrightarrow \{\text{subgroups of } G\}/\text{ conjugation} \\ H_i &= gH_jg^{-1}, \text{some } g \in G. \end{aligned}$$

Given a G, then, we classify all transitive G-sets of a given size n, up to isomorphism, by classifying conjugacy classes of subgroups of "index n" $:= [G : H] = \frac{\#G}{n} = \#(G/H)$.

⊛ **Example 1.3**:

0. $G, \{e\}$ are always subgroups of any G, which give rise to the coset spaces $X = \{\star\}, G$ respectively. The first is "not faithful" (not injective into the group of permutations), and the second gives rise to an injection $G \hookrightarrow S_G$.
1. Let $G = S_n$. We can view $X = \{1, ..., n\}$ as a transitive S_n-set. We should be able to view X as G/H, where $\#(G/H) = \#X = n = \frac{\#G}{\#}(H) = \frac{n!}{\#H}$, i.e. we seek an $H \subset G$ such that $\#H = \frac{n!}{n} = (n-1)!$.

Moreover, we should have H as the stabilizer of some element $x_0 \in \{1, ..., n\}$; so, fixing for instance $1 \in \{1, ..., n\}$, we have $H = \mathrm{Stab}(1)$, i.e. the permutations of $\{1, ..., n\}$ that leave 1 fixed. But we can simply see this as the permutation group on $n-1$ elements, i.e. S_{n-1}, and thus $X \cong S_n/S_{n-1}$. Remark moreover that this works out with the required size of the subgroup, since $\#S_{n-1} = (n-1)!$.

2. Let $X =$ regular tetrahedron and consider

$$G = \mathrm{Aut}(X) := \{\text{rotations leaving } X \text{ globally invariant}\}.$$

We can easily compute the size of G without necessarily knowing G by utilizing the orbit-stabilizer theorem (and from there, somewhat easily deduce G). We can view the tetrahedron as the set $\{1, 2, 3, 4\}$, labeling the vertices, and so we must have

$$\#G = \#X \cdot \#\,\mathrm{Stab}(1),$$

where $\mathrm{Stab}(1) \cong \mathbb{Z}/3\mathbb{Z}$. Hence $\#G = 12$.

From here, there are several candidates for G; for instance, $\mathbb{Z}/12\mathbb{Z}$, D_{12}, A_4, ... Since X can be viewed as the set $\{1, 2, 3, 4\}$, we can view $X \twoheadrightarrow G \hookrightarrow S_4$, where $\hookrightarrow$ an injective homomorphism, that is, embed G as a subgroup S_4. We can show both D_{12} and $\mathbb{Z}/12\mathbb{Z}$ cannot be realized as such (by considering the order of elements in each; there exists an element in D_{12} of order 6, which does not exist in S_4, and there exists an element in $\mathbb{Z}/12\mathbb{Z}$ of order 12 which also doesn't exist in S_4). We can embed $A_4 \subset S_4$, and moreover $G \cong A_4$. If we were to extend G to include planar reflections as well that preserve X, then our G is actually isomorphic to all of S_4.

4. Let X be the cube, $G = \{\text{rotations of } X\}$. There are several ways we can view X as a transitive G sets; for instance $F =$ faces, $E =$ edges, $V =$ vertices, where $\#F = 6, \#E = 12, \#V = 8$. Let's work with F, being the smallest. Letting $x_0 \in F$, we have that $\mathrm{Stab}\,(x_0) \cong \mathbb{Z}/4\mathbb{Z}$ so the orbit-stabilizer theorem gives $\#G = 24$.

This seems to perhaps imply that $G = S_4$, since $\#S_4 = 24$. But this further implies that if this is the case, we should be able to consider some group of size 4 "in the cube" on which G acts.

§1.3 Homomorphisms, Isomorphisms, Kernels

↪**Proposition 1.6**: If $\varphi : G \to H$ a homomorphism, φ injective iff φ has a trivial kernel, that is, $\ker \varphi = \{a \in G : \varphi(a) = e_H\} = \{e\}$.

↪**Definition 1.8** (Normal subgroup): A subgroup $N \subset G$ is called *normal* if for all $g \in G, h \in N$, then $ghg^{-1} \in N$.

↪**Proposition 1.7**: The kernel of a group homomorphism $\varphi : G \to H$ is a normal subgroup of G.

↪**Proposition 1.8**: Let $N \subset G$ be a normal subgroup. Then $G/N = N \setminus G$ (that is, $gN = Ng$) and G/N a group under the rule $(g_1N)(g_2N) = (g_1g_2)N$.

↪**Theorem 1.3** (Fundamental Isomorphism Theorem): If $\varphi : G \to H$ a homomorphism with $N := \ker \varphi$, then φ induces an injective homomorphism $\overline{\varphi} : G/N \hookrightarrow H$ with $\overline{\varphi}(aN) := \varphi(a)$.

↪**Corollary 1.2**: $\mathrm{im}(\varphi) \cong G/N$, by $\overline{\varphi}$ into $\mathrm{im}(\overline{\varphi})$.

⊛ **Example 1.4**: We return to the cube example. Let $\tilde{G} = \widetilde{\mathrm{Aut}}(X) =$ rotations and reflections that leave X globally invariant. Clearly, $G \subset \tilde{G}$, so it must be that $\#\tilde{G}$ a multiple of 24. Moreover, remark that relfections reverse orientation, while rotations preserve it; this implies that the index of G in $\tilde{G}$ is 2. Hence, the action of $\tilde{G}$ on a set $O = \{\text{orientations on} \mathbb{R}^3\}$ with $\#O = 2$ is transitive. We then have the induced map

$$\eta : \tilde{G} \to \mathrm{Aut}(O) \cong \mathbb{Z}/2$$

with kernel given by all of G; G fixes orientations after all.

Remark now the existence of a particular element in $\tilde{G}$ that "reflects through the origin", swapping each corner that is joined by a diagonal. This is not in G, but notice that it actually commutes with every other element in $\tilde{G}$ (one can view such an element by the matrix $\begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$ acting on $\mathbb{R}^3$). Call this element τ. Then, since $\tau \notin G$, $\tau g \neq g$ for any $g \in G$. Hence, we can write $\tilde{G} = G \sqcup \tau G$; that is, $\tilde{G}$ is a disjoint union of two copies of S_4, and so

$$\tilde{G} \cong S_4 \times \mathbb{Z}/2\mathbb{Z}$$

$$f : S_4 \times \mathbb{Z}/2\mathbb{Z} \to \tilde{G}, \quad (g, j) \mapsto \tau^j g.$$

§1.4 Conjugation and Conjugacy

↪**Definition 1.9**: Two elements $g_1, g_2 \in G$ are *conjugate* if $\exists\, h \in G$ such that $g_2 = hg_1h^{-1}$.

Recall that we can naturally define G as a G-set in three ways; by left multiplication, by right multiplication (with an extra inverse), and by conjugation. The first two are always transitive, while the last is never (outside of trivial cases); note that if $g^n = 1$, then $\left(hgh^{-1}\right)^n = 1$, that is, conjugation preserves order, hence G will preserve the order of 1 of the identity element, and conjugation will thus always have an orbit of size 1, $\{e\}$.

An orbit, in this case, is called a *conjugacy class*.

↪**Proposition 1.9**: Conjugation on S_n preserves cycle shape.

PROOF. Just to show an example, consider $(13)(245) \in S_5$ and let $g \in S_5$, and put $\sigma := g(13)(245)g^{-1}$. Then, we can consider what $\sigma g(k)$ is for each k;

$$\begin{aligned}\sigma(g(1)) &= g(3)\\ \sigma(g(3)) &= g(1)\\ \sigma(g(2)) &= g(4)\\ \sigma(g(4)) &= g(5)\\ \sigma(g(5)) &= g(2),\end{aligned}$$

hence, we simply have $\sigma = (g(1)g(3))(g(2)g(4)g(5))$, which has the same cycle shape as our original permutation. A similar logic holds for general cycles. ■

↪**Definition 1.10**: The cycle shape of $\sigma \in S_n$ is the partition of n by σ. For instance,

$$\begin{aligned}1 &\leftrightarrow 1+1+\ldots+1\\ \sigma = (12\ldots n) &\leftrightarrow n.\end{aligned}$$

⊛ **Example 1.5**: We compute all the "types" of elements in S_4 by consider different types of partitions of 4:

Partition	Size of Class
$1+1+1+1$	1
$2+1+1$	$\binom{4}{2} = 6$
$3+1$	$4 \cdot 2 = 8$ (4 points fixed, 2 possible orders)
4	$3! = 6$ (pick 1 first, then 3 choices, then 2)
$2+2$	3

The converse of ↪Proposition 1.9 actually holds as well:

↪**Theorem 1.4**: Two permutations in S_n are conjugate if and only if they induce the same cycle shape.

Proof. We need to show the converse, that if two permutations have the same cycle shape, then they are conjugate.

We show by example. Let $g = (123)(45)(6), g' = (615)(24)(3) \in S_6$. We can let $h \in S_6$ such that it sends $1 \mapsto 6, 2 \mapsto 1, 3 \mapsto 5$, etc; precisely

$$h = (163542).$$

Remark that h need not be unique! Indeed, we could rewrite $g' = (156)(42)(3)$ (which is the same permutation of course), but would result in

$$h = (1)(25)(36)(4),$$

which is not the same as the h above. ■

⊛ **Example 1.6**: We return to ⊛ Example 1.5, and recall that $S_4 \cong \mathrm{Aut}(\text{cube})$. Can we identify the conjugacy classes of S_4 with "classes" of symmetries in the cube?

Conjugation Class	#	**Cube Symmetry**
1	1	id
(12)	6	Rotations about edge diagonals
(12)(34)	3	Rotations about the face centers by π
(123)	8	Rotations about principal diagonals
(1234)	6	Rotations about the face centers by $\frac{\pi}{2}$

⊛ **Example 1.7**: Let $\mathbb{F}$ be a field and consider the vector space $V = \mathbb{F}^n$. Then

$$\mathrm{Aut}(V) = \mathrm{GL}_n(\mathbb{F}) = \{\text{invertible } n \times n \text{ matrices}\}.$$

Recall that linear transformations are described by matrices, after choosing a basis for V; i.e.

$$\{\text{linear transformations on } V\} \longleftrightarrow M_n(\mathbb{F}) := \{n \times n \text{ matrices with entries in } \mathbb{F}\}.$$

However, this identification *depends* on the choose of basis; picking a different basis induces a different bijection. Suppose we have two bases β, β', then $\beta' = P\beta$ for some $P \in \mathrm{GL}_n(\mathbb{F})$ (P called a "change of basis matrix"). Then for $T : V \to V$, and with $M := [T]_\beta, M' := [T]_{\beta'}$, then as discussed in linear algebra, $M' = PMP^{-1}$. Hence, understanding $M_n(\mathbb{F}) \leftrightarrow \mathrm{Hom}(V \to V)$ can be thought of as understanding $M_n(\mathbb{F})$ as a G-set of $G = \mathrm{GL}_n(\mathbb{F})$ under conjugation; then orbits $\leftrightarrow$ conjugacy classes.

Conjugacy Invariants

- The trace tr and determinant det are invariant under conjugation; $\mathrm{tr}\left(PMP^{-1}\right) = \mathrm{tr}(M)$ and $\det\left(PMP^{-1}\right) = \det(M)$
- spec (M) := set of eigenvalues is a conjugate invariant (with caveats on the field, etc)
- Characteristic polynomial, minimal polynomial

§1.5 The Sylow Theorems

Recall that if $H \subseteq G$ a subgroup, then Lagrange's gives us that $\#H \mid \#G$. We are interested in a (partial) converse; given some integer n such that $n \mid \#G$, is there a subgroup of cardinality n?

To see that this is not true in general, let $G = S_5$. $\#G = 120$; the divisors and the (if existing) subgroups:

$$\begin{aligned}
1 &\to \{1\} \\
2 &\to \{1, (12)\} \\
3 &\to \mathbb{Z}/3\mathbb{Z} \\
4 &\to \mathbb{Z}/4\mathbb{Z}, \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \\
5 &\to \mathbb{Z}/5\mathbb{Z} \\
6 &\to \langle (12)(345) \rangle \cong \mathbb{Z}/6\mathbb{Z}, S_3 \\
8 &\to D_8 \\
10 &\to D_{10} \\
12 &\to A_4 \\
15 &\to \text{None :(}
\end{aligned}$$

There is a unique group of order 15, $\mathbb{Z}/15\mathbb{Z}$; but this would need an element of order 15, which doesn't exist in S_5.

↪**Theorem 1.5** (Sylow 1): Fix a prime number p. If $\#G = p^t \cdot m$ with $p \nmid m$, then G has a subgroup of cardinality p^t.

We often call such a subgroup a *Sylow-p* subgroup of G.

⊛ **Example 1.8**: We consider the Sylow subgroups of several permutation groups.

(S_5) $\#S_5 = 120 = 2^3 \cdot 3 \cdot 5$, so by the Sylow theorem, S_5 contains subgroups of cardinality 8, 3, and 5.

(S_6) We have $\#S_6 = 720 = 2^4 \cdot 3^2 \cdot 5$, so by the Sylow theorem we have subgroups H of order $16, 9$, and 5.

$\#H = 9$ is given by

$$H = \langle (123), (456) \rangle := \left\{ (123)^i (456)^j : 0 \leq i, j \leq 2 \right\} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z},$$

or similarly for any other two disjoint cycles of three elements.

$\#H = 16$ is given by $H \cong D_8 \times S_2$.

(S_7) We have $\#S_7 = 2^4 \cdot 3^2 \cdot 5 \cdot 7$. Subgroups of size $16, 9, 5$ can be simply realized as those from S_6, and of size 7 as just the cyclic subgroup generated by an element of order 7.

(S_8) We have $\#S_8 = 2^7 \cdot 3^2 \cdot 5 \cdot 7$ so we have subgroups of size $128, 9, 5, 7$. The latter 3 subgroups are easy to find; the first is found by

$$H \cong \langle (15)(26)(48)(37), D_8 \times D_8 \rangle,$$

where we can view the first copy of D_8 acting on a square labeled $1, 2, 3, 4$, the second acting on a square labeled $5, 6, 7, 8$, and the distinguished permutation swapping all the vertices ??

↪**Theorem 1.6**: Fix a group G and a prime p. TFAE:

1. There exists a G-set X of cardinality prime to p with no orbit of size 1.
2. There is a transitive G-set of cardinality > 1 and of cardinality prime to p.
3. G has a proper subgroup of index prime to p.

Proof. (1. $\Rightarrow$ 2.) We can write $X = X_1 \sqcup X_2 \sqcup \ldots \sqcup X_t$ where X_i the orbits of the action; note that the action of G on each X_i transitive. Since $p \nmid \#X$, then $\exists\, i$ for which $p \nmid \#X_i$. $\#X_i > 1$ necessarily, since X was assumed to have no orbit of size 1.

(2. $\Rightarrow$ 3.) We have $X \cong G/H$ for some subgroup H, where $H = \mathrm{Stab}_G(x_0)$ for some $x_0 \in X$. Moreover, $\#X = [G : H]$ hence $p \nmid [G : H]$.

(3 $\Rightarrow$ 1.) Given H, take $X = G/H$. Then G necessarily acts transitively on X so X has no orbit of size 1, and has cardinality $\#X = [G : H]$, so X also has cardinality prime to p as $[G : H]$ prime to p. ■

↪**Theorem 1.7**: For any finite group G and any prime $p \mid \#G$ with $\#G = p^t \cdot m$, $m \neq 1$, then (G, p) satisfies the properties of the previous theorem.

Proof. It suffices to prove 1. holds. Let

$$X = \{\text{all subsets of } G \text{ of size } p^t\}.$$

X a G-set; for any $A \in X$, gA also a set of size p^t hence $gA \in X$. Moreover, G acts on X without fixed points; that is, there is no element x in X such that $gx = x$ for every $g \in G$. We have in addition

$$\#X = \binom{p^t \cdot m}{p^t} = \frac{(p^t m)(p^t m - 1)(\cdots)(p^t m - p^t + 1)}{(p^t)!} = \prod_{j=0}^{p^t-1} \left(\frac{p^t m - j}{p^t - j} \right).$$

The max power of p dividing $p^t m - j$ will be the same as the maximum power of p dividing j itself (since $p \mid p^t m$), and by the same logic the same power that divides $p^t - j$. That is, then, the max power of p that divides both numerator and denominator in each term of this product for each j, hence they will cancel identically in each term. Thus, $p \nmid \#X$ as desired.

■

Proof. (Of ↪Theorem 1.5) Fix a prime p and let G be a group of minimal cardinality for which the theorem fails for (G, p). By 3. of ↪Theorem 1.6, there exists a subgroup $H \subsetneq G$ such that $p \nmid [G : H]$. We have $\#G = p^t m$, and $\#H = p^t m'$; since $p \nmid \frac{\#G}{\#H} = \frac{p^t m}{p^t m'} = \frac{m}{m'}$.

Then, by our hypothesis H contains a subgroup N of cardinality p^t; N is also a subgroup of G and thus a Sylow-p subgroup of G, contradicting our hypothesis of minimality. ■

Proof. (A Second Proof of ↪Theorem 1.5) We may write

$$G = C_1 \sqcup C_2 \sqcup \cdots \sqcup C_h,$$

where $C_j = \{gag^{-1} : g \in G\}$. We must have (at least one) some C_j where $\#C_j = 1$, so $C_j = \{a\}$; it must be that a commutes with every $g \in G$. Consider the center of G,

$$Z(G) = \{a : ag = ga \; \forall \, g \in G\}.$$

Note that $Z(G)$ is a subgroup of G;

$$G = Z(G) \sqcup C_1 \sqcup \cdots \sqcup C_{h'},$$

where C_j are the conjugacy classes of size > 1 (all the conjugacy classes of size 1 are included in $Z(G)$). By the orbit-stabilizer theorem, the cardinality of each C_j divides the cardinality of G (and as $Z(G)$ a subgroup, so does the cardinality of $Z(G)$). We call this decomposition a "class equation of G".

With this setup, we assume again G is the smallest group for which the theorem fails for p. We consider the following cases:

Case 1: $p \nmid \#Z(G)$, then at least one C_j must be of cardinality prime to p (if all were divisible by p, then we'd have

$$\#G \underset{p}{\equiv} 0 \underset{p}{\equiv} (\text{not } 0) + 0 + \cdots + 0,$$

which is impossible). Then, $C_j \cong G/H$ for some subgroup H of G, with $\#H = p^t m' < \#G$, so by our assumption H has a Sylow p-subgroup, and thus so does G.

Case 2: $p \mid \#Z(G)$. $Z(G)$ an abelian subgroup, so there exists a subgroup $Z \subseteq Z(G)$ with $\#Z = p$ (why? For every abelian group with $p \mid \#Z(G)$, $Z(G)$ has an element of that order, hence take the cyclic subgroup generated by that element; see following lemma). Then, since Z is a normal subgroup, and we may consider $\overline{G} = G/Z$, which is then a group with cardinality

$$\#\overline{G} = \frac{\#G}{\#Z} = \frac{p^t m}{p} = p^{t-1} \cdot m < \#G,$$

so we may apply the induction hypothesis to $\overline{G}$, i.e. $\overline{G}$ has a Sylow p-subgroup $\overline{H}$ of cardinality p^{t-1}. We have a natural, surjective homomorphism

$$\pi : G \twoheadrightarrow \overline{G} \supseteq \overline{H}.$$

Take

$$H = \bigcup_{gZ \in \overline{H}} gZ,$$

or equivalently, $H = \pi^{-1}\left(\overline{H}\right)$. We have an induced surjective homomorphism

$$\pi : H \twoheadrightarrow \overline{H}$$

with $\ker(\pi) = Z$, so $\overline{H} \cong H/Z$, and thus $\#H = \#\overline{H} \cdot \#Z = p^t$, and thus H a Sylow p-subgroup of G. ■

↪**Lemma 1.1**: Let p be prime. If G a finite abelian group and $p \mid \#G$, then G has an element of order p, i.e. there is a subgroup $Z \subset G$ of cardinality p.

Proof. We can write $\#G = p \cdot m$. Remark that it suffices to find an element g of order t such that $p \mid t$; indeed, then the element $g^{\frac{t}{p}}$ has order p, which exists since then $\frac{t}{p}$ an integer.

Let $g_1, g_2, \ldots, g_t$ be a set of generators for G and put $n_j := \operatorname{ord}\left(g_j\right)$. Define now

$$\varphi : (\mathbb{Z}/n_1\mathbb{Z}) \times (\mathbb{Z}/n_2\mathbb{Z}) \times \cdots \times (\mathbb{Z}/n_t\mathbb{Z}) \to G,$$

$$(a_1, a_2, \ldots, a_t) \mapsto g_1^{a_1} g_2^{a_2} \cdots g_t^{a_t}.$$

One can show that this is a homomorphism; moreover, it is surjective, since any element in G can be written in terms of these generators. Hence, $\#G \mid \#(\mathbb{Z}/n_1\mathbb{Z}) \times (\mathbb{Z}/n_2\mathbb{Z}) \times \cdots \times (\mathbb{Z}/n_t\mathbb{Z}) = n_1 n_2 \cdots n_t$. Since $p \mid \#G$, then it follows too that $p \mid n_1 n_2 \cdots n_t$

and thus there is some n_i such that $p \mid n_i$ ("Gauss's Lemma"). Hence, g_j has order divisible by p. ■

↪**Theorem 1.8** (Sylow 2): If H_1, H_2 are Sylow p-subgroups of G, then $\exists g \in G$ s.t. $gH_1g^{-1} = H_2$.

Proof. Consider G/H_1 as an H_2-set. We may write

$$G/H_1 = X_1 \sqcup X_2 \sqcup \cdots \sqcup X_N,$$

where the X_j's are disjoint orbits, then $\#X_j \mid \#H_2$, so $\#X_j = p^a$, some $a \leq t$. Then, there must be some orbit X_j of cardinality 1; since $p \mid \#X_j$, but $p \nmid G/H_1$, but each must be a power of p hence the power a of some cardinality must be 0. Then, we may write $X_j = \{gH_1\}$. This is fixed by every element in H_2, i.e. $\forall\, h \in H_2, hgH_1 = gH_1$ i.e.

$$(g^{-1}hg)H_1 = H_1,$$

i.e. $g^{-1}hg \in H_1$ for all $h \in H_2$, and thus $g^{-1}H_2g = H_1$. ■

↪**Theorem 1.9** (Sylow 3): The number N_p of distinct Sylow p-subgroups satisfies
1. $N_p \mid m$,
2. $N_p \underset{p}{\equiv} 1$,

where $\#G = p^t m$.

Proof.

1. Let $X = \{\text{all Sylow } p\text{-subgroups}\}$. By Sylow 2, G acts transitively on X by conjugation. Then, by the orbit-stabilizer theorem,

$$\#X = \frac{\#G}{\#N},$$

where N the *normalizer* of $H = \{g \in G : gHg^{-1} = H\}$. We know that $H \subset N$, hence $p^t = \#H \mid \#N$, so $\#X \mid \frac{\#G}{\#H} = m$ and so $\#X \mid m$.

2. Let H be a Sylow p-subgroup and let X be the set of all Sylow p-subgroups as above, viewed as a G-set by conjugation. Again, this is a transitive action. We can also view X as an H-set. Then,

$$X = X_1 \sqcup \cdots \sqcup X_a,$$

where

$$\#X_j \mid \#H = p^t,$$

i.e. $\#X_j = 1, p, p^2, \ldots, p^t$. We claim there is exactly one X_j of size 1. Let $X_j = \{H'\}$ be an orbit of size 1 (remarking that there exists at least one, namely just H itself.) Then, we must have $aH'a^{-1} = H'$ for all $a \in H$. Then, H is contained in the normalizer of H', N,

$$H \subset N = \left\{a \in G : aH'a^{-1} = H'\right\}.$$

$H' \subset N$, but moreover, H' a normal subgroup of N. Then,

$$p \nmid \#(N/H').$$

We have the natural map

$$\varphi : N \to N/H',$$

and we consider $\varphi(H)$; its cardinality must be 1, since it must simultaneously divide p^t and something prime to p. Thus, $H \subset \ker(\varphi) = H'$. But $\#H = \#H'$, and thus $H = H'$. Hence, there is a unique orbit of size 1, just H itself.

Thus, the cardinality of X will be, modulo p, 1. ■

1.5.1 Illustrations of the Sylow Theorems

1. $G = S_4$; $\#G = 2^3 \cdot 3$. The Sylow 8-subgroup is D_8,

$$\{1, (1234), (13)(24), (1432), (12)(34), (14)(23), (13), (24)\}.$$

N_2 must divide 3 and must equal 1 modulo 2, so $N_2 = 1$ or 3. In this case, $N_2 = 3$ indeed; D_8 is not normal in S_4, which it would have to be if $N_2 = 1$. Inside S_4, we also have the "Klein group"

$$V = \{1, (12)(34), (13)(24), (14)(23)\},$$

which is normal in S_4. The resulting quotient

$$S_4/V$$

is then a group of cardinality 6, isomorphic to S_3. Consider the homomorphism

$$\varphi : S_4 \to S_3.$$

S_3 has 3 elements of order 2, $(ab), (ac), (bc)$ which generate subgroups of order 2. If A one of these subgroups of order 2, then $\varphi^{-1}(A)$ is a Sylow 2-subgroup.

2.

↪**Theorem 1.10**: Let p, q be primes with $p < q$, $p \nmid q - 1$. If G is a group of cardinality $p \cdot q$, then $G \cong \mathbb{Z}/pq\mathbb{Z}$.

What if $p \mid q - 1$? Consider, for instance, $p = 2, q = 3$, then S_3 has cardinality $p \cdot q$. More generally, suppose $p = 2$ and q any odd prime. Then $p \mid q - 1$ always, and we may consider D_{2q}.

For $p \neq 2$, consider the field $\mathbb{F}_q = \mathbb{Z}/q\mathbb{Z}$, and let

$$G = \left\{T_{a,b} : \mathbb{F}_q \to \mathbb{F}_q, T_{a,b}(x) := ax + b : a \in \mathbb{F}_q^\times, b \in \mathbb{F}_q\right\}$$

be the group of affine-linear transformations on the field. We have that $\#G = (q-1)q$ ($q-1$ choices for a, q choices for b), and that G *not* abelian;

$$\left(T_{a_1,b_1} \circ T_{a_2,b_2}\right)(x) = a_1(a_2x + b_2) + b_1 = a_1a_2x + a_2b_2 + b_1 = T_{a_1a_2,a_2b_2+b_1}(x) \neq \left(T_{a_2,b_2} \circ T_{a_1,b_1}\right)(x).$$

There exists a subgroup $H \subset \mathbb{F}_q^\times$ with $\#H = p$, since $\mathbb{F}_q^\times$ abelian and $p \mid \#\mathbb{F}_q^\times = q-1$, so we may consider the subgroup of G given by

$$G_{pq} = \left\{T_{a,b} : \mathbb{F}_q \to \mathbb{F}_q : a \in H, b \in \mathbb{F}_q\right\} \subset G,$$

with $\#G_{pq} = p \cdot q$. Let us consider the Sylow subgroups of G_{pq}.

A Sylow p-subgroup can be given by $P := \{T_{a,0} : a \in H\}$, and a Sylow q-subgroup can be given by $\left\{T_{1,b} : b \in \mathbb{F}_q\right\}$. Let N_p, N_q the number of Sylow p-, q-subgroups. By Sylow 3, we know that $N_p \underset{p}{\equiv} 1$ and $N_p \mid q$, hence it must be that $N_p = 1$ *or* q. Similarly, $N_q \underset{q}{\equiv} 1$ and $N_q \mid p$, so it must be that $N_q = 1$ so the Sylow q-subgroup we found is unique, and moreover normal.

Remark that the map

$$T_{a,b} \mapsto a, \qquad G \to \mathbb{F}_q^\times \text{ and } G_{pq} \to H$$

is a homomorphism.

To further investigate if $N_p = 1$ or q, we can see how P behaves under conjugation; if it is normal, then it is unique and so $N_p = 1$, else if we can find any second conjugate subgroup then it must be that $N_p = q$. Consider

$$\left(T_{1,1} \circ T_{a,0} \circ T_{1,-1}\right)(x) = a(x-1) + 1 = ax - a + 1 = T_{a,-a+1}(x) \notin P \text{ if } a = 1,$$

hence P not normal and thus $N_p = q$.

§1.6 Burnside's Counting Lemma

↪**Definition 1.11** (Fixed Point Set): Let G a finite group and X a finite G-set. Given $g \in G$, we denote

$$X^g := \{x \in X \mid gx = x\}.$$

the *fixed-point set* of g, and

$$\mathrm{FP}_X(g) := \#X^g.$$

⊛ **Example 1.9**: If $G = S_4$ acting on $X = \{1,2,3,4\}$, then for instance

$$\mathrm{FP}_X((12)) = 2, \mathrm{FP}_X((12)(34)) = 0.$$

↪**Proposition 1.10**: $\mathrm{FP}_X\left(hgh^{-1}\right) = \mathrm{FP}_X(g)$; we say FP_X a *class function* on G, being constant on conjugacy classes.

Proof. Define $X^g \to X^{hgh^{-1}}$ by $x \mapsto hx$, noting $hgh^1hx = x$ for $x \in X^g$; this is a bijection. ■

↪**Theorem 1.11** (Burnside):

$$\frac{1}{\#G}\sum_{g\in G}\mathrm{FP}_X(g) = \#(X/G) = \#G - \text{orbits on } X.$$

Proof. Let $\Sigma \subseteq G \times X$ such that

$$\Sigma = \{(g,x) : gx = x\}.$$

We will count $\#\Sigma$ in two different ways, by noting that we can "project" Σ either to G or X on the first or second coordinate, respectively. On the one hand (the "G view"), we have

$$\#\Sigma = \sum_{g\in G}\mathrm{FP}_X(g),$$

and on the other (the "X view")

$$\#\Sigma = \sum_{x\in X}\#\mathrm{Stab}_G(x) = \sum_{O\in X/G}\sum_{x\in O}\#\mathrm{Stab}_G(x).$$

The orbit-stabilizer theorem gives us that for any $x \in O$, $\#\mathrm{Stab}_G(x)\cdot\#O = \#G$, hence further

$$\#\Sigma = \sum_{O\in X/G}\sum_{x\in O}\frac{\#G}{\#O} = \sum_{O\in X/G}\#G,$$

where the simplification in the final equality comes from the fact that we remove dependence on x in the inner summation, and we are just summing a constant $\#O$ times. Hence,

$$\#\Sigma = \#(X/G)\cdot\#G,$$

and so bringing in our original computation ("G view"),

$$\sum_{g\in G}\mathrm{FP}_X(g) = \#(X/G)\cdot\#G \Rightarrow \frac{1}{\#G}\sum_{g\in G}\mathrm{FP}_X(g) = \#(X/G),$$

completing the proof. ■

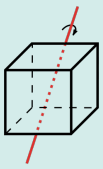
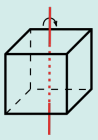
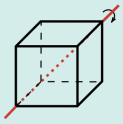
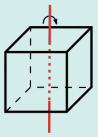
↪**Corollary 1.3**: If X is a transitive G-set with $\#X > 1$, then $\exists g \in G$ such that $\mathrm{FP}_X(g) = 0$.

Proof. By Burnside's,

$$\frac{1}{\#G}\sum_{g\in G}\mathrm{FP}_X(g) = 1,$$

but we have that $\mathrm{FP}_X(1) = \#X > 1$ since 1 fixes everything, so there must be at least a g such that $\mathrm{FP}_X(g) = 0$. ■

⊛ **Example 1.10** (Application of Burnside's): Let $G = S_4 = \mathrm{Aut}(\text{cube})$. We can realize several different (transitive) G-sets; for instance $X = \{1,2,3,4\}, F = \{\text{faces}\}, E = \{\text{edges}\}, V = \{\text{vertices}\}$. We can compute the number of fixed points $\mathrm{FP}_X(g)$ of different elements of G on these G-sets. Recall that it suffices to check one element per conjugacy class of G.

Conj. Class	#	X	F	E	V	**Geometric Desc.**	
id	1	4	6	12	8	id	
(12)	6	2	0	2	0	Rotations about "edge diagonals"	
(12)(34)	3	0	2	0	0	Rotations about "face diagonals", π	
(123)	8	1	0	0	2	Rotations about "principal diagonals"	
(1234)	6	0	2	0	0	Rotations about "face diagonals", $\pi/2$	
$\frac{1}{\#G}\sum \mathrm{FP}_{"X"}(g)$:		1	1	1	1		

The number of orbits, hence, in each case is 1, as we already knew since G acts transitively on all of these sets.

Remark that for two G-sets X_1, X_2, $\mathrm{FP}_{X_1 \times X_2}(g) = \mathrm{FP}_{X_1}(g) \cdot \mathrm{FP}_{X_2}(g)$, where the action of G on $X_1 \times X_2$ defined by $g(x_1, x_2) = (gx_1, gx_2)$. Using this we can consider actions on "pairs" of elements;

Conj. Class	$F \times F$	$F \times V$	$V \times V$
id	36	48	64
(12)	0	0	0
(12)(34)	4	0	0
(123)	0	0	4
(1234)	4	0	0
$\frac{1}{\#G}\sum \mathrm{FP}_{"X"}(g)$:	3	2	4

↪**Definition 1.12** (Colorings of a G-set): Let $C := \{1, 2, ..., t\}$ be a set of "colors". A coloring of X by C is a function $X \to C$. The set of all such functions is denoted C^X. Then, G acts on C^X naturally by

$$G \times C^X \to C^X, \qquad (g, f) \mapsto gf : X \to C, \qquad gf(x) := f\left(g^{-1}x\right).$$

⊛ **Example 1.11**: How many ways may we color the *faces* of a cube with t colors? There are 6 faces with t choices per face, so t^6 faces. More interestingly, how many *distinct* ways are there, up to an automorphism (symmetry) of the cube? G acts on F, and hence on the set of "t-colorings". Let F again be the set of faces and $X := C^F$. Then,

$$\#X = t^6.$$

We would like to calculate the number of orbits of G acting on X, namely $\#(X/G)$. We compute the number of fixed points for each conjugacy class of G; in general, $\#\left(C^F\right)^g = t^{\#(F/\langle g\rangle)} = t^{\#\text{ orbits of } <g> \text{ on } F}$. $(g \leftrightarrow (abc)(de)(f)(g)$ for each element a, say, we have t choices for the coloring of a. Then b, c must be the same color. This repeats for each transposition. etc

Conj. Class	#	F	Shape	X
id	1	6	1^6	t^6
(12)	6	0	$(ab)(cd)(ef)$	t^3
(12)(34)	3	2	$(ab)(cd)$	t^4
(123)	8	0	$(abc)(def)$	t^2
(1234)	6	2	$(abcd)$	t^3

By Burnside's then,

$$\begin{aligned}\#\left(C^F/G\right) &= \frac{1}{24}\sum_{g\in G}\mathrm{FP}_{C^F}(g)\\ &= \frac{1}{24}\left(t^6 + 6t^3 + 3t^4 + 8t^2 + 6t^3\right)\\ &= \frac{1}{24}\left(t^6 + 3t^4 + 12t^3 + 8t^2\right).\end{aligned}$$

Remark that this polynomial does not have integer coefficients, but indeed must have integer outputs for integer t's. This is not obvious.

§1.7 The Exceptional Outer Automorphism of S_6

We consider the fixed points of S_5 acting on various sets, in particular the quotient space S_5/F_{20}, where F_{20} the *Frobenius group* of affine linear transformations $\sigma : x \mapsto ax + b, a \in \mathbb{F}_5^\times, b \in \mathbb{F}_5$. The possible orders of elements $\sigma \in F_{20} \subset S_5$ are

$$1 \leftrightarrow 1^5, 5 \leftrightarrow (01234), 4 \leftrightarrow (1243), 2 \leftrightarrow (14)(23).$$

In particular, each (non-identity) permutation has *at most* one fixed point. One can verify that elements of these types indeed exist in F_{20}.

Remark that to find the cycle shape when acting on S_5/F_{20}, it suffices to check if the permutation given is conjugate to an element in F_{20}, since $(12)gF_{20} = gF_{20} \Leftrightarrow g^{-1}(12)g \in F_{20}$. So, in short, $\sigma \in C$ for some conjugacy class $C \subset S_5$ has no fixed points if it is not conjugate to an element in F_{20}. This holds more generally.

C	#	$\{1,2,3,4,5\}$	$\{1,2,3,4,5,6\}$	S_5/F_{20}	Shape on S_5/F_{20}	Reasoning
id	1	5	6	6	$()$	Identity
(12)	10	3	4	0	$(ab)(cd)(ef)$	Order 2 with no fixed points
$(12)(34)$	15	1	2	2	$(ab)(cd)$	Square of (1234)
(123)	20	2	3	0	$(abc)(def)$	Order 3 with no fixed points
(1234)	30	1	2	2	$(abcd)$	Order 4
(12345)	24	0	1	1	$(abcde)$	Order 5
$(123)(45)$	20	0	1	0	$(abcdef)$	Order 6 with no fixed points

Since F_{20} of index 6 in S_5, we have then a natural injection

$$f : S_5 \to \mathrm{Aut}(S_5/F_{20}) \cong S_6,$$

with image $\widetilde{S_5} := \mathrm{im}(f) \subset S_6$. The cycle shapes of elements in $\widetilde{S_5}$ are precisely those listed in the 2nd-right-most column above.

Now, we can realize $S_5 \subset S_6$ naturally as the permutations that fix, say, 6. However, its clear that while $S_5 \cong \widetilde{S_5}$, they are not conjugate to each other; indeed, $\widetilde{S_5}$ contains cycle shapes that S_5 does not, and since conjugation preserves cycle shape, they certainly cannot be conjugate.

We have that S_6 acts transitively on S_6/S_5, which is isomorphic as a G-set to the typical action of S_6 on 6 numbers. This induces a natural map $S_6 \to \mathrm{Aut}(S_6/S_5) \cong S_6$. One can show that this map is actually an automorphism of S_6, more specifically an *inner automorphism*, one that may be realized as conjugation by an element of S_6. But we can also view S_6 acting on $S_6/\widetilde{S_5}$, which will also be a transitive group action and also induce an automorphism $S_6 \to \mathrm{Aut}\left(S_6/\widetilde{S_5}\right) \cong S_6$. To view how this automorphism acts on elements of S_6, we view how elements of distinct conjugacy classes of S_6 affect $\widetilde{S_5}$. To do so, we need only consider that 1) automorphisms preserve order and 2) automorphisms induce bijections when restricted to conjugacy classes, namely, given a conjugacy class C, it must entirely map to another conjugacy class of the same size. We use the notation (order)(letter if more than one of that order) for conjugacy classes.

C	#	S_6/S_5	$S_6/\widetilde{S_5}$
$1A$	1	$()$	$()$
$2A$	15	(12)	$(ab)(cd)(ef)$
$2B$	45	$(12)(34)$	$(ab)(cd)$
$2C$	15	$(12)(34)(56)$	(ab)
$3A$	40	(123)	$(abc)(def)$
$3B$	40	$(123)(456)$	(abc)
$4A$	90	(1234)	$(abcd)$
$4B$	90	$(1234)(56)$	$(abcd)(ef)$
$5A$	144	(12345)	$(abcde)$
$6A$	120	(123456)	$(abc)(de)$
$6B$	120	$(123)(45)$	$(abcdef)$

In particular, the automorphism $S_6 \to \mathrm{Aut}\left(S_6/\widetilde{S_5}\right)$ interchanges the conjugacy classes $2A$ and $2C$, $3A$ and $3B$, and $6A$ and $6C$.

§2 Rings and Fields

Groups are to symmetries as rings are to numbers.

§2.1 Definitions

↪**Definition 2.1** (Ring): A *ring* is a set R endowed with two operations, $+, \times : R \times R \to R$ satisfying

- (*addition*) $(R, +)$ is an abelian group, with neutral element 0_R and (additive) inverses of $a \in R$ denoted $-a$;
- (*multiplication*) $(R, \times)$ is a *monoid* i.e. it has a neutral element 1_R and is associative;
- (*distribution 1*) $a \times (b + c) = a \times b + a \times c$ for all $a, b, c \in R$;
- (*distribution 2*) $(b + c) \times a = b \times a + c \times a$ for all $a, b, c \in R$.

Remark 2.1:

1. Conventions differ; some texts do not require 1, and call such objects with one a "ring with unity".
2. We will blanketly assume $1 \neq 0$, else R is trivial.
3. 0 is never invertible; $1 \times a = (0 + 1) \times a = 0 \times a + 1 \times a \Rightarrow 0 \times a = 0$ for any $a \in \mathbb{R}$, so in particular there is no a such that $0 \times a = 1$.
4. Exercise: show $(-a) \times (-b) = a \times b$.

⊛ **Example 2.1** (Examples of Rings):

1. $\mathbb{Z}$
2. $\mathbb{Q}$ (essentially $\mathbb{Z}$ appending inverses)
3. Completions of $\mathbb{Q}$; taking $\{\text{Cauchy Sequences}\}/\{\text{Null Sequences}\} = \mathbb{R}$ under the standard distance $d(x,y) = |x-y|$. Alternatively, let p be a prime and take the p-adic metric $|x-y|_p := p^{-\text{ord}_p(x-y)}$ on $\mathbb{Q}$, and consider the completion with respect to $|\cdot|_p$, denoted $\mathbb{Q}_p$, called the *field of p-adic numbers*.
4. $\mathbb{C} := \mathbb{R}[i] = \{a+bi : a,b \in \mathbb{R}\}$
5. Polynomial rings; given a ring R, we may define $R[x] := \{a_0 + a_1 x + \cdots + a_n x^n : a_i \in R\}$ where x a "formal" indeterminate variable.
6. The *Hamilton quaternions*, $\mathbb{H} = \{a+bi+cj+dk : a,b,c,d \in \mathbb{R}\}$, where $i^2 = j^2 = k^2 = -1$ and $ij = -ji = k, jk = -kj = i, ik = -ki = -j$.
7. For any commutative ring R, $M_n(R) = n \times n$ matrices with entries in R is a ring. In particular, associativity of multiplication in $M_n(R)$ follows from the identification of matrices with R-linear functions $R^n \to R^n$ and the fact that function composition is associative.
8. Given a ring R, we can canonically associate two groups, $(R,+,0)$ ("forgetting" multiplication) and $(R^\times,\times,1)$ ("forgetting" addition and restricting to elements with inverses, i.e. *units*).
9. If G is any finite group and R a ring, we may consider $R[G] = \left\{\sum_{g\in G} a_g g : a_g \in R\right\}$, a *group ring*. Addition is defined component-wise, and multiplication

$$\left(\sum_{g\in G} a_g g\right)\left(\sum_{h\in G} b_h h\right) = \sum_{g,h\in G} a_g b_h \cdot gh = \sum_{g\in G}\left(\sum_{h_1\cdot h_2 = h\in G} a_{h_1} b_{h_2}\right) g.$$

§2.2 Homomorphisms

↪**Definition 2.2** (Homomorphism of Rings): A *homomorphism* from a ring R_1 to a ring R_2 is a map $\varphi : R_1 \to R_2$ satisfying:

- $\varphi(a+b) = \varphi(a) + \varphi(b)$ for all $a,b \in R_1$ (that is, φ a group homomorphism of the additive groups $(R_1,+)$, $(R_2,+)$)
- $\varphi(ab) = \varphi(a)\varphi(b)$
- $\varphi\left(1_{R_1}\right) = \varphi\left(1_{R_2}\right)$

↪**Definition 2.3** (Kernel): The *kernel* of a ring homomorphism φ is the kernel as a homomorphism of additive groups, namely

$$\ker(\varphi) = \left\{a \in R_1 : \varphi(a) = 0_{R_2}\right\}.$$

↪**Definition 2.4** (Ideal): A subset $I \subseteq R$ is an *ideal* of R if

1. I an additive subgroup of $(R, +)$, in particular $0 \in I$, I closed under addition and additive inverses
2. I closed under multiplication by elements in R, i.e. for all $a \in R, b \in I$, $ab \in I$ *and* $ba \in I$ (the second condition only being necessary when R non-commutative.)

↪**Proposition 2.1**: If φ is a ring homomorphism, then $\ker(\varphi)$ is an ideal of R_1.

Proof. The first requirement follows from the fact that φ an additive group homomorphism. For the second, let $a \in R_1, b \in \ker(\varphi)$, then $\varphi(ab) = \varphi(a)\varphi(b) = \varphi(a) \cdot 0 = 0$ so $ab \in \ker(\varphi)$. ■

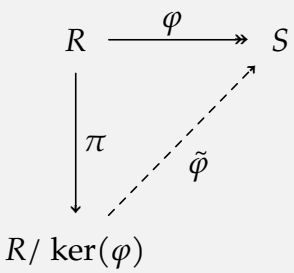
↪**Proposition 2.2**: If I an ideal of R_1, then there exists a ring R_2 and a ring homomorphism $\varphi : R_1 \to R_2$ such that $I = \ker(\varphi)$.

Proof. Let $R_2 = R_1/I = \{a + I : a \in R_1\}$ be the quotient group of R_1 additively. We can define multiplication by $(a + I)(b + I) := ab + I$. One may verify this indeed makes R_2 a ring. Then take φ to be the quotient map

$$\varphi : R_1 \to R_2, \qquad a \mapsto a + I.$$

Then, this is indeed a (surjective) ring homomorphism, with $\ker(\varphi) = I$. ■

↪**Theorem 2.1** (Isomorphism Theorem): Let R be a ring (group) and φ be a surjective homomorphism of rings (groups) $\varphi : R \twoheadrightarrow S$. Then, S is isomorphic to $R/\ker(\varphi)$.



Proof. Define

$$\tilde{\varphi} : R/\ker(\varphi) \to S, \qquad a + \ker(\varphi) \mapsto \varphi(a).$$

One can verify this indeed an isomorphism. ■

§2.3 Maximal and Prime Ideals

↪**Definition 2.5** (Maximal): An ideal $I \subseteq R$ is *maximal* if it is not properly contained in any proper ideal of R, namely if $I \subsetneq I'$ for any other ideal I', then $I' = R$.

↪**Definition 2.6** (Prime): An ideal $I \subseteq R$ is *prime* if $ab \in I$, then a or b in I.

⊛ **Example 2.2**: Let $R = \mathbb{Z}$ and $I = (n) = n\mathbb{Z} = \{na : a \in \mathbb{Z}\}$ for some $n \in \mathbb{N}$. We claim (n) is prime iff n is prime. If n prime, then if $ab \in I$, then $n \mid ab$. By Gauss's Lemma, then n divides at least one of a or b, and hence either a or b in I. Conversely, if n not prime, then we may write $n = ab$ where $|a|, |b| < n$. Then, $a \cdot b \in I$, but n divides neither and so $a, b \notin I$.

↪**Theorem 2.2**: If $I \subseteq \mathbb{Z}$ an ideal, then there exists $n \in \mathbb{Z}$ such that $I = (n)$.

Proof. Consider $\mathbb{Z}/I$. As an abelian group, it is cyclic, generated by $1 + I$. Let $n = \#(\mathbb{Z}/I) = \operatorname{ord}(1 + I)$. If $n = \infty$, then $\mathbb{Z} \to \mathbb{Z}/I$ is injective and $I = (0)$. Else, $I = (n)$.

Alternatively, assume that $I \neq (0)$. Let $n = \min\{a \in I : a > 0\}$. Let $a \in I$, then we may write $a = q \cdot n + r$ where $0 \leq r < n$. $a \in I$ by assumption as is n, and thus so must be $qn \in I$. Hence, $a - qn = r \in I$, and so $r = 0$, by assumption on the minimality of n. ■

↪**Definition 2.7** (Principal Ideal): An ideal of a ring R which is of the form $aR = (a)$ is called *principal*. A ring in which every ideal is of this type is called a *principal ideal ring*.

⊛ **Example 2.3**: $\mathbb{Z}$ is a principal ideal ring. Another example is $R = \mathbb{F}[x]$ where $\mathbb{F}$ a field.

↪**Theorem 2.3**: If I an ideal of $\mathbb{F}[x]$, then I is a principal ideal.

Proof. Let $f(x) \in I$ non-zero and of minimal degree, which necessarily exists if $I \neq (0)$. Put $d := \deg f$. If $g(x) \in I$, then $g(x) = f(x)q(x) + r(x)$ where $\deg r < d$. Then, we have that $r \in I$, so by minimality of d it must be that $r = 0$. ■

Remark 2.2: We conventionally take $\deg(0) = -\infty$ for the sake of the formula $\deg(fg) = \deg f + \deg g$ and taking $-\infty + k = 0$ for any k.

⊛ **Example 2.4**: Consider $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Let $I \subseteq \mathbb{Z}/n\mathbb{Z}$ and consider $\varphi^{-1}(I)$; this is an ideal of $\mathbb{Z}$ and so is principal ie $\varphi^{-1}(I) = (a)$ for some $a \in \mathbb{Z}$. Then, $I = (a + n\mathbb{Z}) \subseteq \mathbb{Z}/n\mathbb{Z}$.

⊛ **Example 2.5**: Let $R = \mathbb{Z}[x] = \{a_n x^n + \cdots + a_1 x + a_0 : a_i \in \mathbb{Z}\}$. Take $I = \{f(x) : f(0) \text{ even}\} \subsetneq \mathbb{Z}[x]$. We claim this an ideal. The subgroup property is clear. If $f(x) \in \mathbb{Z}[x]$ and $g(x) \in I$, then $f(0)g(0) =$ some integer $\cdot$ even integer $=$ even. Elements of I include 2, x, ... any polynomial with a_0 even. If I were principal, then there must exist some element in R dividing both 2 and x; the only possibilities are 1 and -1. This would imply, then, that $I = R$, which is not possible and so I not principal. We may say, however, that I is generated by 2 elements, $I = (2, x) = 2\mathbb{Z}[x] + x\mathbb{Z}[x]$.

⊛ **Example 2.6**: Let $R = \mathbb{F}[x, y]$. Consider $(x, y) = Rx + Ry = \{f : f(0,0) = 0\}$ with typical element $xf(x, y) + yg(x, y)$; these will not have constant terms.

↪**Proposition 2.3**: I is a prime ideal of R iff R/I has no zero divisors (namely an element $x \neq 0$ such that $xy = 0$ for some $y \neq 0$); such a ring is called an *integral domain*.

Proof. Given $a + I, b + I \in R/I$, $(a + I)(b + I) = 0 \Rightarrow ab + I = 0 \Rightarrow ab \in I$. By primality of I, then at least one of $a, b \in I$, so at least one of $a + I, b + I = 0$. ■

Remark 2.3: If R an integral domain, then it satisfies the "cancellation law", namely $\forall\, a \neq 0, ax = ay \Rightarrow x = y$, since we may write $a(x - y) = 0$ hence it must be $x - y = 0 \Rightarrow x = y$.

↪**Theorem 2.4**: I is a maximal ideal $\Leftrightarrow$ R/I is a field.

Proof. ($\Rightarrow$) Let $a + I \in R/I$. If $a + I \neq 0$, then consider the ideal $Ra + I \supsetneq I$. By maximality of I, it must be that $Ra + I = R$. So, anything in R can be written as a "multiple" of a plus an element of I, so in particular $1 \in R$ may be written $1 = ba + i$ for some $i \in I, b \in R$. Passing to the quotient, we find

$$1 + I = (b + I)(a + I) \Rightarrow b + I = (a + I)^{-1} \in R/I,$$

so we indeed have multiplicative inverses.
($\Leftarrow$) Given $J \supsetneq I$, let $a \in J - I$. Then, $a + I \neq 0 \in R/I$, so there exists a b such that $ba + I = 1 + I$ since R/I a field, and hence $1 \in J$ so $J = R$ and thus I maximal. ■

§2.4 Quotients

⊛ **Example 2.7**: Let $R = \mathbb{Z}, I = (n)$, and consider

$$\begin{aligned} R/I = \mathbb{Z}/n\mathbb{Z} &= \{a + n\mathbb{Z} : a + \mathbb{Z}\} \\ &= \{0, 1, 2, \ldots, n-1\} \pmod{n}. \end{aligned}$$

Let $R = \mathbb{F}[x], I = (f(x))$, and consider

$$\begin{aligned} R/I = \mathbb{F}[x]/(f(x)) &= \{p(x) + f(x)\mathbb{F}[x]\} \\ &= \{p(x) : \deg p \leq d - 1 \text{ where } d := \deg f\}. \end{aligned}$$

Remark 2.4: In $R/I, a + I = b + I \leftrightarrow a - b \in I$. If $I = (d)$ principal, then $a + I = b + I \Leftrightarrow d \mid b - a$. For more general quotients (namely, more general ideals) this is a more difficult question.

⊛ **Example 2.8**: Let $R = \mathbb{Z}[x], I = (2, x) = \{f(x) : f(0) \text{ even}\}$, then $\mathbb{Z}[x]/I$ has precisely two elements, and indeed is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. To see this, consider the map

$$\varphi : \mathbb{Z}[x] \to \mathbb{Z}/2\mathbb{Z}, \qquad f(x) \mapsto f(0) \bmod 2,$$

a surjective homomorphism, with

$$\ker(\varphi) = \left\{f(x) : f(0) \underset{2}{\equiv} 0\right\} = \{f(x) : f(0) \text{ even}\} = I,$$

so by the isomorphism theorem, $\mathbb{Z}[x]/\ker(\varphi) \cong \operatorname{im}(\varphi) \Rightarrow \mathbb{Z}[x]/I \cong \mathbb{Z}/2\mathbb{Z}$.

⊛ **Example 2.9**: Let $R = \mathbb{F}[x, y] = \left\{\sum_{i,j=1}^{N} a_{i,j} x^i y^j : a_{i,j} \in \mathbb{F}\right\}$ and $I = (x, y) = \{f(x, y) : f(0, 0) = 0\}$. Then, $R/I \cong \mathbb{F}$ by the map $f(x, y) + I \mapsto f(0, 0)$.

⊛ **Example 2.10**: Let $R = \mathbb{F}[x_1, \ldots, x_n]$ and $I = (f_1, \ldots, f_t)$, for $f_j(x_1, \ldots, x_n) \in R$. Then, consider R/I; this is hard. Let

$$V(I) := \{(x_1, \ldots, x_n) : f_i(x_1, \ldots, x_n) = 0 \text{ for all } i = 1, \ldots, t\}.$$

Then, we may identify $R/I \to$ functions on $V(I)$.

§2.5 Adjunction of Elements

↪**Theorem 2.5**: Given a ring R and $p(x) \in R[x]$, there exists a ring S containing both R and a root of $p(x)$.

Proof. Let $S = R[x]/(p(x))$, $R \to S$ by $a \mapsto a + (p(x))$. Let $\alpha = x + (p(x))$; then $p(\alpha) = p(x) + (p(x)) = 0 + (p(x))$. ∎

↪**Theorem 2.6**: Let $\mathbb{F}$ a field and $f(x) \in \mathbb{F}[x]$ an irreducible, non-zero polynomial. Then, there is a field $\mathbb{K} \supset \mathbb{F}$ such that $\mathbb{K}$ contains a root of $f(x)$

Proof. Let $\mathbb{K} = \mathbb{F}[x]/(f(x))$.

1. This is a field, since $(f(x))$ maximal. To see this, suppose otherwise that $(f(x)) \subsetneq I \subsetneq \mathbb{F}[x]$ for some ideal I of $\mathbb{F}[x]$. Since $\mathbb{F}[x]$ principal, then $I = (g(x))$ for some $g \in \mathbb{F}[x]$. Then, $g(x) | f(x)$ by assumption, but by irreducibility $g(x) = 1$, which implies $I = \mathbb{F}[x]$, or $g(x) = \lambda \cdot f(x)$ for some non-zero $\lambda \in \mathbb{F}$, which implies $I = (f(x))$. In either case, we conclude $(f(x))$ indeed maximal.
2. $\mathbb{F} \hookrightarrow \mathbb{K}$ by the map $\lambda \mapsto \lambda + (f(x))$.
3. We can view $f(t) \in \mathbb{F}[t] \subset \mathbb{K}[t] = (\mathbb{F}[x]/(f(x)))[t]$; indeed, f gains a root in $\mathbb{K}[t]$. Let $\alpha = x + (f(x)) \in \mathbb{K}$. $f(t)$, again viewed as an element of $\mathbb{K}[t]$, evaluated at this $\alpha \in \mathbb{K}$, gives

$$f(\alpha) = f(x + (f(x))) = f(x) + (f(x)) = (f(x)) = 0 \in \mathbb{K},$$

i.e. α indeed a root of $f(x)$.

■

Remark 2.5: In some general R/I with $f(x) \in R[x]$,

$$f(a + I) = f(a) + I$$

for any coset $a + I \in R/I$. To see this, we have $(a + I)^k = (a + I)(a + I) \cdots (a + I) = a^k + I$; we can expand this for more general polynomials in the same manner.

⊛ **Example 2.11**: Let $R = \mathbb{R}$ and $p(x) = x^2 + 1$. Then,

$$\mathbb{C} = \mathbb{R}[x]/\left(x^2 + 1\right) = \left\{a + bx + \left(x^2 + 1\right) : a, b \in \mathbb{R}\right\}.$$

We have that for $x \in \mathbb{C}$, $x^2 = -1 \bmod x^2 + 1$.

⊛ **Example 2.12**: Let $\mathbb{F} = \mathbb{Q}$ and $f(x) = x^2 - 2$. $\sqrt{2}$ irrational so f irreducible over $\mathbb{F}[x]$. Then

$$\mathbb{K} = \mathbb{Q}[x]/\left(x^2 - 2\right) =: \mathbb{Q}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2} : a, b \in \mathbb{Q}\right\}.$$

We can verify this indeed a field; for some arbitrary element $a + b\sqrt{2}$, $\frac{1}{a+b\sqrt{2}}$ a naive inverse, which is indeed equal to $\frac{a-b\sqrt{2}}{a^2-2b^2}$ which one can check indeed in $\mathbb{Q}\left[\sqrt{2}\right]$.

§2.6 Finite Fields

↪**Proposition 2.4**: If $\mathbb{F}$ a finite field, then $\#\mathbb{F} = p^t$ where p a prime number.

Proof. If R is any ring, then there is a unique homomorphism $\mathbb{Z} \to R$ entirely determined by the necessary $0 \mapsto 0_R, 1 \mapsto 1_R$. Then, the map $\varphi : \mathbb{Z} \to \mathbb{F}$ can never be injective, since $\mathbb{Z}$ infinite and $\mathbb{F}$ not. Put $I = \ker(\varphi)$. Then we have an induced injection $\overline{\varphi} : \mathbb{Z}/I \hookrightarrow \mathbb{F}$ by the first isomorphism theorem for rings. We can view then $\mathbb{Z}/I$ as a subring of $\mathbb{F}$, and since $\mathbb{F}$ an integral domain so must be $\mathbb{Z}/I$ and thus I a prime ideal. Prime ideals in $\mathbb{Z}$ are necessarily generated by some prime p, namely $I = (p)$.

Then, $\mathbb{F}$ contains the subfield $\mathbb{Z}/p\mathbb{Z}$. We can view $\mathbb{F}$ as a vector space over $\mathbb{Z}/p\mathbb{Z}$, necessarily finite dimensional. Let $t = \dim(\mathbb{F})$, the dimension as a vector space. Then, we have $\mathbb{F} \cong (\mathbb{Z}/p\mathbb{Z})^t$ as a vector space, and thus the cardinality of $\mathbb{F}$ is p^t. ■

Remark 2.6: One may ask the converse; given p, t, is there a field of cardinality p^t? If so, how many?

If we can find $f(x) \in \mathbb{Z}/p\mathbb{Z}[x]$ irreducible of degree t, then we'd have $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}[x]/(f(x))$ a field of cardinality p^t.

↪**Theorem 2.7**: For all primes p and integers $t \geq 1$, there exists a unique field $\mathbb{K}$ with $\#\mathbb{K} = p^t$.

§3 Modules and Vector Spaces

Groups G are to G-sets as rings R are to R-modules.

§3.1 Modules

↪**Definition 3.1** (Module): An *R-module* is an abelian group M equipped with a map $R \times M \to M$ satisfying, for $\lambda \in R, m_1, m_2, m \in M$:

1. $\lambda(m_1 + m_2) = \lambda m_1 + \lambda m_2$;
2. $\lambda(-m) = -\lambda m$;
3. $\lambda \cdot 0_M = 0_M$;

(that is, for all fixed $\lambda \in R$, left-multiplication by λ $M \to M, m \mapsto \lambda m$ is a group homomorphism from M to itself) and for all $\lambda_1, \lambda_2 \in R, m \in M$,

4. $(\lambda_1 + \lambda_2)m = \lambda_1 m + \lambda_2 m$;
5. $(\lambda_1 \lambda_2)m = \lambda_1(\lambda_2 m)$;
6. $1_R \cdot m = m$.

(that is, multiplication $R \times M \to M$ defines a ring homomorphism $R \to \mathrm{End}(M)$)

Remark 3.1: For an abelian group M,

$$\mathrm{End}(M) := \{f : M \to M \mid f \text{ a group homomorphism}\}$$

is a *ring*, with pointwise addition $(f + g)(m) := f(m) + g(m)$ and multiplication given by composition $(f \circ g)(m) := f(g(m))$.

Remark 3.2: When R a field, we call M a vector space.

↪**Definition 3.2** (Spanning Set): Let M an R-module. A set $\Sigma \subset M$ is called a *spanning set* if for all $m \in M$, there exists $m_1, ..., m_t \in \Sigma$ and $\lambda_1, ..., \lambda_t \in R$ such that

$$m = \lambda_1 m_1 + \cdots + \lambda_t m_t.$$

Remark 3.3: We do not assume Σ finite.

↪**Definition 3.3** (Linear Dependence): A set $\Sigma \subset M$ is *linearly independent* if for all $m_1, ..., m_t \in \Sigma$ and $\lambda_1, ..., \lambda_t$,

$$\lambda_1 m_1 + \cdots + \lambda_t m_t = 0 \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_t = 0.$$

↪**Definition 3.4** (Basis): A set $\Sigma \subset M$ is a *basis* if it is both a spanning set and linearly independent.

Equivalently, Σ a basis if every element in M may be written in a unique way with elements in Σ and scalars in R.

↪**Theorem 3.1**: If $R = \mathbb{F}$ a field and V a vector space over $\mathbb{F}$, then V has a basis.

Proof. Let $\mathcal{L}$ be the set of all linearly independent subsets of V. Inclusion gives a partial ordering on $\mathcal{L}$ ($W_1 \leq W_2$ for $W_1 \subseteq W_2 \in \mathcal{L}$). With this order, $(\mathcal{L}, \leq)$ satisfies the "maximal chain condition"; namely, if $S \subseteq \mathcal{L}$ totally ordered under $\leq$, then there exists an element $\Sigma \in \mathcal{L}$ such that $\Sigma \supseteq B$ for every $B \in S$. Indeed, simply taking $\Sigma = \bigcup_{B \in S} B$ satisfies this condition, remarking that $\Sigma \in \mathcal{L}$ indeed.

We appeal now to Zorn's Lemma; since the maximal chain condition holds, there is an element B in $\mathcal{L}$ which is maximal in the sense that if $B \subsetneq B'$, then $B' \notin \mathcal{L}$. We claim B a basis for V. By definition, it is linearly independent (being a member of $\mathcal{L}$) so it remains to show B spans.

Suppose B is not spanning. Then, there exists some $v \in V$ such that $v \notin \text{Span}(B)$. Consider the set $B \cup \{v\}$; this set is linearly independent. To see this, suppose we take v and $v_1, ..., v_n \in B$, and scalars $\lambda_0, \lambda_1, ..., \lambda_n$ such that $\lambda_0 v + \cdots + \lambda_n v_n = 0$.

If $\lambda_0 = 0$, then by linear independence of B $\lambda_1 = \cdots = \lambda_n = 0$.

Otherwise, if $\lambda_0 \neq 0$, then *since* $\mathbb{F}$ *a field*, we may invert λ_0 and write

$$v = -\lambda_0^{-1}\lambda_1 v_1 + \cdots + -\lambda_0^{-1}\lambda_n v_n \Rightarrow v \in \text{span } (B),$$

a contradiction. Hence, B indeed spanning, and thus a basis. ■

⊛ **Example 3.1**:
1. V is *finitely generated* if it admits a finite spaning set.
2. Suppose $V = \mathbb{R}$ over $\mathbb{F} = \mathbb{Q}$; this is called the "Hamel basis".

Remark 3.4: Existence of bases is *not true* in general for modules over rings; notice in our proof we used the existence of inverses.

Remark 3.5: If $M \subset R$, then M is an R-module if it is an ideal of R.

⊛ **Example 3.2**:
1. Consider $M = \mathbb{Z}^n$ as a $\mathbb{Z}$-module; this has a basis, the standard $\{e_i : i = 1, ..., n\}$ e_i the vector with all zero entries but the ith.
2. Consider $M = \mathbb{Q}$ as a $\mathbb{Z}$-module. Notice that (a) Any two elements in M are linearly dependent; given $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$,
$$(cb)\left(\frac{a}{b}\right) - (ad)\left(\frac{c}{d}\right) = 0.$$
(b) Any finite set in $\mathbb{Q}$ does not span $\mathbb{Q}$ over $\mathbb{Z}$. For instance, if we had
$$S = \left\{\frac{a_1}{b_1}, ..., \frac{a_N}{b_N}\right\} \subset \mathbb{Q},$$
then, for instance,
$$\frac{1}{b_1 b_2 \cdots b_N + 1} \notin \text{Span}(S).$$
As such, $\mathbb{Q}$ has no basis over $\mathbb{Z}$.
3. Consider $M = \mathbb{Z}/n\mathbb{Z}$ as a $\mathbb{Z}$-module, and consider $\{1\}$. It spans $\mathbb{Z}/n\mathbb{Z}$, but is not linearly independent, since for instance, taking $n \in \mathbb{Z}$, $n \cdot 1 = n \equiv 0$ in $\mathbb{Z}/n\mathbb{Z}$.
4. If I is an ideal of R, I has a basis iff I is principal, $I = (a)$, and a is not a zero divisor. If $a, b \in I$, then $ba - ab = 0$ so a, b necessarily linearly dependent.

↪**Definition 3.5** (R-module homomorphism): An *R-module homomorphism* is a map

$$f : M_1 \to M_2$$

between two R-modules M_1, M_2, such that
1. f a group homomorphism;
2. $f(\lambda m) = \lambda f(m)$ for every $\lambda \in R, m \in M$.

Define

$$\ker(f) = \{m \in M_1 \mid f(m) = 0\}.$$

↪**Proposition 3.1**: $\ker(f)$ a subgroup of M_1, closed under scalar multiplication. In particular, it is an *R-submodule* of M_1.

Proof. The subgroup property comes from the fact that f a group homomorphism. For $\lambda \in R, m \in \ker(f), f(\lambda m) = \lambda f(m) = \lambda 0 = 0 \Rightarrow \lambda m \in \ker(f)$. ■

We have in summary the following general properties concerning kernels of homomorphisms in the different categories we've discussed so far:

	kernels	closure property
Non-Abelian Groups	Normal subgroup	Conjugation by G
Rings	Ideal	Multiplication by R
R-modules	R-submodules	Multiplication by R

Just as with groups and rings, we can similarly talk about quotienting modules by submodules.

§3.2 Quotients

↪**Definition 3.6**: If $N \subset M$ are R-modules, then M/N is an R-module with operation defined

$$\lambda \in R, a + N \in M/N \Rightarrow \lambda(a + N) = \lambda a + N.$$

↪**Theorem 3.2** (Isomorphism Theorem): If $f : M_1 \to M_2$ an R-module homomorphism, then it induces an injective homomorphism

$$\bar{f} : M_1/\ker(f) \to M_2, \qquad a + \ker(f) \mapsto f(a).$$

Proof. We just check injectivity.

$$\bar{f}(a + \ker(f)) = 0 \Rightarrow f(a) = 0 \Rightarrow a \in \ker(f) \Rightarrow a + \ker(f) = 0 \in M_1/\ker(f),$$

i.e. $\bar{f}$ has trivial kernel and so is injective. ■

§3.3 Free Modules

↪**Definition 3.7** (Free Module): An R-module M is said to be *free* if it has a basis.

If M is free with a finite basis $e_1, ..., e_n$, then as a module, $M \cong R^n$.

↪**Theorem 3.3**: If M is a free R-module with a finite basis, then any two bases of M have the same cardinality.

Proof. Let I be a proper maximal ideal of R (which exists by a similar argument to the existence of a basis of a vector space, i.e. via Zorn's Lemma argument). Let $F = R/I$; this is a field by maximality. Let

$$IM := \text{span}\{\lambda m : \lambda \in I, m \in M\}.$$

IM is an R-submodule of M. Consider M/IM; this is an R-module as well, but is in fact actually an F-vector space, since I acts as 0 on M/IM. That is, for $\lambda \in R$,

$$(\lambda + I)(m + IM) = \lambda m + IM.$$

If M has a basis of size n, $M \cong R^n$. Then,

$$M/IM \cong F^n$$

as an F-vector space. Then, supposing M has bases $\{e_1, ..., e_n\}, \{f_1, ..., f_m\}$ are two bases of M, then we have that

$$M \cong R^n \cong R^m$$

and so

$$M/IM \cong F^n \cong F^m$$

as F-vector spaces, but by the same theorem for specifically vector spaces, it must be that $n = m$. ■

↪**Definition 3.8** (Rank): If M is free, the cardinality of a basis is called the *rank* of M over R.

Remark 3.6: If M, N are free over R, then M/N need not be free.

For instance, taking $M = \mathbb{Z}, N = m\mathbb{Z}$, for some m, both are free (generated by $1, m$ resp.), but $M/N = \mathbb{Z}/n\mathbb{Z}$ is not free, as any element is linearly dependent.

However, if $R = F$ a field and $W \subset V$ are F-vector spaces, then V/W a vector space, and moreover $\dim(V) = \dim(W) + \dim(V/W)$:

↪**Theorem 3.4**: $\dim(V) = \dim(W) + \dim(V/W)$, where $V \supset W$ are vector spaces over a field F.

Proof. Let $m := \dim(W), n := \dim(V)$. Let $\{v_1, ..., v_m\}$ a basis for W. We complete this to a basis $\{v_1, ..., v_m, v_{m+1}, ..., v_n\}$ for V (note that this is a procedure we *cannot* do, in general, for modules). We claim that

$$\{v_{m+1} + W, ..., v_n + W\}$$

defines a basis for V/W.

- Spanning: given $v + W \in V/W$, $v = \lambda_1 v_1 + ... + \lambda_n v_n$ so $v + W = \lambda_{m+1}(v_{m+1} + W) + \cdots + \lambda_n(v_n + W)$.
- Linear independence: suppose $\lambda_{m+1}, ..., \lambda_n \in F$ are such that $\lambda_{m+1}(v_{m+1} + W) + \cdots + \lambda_n(v_n + W) = 0$. We may rewrite as

  $$(\lambda_{m+1}v_{m+1} + \cdots + \lambda_n v_n) + W = 0,$$

  so there exist $\lambda_1, ..., \lambda_m \in F$ such that

  $$\lambda_{m+1}v_{m+1} + \cdots + \lambda_n v_n = -\lambda_1 v_1 - \cdots - \lambda_m v_m,$$

  which gives a linear combination of our original basis vectors, hence is only possible if $\lambda_1 = \cdots = \lambda_n = 0$, and independence follows.

Hence, indeed our basis is a basis for V/W and so $\dim(V/W) = n - m = \dim(V) - \dim(W)$. ■

3.3.1 Changing Bases

Given a basis $B = (m_1, ..., m_n)$ of an R-module M, we have a natural isomorphism

$$R^n \xrightarrow{\varphi_B} M, \qquad \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto B \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 m_1 + \cdots + \lambda_n m_n.$$

Namely, such an isomorphism is dependent on B. Given another basis $B' = (m'_1, ..., m'_n)$, then there exists some invertible matrix $P \in \mathrm{GL}_n(R)$ such that

$$B' = BP, \qquad (m'_1, ..., m'_n) = (m_1, ..., m_n)P$$

thinking of B, B' as vectors, where the jth column of P is the coordintes of m'_j relative to B.

3.3.2 Homomorphisms Between Free Modules

↪**Definition 3.9** (Free Module Homomorphism): A map

$$T : M_1 \to M_2$$

between two free R-modules of rank n, m respectively is a *free module homomorphism* if T a group homomorphism and is R-linear, i.e.

$$T(\lambda m) = \lambda T(m)$$

for every $\lambda \in R, m \in M$.

↪**Definition 3.10** (Matrix Representation of Module Homomorphism): If $B_1 = (e_1, ..., e_n)$ and $B_2 = (f_1, ..., f_m)$ bases for M_1, M_2 resp., and $T : M_1 \to M_2$ a free module homomorphism, then let

$$M_{T,B_1,B_2} \in M_{m\times n}(R)$$

$$M^{(j)}_{T,B_1,B_2} := j\text{-th column of } M_{T,B_1,B_2} = \text{coordinates of } T\left(e_j\right) \text{ relative to } B_2 =: \left[T\left(e_j\right)\right]_{B_2}.$$

In other words, the following diagram commutes:

$$\begin{array}{ccc} M_1 & \xrightarrow{\quad T \quad} & M_2 \\ \varphi_{B_1} \uparrow \wr & & \wr \uparrow \varphi_{B_2} \\ R^n & \xrightarrow[M_{T,B_1,B_2}]{} & R^m \end{array}$$

i.e., $M_{T,B_1,B_2} = \varphi_{B_2^{-1}} \circ T \circ \varphi_{B_1}$.

↪**Proposition 3.2**: Suppose $M_1 = M_2 =: M$, and $B_1 = B_2 =: B$, i.e. T a homomorphism from M to itself. Consider $M_{T,B} := M_{T,B,B} \in M_n(R)$. Given another basis B', then $M_{T,B}, M_{T,B'}$ are conjugate, namely there exists some $P \in \mathrm{GL}_n(R)$ such that

$$M_{T,B} = PM_{T,B'}P^{-1}.$$

Proof. Let P be such that $B' = BP$. Then for $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n$,

$$\varphi_{B'}\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = B' \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = BP\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\varphi_B \circ P)\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right),$$

i.e. $\varphi_{B'} = \varphi_B \circ P$. We have too $M_{T,B} = \varphi_B^{-1} T \varphi_B, M_{T,B'} = \varphi_{B'}^{-1} T \varphi_{B'}$, so in particular

$$M_{T,B'} = \varphi_{B'}^{-1} T \varphi_{B'} = P^{-1}\left(\varphi_B^{-1} \circ T \circ \varphi_B\right)P = P^{-1} M_{T,B} P,$$

so indeed,

$$M_{T,B} = PM_{T,B'}P^{-1}.$$

■

3.3.3 Matrices Up To Conjugation

Given $M, M' \in M_n(R)$, we'd like to be able to tell when two matrices are conjugate. In particular, we'd like to study $M_n(R)$ (which is a free R-module of rank n^2), modulo conjugation. Namely, we can view $\mathrm{GL}_n(R)$ acting naturally on $M_n(R)$ by conjugation, and we'd like to study the orbits of this action, namely the set

$$M_n(R)/\mathrm{GL}_n(R).$$

Suppose for now $R = F$ a field. Given $M \in M_n(F)$, consider an "evaluation" homomorphism

$$\mathrm{ev}_M : F[x] \to M_n(F), \qquad f(x) \mapsto f(M),$$
$$f(x) := a_n x^n + \cdots + a_1 x + a_0 \mapsto a_n M^n + \cdots + a_1 M + a_0 I_n =: f(M).$$

↪**Proposition 3.3**: ev_M is a homomorphism from $F[x]$ to $M_n(F)$.

ev_M is not injective; notice that $F[x], M_n(F)$ are both vector spaces over F, but $\dim_F(F[x]) = \infty$ (we actually have an explicit basis $B = \{1, x, x^2, \ldots\}$) and $\dim_F(M_n(F)) = n^2$ (with basis $\{E_{i,j} : 1 \leq i, j \leq n\}$). Hence, $\ker(\mathrm{ev}_M)$ is an infinite-dimensional ideal of $F[x]$. $F[x]$ a principal ideal domain, so it must be that

$$\ker(\mathrm{ev}_M) = (p_M(x)),$$

for some $p_M(x) \in F[x]$; let us require that $p_M(x)$ monic, namely the coefficient of its highest power is 1. We can always do this, and in particular makes the generator p_M unique.

↪**Definition 3.11** (Minimal Polynomial): $p_M(x)$ is called the *minimal polynomial* of M.

↪**Proposition 3.4**: $p_M(M) = 0$. In particular, if $f(M) = 0$ for some other $f \in F[x]$, then $p_M | f$.

Proof. This follows from the fact that p_M is a generator for the kernel of ev_M. ■

↪**Proposition 3.5**: Let $T : F^n \to F^n$ be a linear map, B a basis for F^n and $M = [T]_B \in M_n(F)$. Then, for any $A \in \mathrm{GL}_n(F)$, then

$$p_{AMA^{-1}} = p_M,$$

namely, the minimal polynomial of a linear transformation is independent of the choice of basis used to represent it as a matrix; namely p_T, the *minimal polynomial of* T, is well-defined.

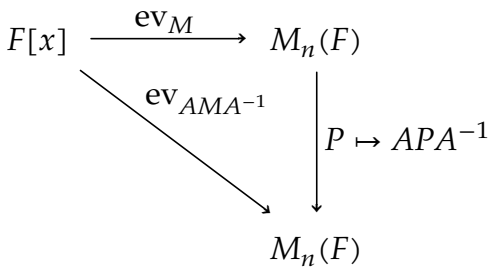
Proof. If $f \in F[x]$, then notice that since

$$\left(AMA^{-1}\right)^k = AM^kA^{-1},$$

and in addition, since conjugation by A linear (namely $A(M_1 + M_2)A^{-1} = AM_1A^{-1} + AM_2A^{-1}$), the map $M \mapsto AMA^{-1}$ is an automorphism of $M_n(F)$. So, in particular we have that

$$f\left(AMA^{-1}\right) = Af(M)A^{-1},$$

so the following commutes,



i.e. $\ker(\mathrm{ev}_M) = \ker\left(\mathrm{ev}_{AMA^{-1}}\right)$, and thus have the same generators. ■

↪**Definition 3.12** (Minimal Polynomial of a Transform): The minimal polynomial p of T is the unique monic polynomial over F satisfying
1. $p(T) = 0$; and
2. if $f(T) = 0$ then $p \mid f$.

More abstractly, we may consider, basis-free, the evaluation homomorphism

$$\mathrm{ev}_T : F[x] \to \mathrm{End}_F(V), \qquad f(x) \mapsto f(T),$$

where $\mathrm{End}_F(V)$ the ring of endomorpisms of V as a F-vector space. We can then equivalently define the minimal polynomial as that which generates $\ker(\mathrm{ev}_T)$.

Note that

$$\deg p_T = \text{smallest non-zero linear combination of } I, T, T^2, T^3, \ldots.$$

Hence, in particular since

$$\dim_F \mathrm{End}_F(V) = n^2,$$

then certainly

$$\deg p_T \leq n^2,$$

since if it were any more, then some power of T would be linearly dependent on another. However, we can bound this further; indeed, we clam that the map ev_T is never surjective (if $n > 1$).

To see this, note that by the isomorphism theorem,

$$\mathrm{im}(\mathrm{ev}_T) \cong F[x]/(p_T(x)).$$

But notice that $F[x]/(p_T(x))$ is a commutative ring, while $\mathrm{End}_F(V)$ is not. Hence, it certainly can't be that $\mathrm{im}(\mathrm{ev}_T)$ equals the whole $\mathrm{End}_F(V)$. What is it?

↪**Theorem 3.5**: $\deg(p_T(x)) \leq n$

Proof. (A first proof) Recall that $f_T(x) := \det(xI - T)$, the characteristic polynomial of T, has degree n. By the Cayley-Hamilton theorem, $f_T(T) = 0$, and thus $p_T \mid f_T$ and so $\deg p_T \leq n$. ■

⊛ **Example 3.3**: Let $T \in \mathrm{GL}_3(\mathbb{F}_2)$ of order 7. What is the minimal polynomial of T?

We know $T^7 - 1 = 0$, so

$$p_T \mid f(x) = (x+1)\big(x^3 + x^2 + 1\big)\big(x^3 + x + 1\big).$$

These resulting polynomials on the RHS are irreducible, hence we know p_T to be some combination of these.

Claim: $\deg p_T \leq 3$

- $p_T(1) \neq 0$. If it were, then T has 1 as an eigenvalue, i.e. $\exists\, v$ such that $Tv = v$. Hence, $T \in \mathrm{Stab}_{\mathrm{GL}_3(\mathbb{F}_2)}(v)$, but $\#\mathrm{Stab}_{\mathrm{GL}_3(\mathbb{F}_2)}(v) = \frac{168}{7} = 24$, and $7 \nmid 24$ so this is impossible. It follows that $x+1$ not a factor of p_T.
- For any $v \neq 0$, $v, T(v), T^2(v)$ are linearly independent, hence are a basis for V. Suppose otherwise, that there exists $a_0, a_1, a_2 \in \mathbb{F}_2$ such that $a_0 v + a_1 T(v) + a_2 T^2(v) = 0$. Then, letting $f(x) = a_2 x^2 + a_1 x + a_0$, we equivalently claim that $f(T)(v) = 0$.

But we know that $\gcd(f, p_T) = 1$ since f irreducible. By the lemma below, $f(T)$ invertible, but this contradicts the fact that there is some v such that $f(T)(v) = 0$, i.e. $f(T)$ has nontrivial kernel. We conclude that indeed $\{v, T(v), T^2(v)\}$ linearly independent.

In short, then, we have that every vector v is a *cyclic vector* for T, namely $\{v, Tv, T^2v\}$ is a basis for V. In particular, there is some $f(x) = x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{F}_2[x]$ such that $f(T)(v) = 0$ i.e. $a_0v + a_1T(v) + a_2T^2(v) + T^3(v) = 0$. We claim $f(T) \equiv 0$. Indeed, we have that

$$0 = T \circ f(T)(v) = f(T)(Tv) = 0 \Rightarrow Tv \in \ker(f(T)),$$

and similarly $T^2v \in \ker(f(T))$, hence $\ker(f(T)) = V$, namely, $f(T)(w) = 0$ for every $w \in V$. Hence,

$$f(T) = 0,$$

so we conclude indeed that p_T indeed of degree at most 3.

So, p_T is one of $x^3 + x^2 + 1$ or $x^3 + x + 1$; it could be either, for general T.

Indeed, there are 2 conjugacy classes in $\mathrm{GL}_3(\mathbb{F}_2)$ of order 7, with the minimal polynomial of those in one conjugacy class $x^3 + x^2 + 1$, the other $x^3 + x + 1$.

↪**Lemma 3.1**: If $\gcd(f, p_T) = 1$, then f is invertible.

Proof. Appealing to the euclidean algorithm, there exist $a, b \in \mathbb{F}[x]$ such that $1 = af + bp_T$. Evaluating on T, we find

$$I = a(T)f(T) + \cancel{b(T)p_T(T)} \Rightarrow I = a(T)f(T) \Rightarrow a(T) = (f(T))^{-1}.$$

■

Proof. (A second proof of ↪Theorem 3.5) If V has a cyclic vector v for T, then we are done, since then

$$\{v, Tv, \ldots, T^{n-1}v\}$$

a basis for V, and so there is some $f(x) \in \mathbb{F}[x]$ of degree n such that

$$f(T)(v) = 0 \Rightarrow f(T)\left(T^k v\right) = T^k f(T)(v) = T^k(0) = 0,$$

i.e. $f(T) \equiv 0$, and thus $p_T \mid f$ and $\deg p_T \leq \deg f = n$.

Otherwise, we proceed by induction on $\dim V$. We prove the statement

$$P_n : \text{if } T \in \operatorname{End}(V) \text{ with } \dim \operatorname{End}(V) = n, \text{then } \deg p_T \leq n.$$

Suppose The case for P_{n-1} and let V be of dimension n. Let $v \in V - \{0\}$, and let

$$W = \operatorname{span}\{v, Tv, T^2v, \ldots\}.$$

If v was cyclic, we are in the previous case and we are done, hence assume $W \subsetneq V$. Remark W a T-stable subspace; T maps vectors in W to vectors in W. Let $T_W : W \to W$ denote the restriction of T to W. This induces a linear transformation

$$\overline{T} : V/W \to V/W, \qquad v + W \mapsto Tv + W,$$

which is well defined, since if $v_1 + W = v_2 + W$, then v_1, v_2 differ by something in W, i.e. $v_1 - v_2 \in W$ so $T(v_1 - v_2) \in W$ as well since W stable under T. It follows that $T(v_1) - T(v_2) \in W$.

We know that $\deg p_{T_W} \leq \dim W$ and likewise $\deg p_{\overline{T}} \leq \dim(V/W)$ by the induction hypothesis. We claim $p_{T_W} p_{\overline{T}}$ vanishes on T, namely $p_T \mid p_{T_W} p_{\overline{T}}$.

Note that $p_{\overline{T}}(T)$ maps V to W, and $p_{T_W}(T)$ maps W to zero. Hence,

$$p_{T_W}(T) \circ p_{\overline{T}}(T) = 0,$$

and so indeed $p_{T_W} p_{\overline{T}}$ vanishes on T and so a multiple of p_T.

So, it follows that $\deg p_T \leq \deg p_{T_W} + \deg p_{\overline{T}}$; by the induction hypothesis, $\deg p_{T_W} \leq \dim W$ and $\deg p_T \leq \dim V/W$ and thus $\deg p_T \leq \dim W + \dim V/W = \dim V$, as we aimed to show. ■

↪**Lemma 3.2**: If $T : V \to V$ a linear transformation and W a T-stable subspace, then T induces

$$\overline{T} : V/W \to V/W, \qquad v + W \mapsto T(v) + W.$$

Remark 3.7: It is *not* always true that W has a T-stable complement.

For instance let $V = F^2$ and $T\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix}, T\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}1\\1\end{pmatrix}$, i.e. $T = \begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$. Then, $W = F\begin{pmatrix}1\\0\end{pmatrix}$ is stable, and $W' = F\begin{pmatrix}\lambda\\1\end{pmatrix}$ is complementary. But notice $T\begin{pmatrix}\lambda\\1\end{pmatrix} = \begin{pmatrix}\lambda+1\\1\end{pmatrix} \notin W'$, for any λ, hence W' not T-stable.

3.3.4 Zeros of the Minimal Polynomial

↪**Definition 3.13** (Eigenvalue): λ is an *eigenvalue* of T if there is a non-zero vector $v \in V$ with the property $Tv = \lambda v$.

↪**Theorem 3.6**: If $p_T(\lambda) = 0$ for some $\lambda \in F$, then λ an eigenvalue of T.

Proof. $p_T(\lambda) = 0$ hence $p_T(x) = (x - \lambda)q(x)$ for some polynomial q of one degree smaller than p_T. Then,

$$0 = p_T(T) = (T - \lambda I) \circ q(T).$$

Note that $q(T) \neq 0$, by assumption of minimality of p_T. Hence, we have in particular that $\text{im}(q(T))$ contained in $\ker(T - \lambda I)$. Let $v \in \text{im}(q(T))$ non-zero. Then, $(T - \lambda I)v = 0 \Rightarrow Tv = \lambda v$. ■

The converse holds as well:

↪**Theorem 3.7**: If λ is an eigenvalue of T, then $p_T(\lambda) = 0$.

Proof. Let v be such that $Tv = \lambda v$. Let $g(x)$ be any polynomial in $F[x]$. Then, $g(T)(v) = g(\lambda)v$, namely, $g(\lambda)$ an eigenvalue of $g(T)$ with vector v. Specify g to be p_T. Then,

$$0 = p_T(T)(v) = p_T(\lambda)v.$$

but $v \neq 0$ hence $p_T(\lambda) = 0$ as desired. ■

§3.4 The Primary Decomposition Theorem

↪**Theorem 3.8** (Primary Decomposition): Let $T : V \to V$ a linear transformation on a vector space V over a field F. Supopse $p_T(x) = p_1(x)p_2(x)$ with $\gcd(p_1, p_2) = 1$. Then, there exists unique subspaces $V_1, V_2 \subseteq F$ such that
1. $V = V_1 \oplus V_2$
2. V_j is stable under T, and the minimal minimal polynomial of $T_j := T|_{V_j}$ is $p_j(x)$.

⊛ **Example 3.4**: If T idempotent i.e. $T^2 = T$, then $p_T(x) = x^2 - x = x(x - 1)$, $p_1(x) = x$, $p_2(x) = x - 1$. Then, $V_1 = \ker(T)$, $V_2 = \text{im}(T)$; on V_1, T acts as 0, on V_2, T acts as the identity.

↪**Proposition 3.6** (Chinese Remainder Theorem):

$$F[x]/(p_T(x)) \cong F[x]/(p_1(x)) \times F[x]/(p_2(x)).$$

Proof. Consider $\varphi : F[x] \to F[x]/(p_1(x)) \times F[x]/(p_2(x))$ given by $f(x) \mapsto (f(x) + (p_1(x)), f(x) + (p_2(x)))$. This has kernel

$$\ker(\varphi) = \{f : p_1|f, p_2|f\} = \{f : p_1p_2|f\} = \{f : p_T|f\} = (p_T(x)),$$

since p_1, p_2 relatively prime. Hence, we have an induced injection

$$\overline{\varphi} : F[x]/\left(p_{T(x)}\right) \hookrightarrow F[x]/(p_1(x)) \times F[x]/(p_2(x)).$$

Moreover, as a vector space

$$\dim F[x]/(p_T(x)) = \deg(p_T(x)), \qquad \dim(F[x]/(p_1(x)) \times F[x]/(p_2(x))) = \deg(p_1(x)) + \deg(p_2(x)),$$

and since $\deg(p_T) = \deg(p_1) + \deg(p_1)$, the dimensions of both sides agree. Hence, the map $\overline{\varphi}$ also surjective and thus the isomorphism we sought. ■

Remark 3.8: Given R_1, R_2 rings, $R_1 \times R_2$ a ring. If M_1, M_2 are modules over R_1, R_2, then $M_1 \times M_2$ is an $(R_1 \times R_2)$-module by the action $(\lambda_1, \lambda_2)(m_1, m_2) = (\lambda_1 m_1, \lambda_2 m_2)$.

↪**Theorem 3.9**: If M is any module over $R_1 \times R_2$, then there are R_j-modules M_j, $j = 1, 2$, such that $M \cong M_1 \times M_2$.

Proof. Consider $i_1 : R_1 \to R_1 \times R_2, a \mapsto (a, 0)$. This is *not* a ring homomorphism ($1 \nrightarrow 1$), but does include $R_1 \subset R_2$ as an ideal in $R_1 \times R_2$. Define $M_1 = (1,0)M, M_2 = (0,1)M$. We claim $M = M_1 \times M_2$.

Given $m \in M$, $m = (1,1)m = (1,0)m + (0,1)m$. Putting $m_1 := (1,0)m, m_2 := (0,1)m$ gives $m = m_1 + m_2$ with $m_1, m_2 \in M_1, M_2$ resp, hence $M_1 \times M_2$ spans M.

We now wish to show $M_1 \cap M_2 = \{0\}$. Let $m \in M_1 \cap M_2$. Then, $m = (1,0)m_1 = (0,1)m_2$. Multiplying both sides by $(1,0)$ gives that $m = (1,0)m_1 = 0$, and the claim follows. ■

Proof. (Of ↪Theorem 3.8) If $p_T(x) = p_1(x)p_2(x)$ as given, notice V is a module over $F[x]/(p_T(x)) = F[x]/(p_1(x)) \times F[x]/(p_2(x))$. By the previous theorem, then, $V = V_1 \oplus V_2$, where V_1 a $F[x]/(p_1(x))$-module, V_2 a $F[x]/(p_2(x))$-module. On V_i, $p_i(T) = 0$. ■

↪**Theorem 3.10** (PDT 2): If $p_T(x) = p_1(x)^{e_1} \cdots p_t(x)^{e_t}$ where $p_1, \ldots, p_t$ irreducible, then

$$V = V_1 \oplus \cdots \oplus V_t,$$

where

$$p_T|_{V_j} = p_j(x)^{e_j}$$

for each $j = 1, \ldots, t$.

↪**Theorem 3.11** (PDT 3): Suppose F is algebraically closed, so

$$p_T(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_t)^{e_t},$$

then

$$V = V_1 \oplus \cdots \oplus V_t,$$

where

$$p_T|_{V_j} = \left(x-\lambda_j\right)^{e_j}.$$

If $e_1 = \cdots = e_t = 1$, then $T|_{V_j}$ = multiplication by λ_j, so in particular V_j the eigenspace for λ_j and T diagonalizable.

↪**Corollary 3.1**: If $p_T(x) = (x-\lambda_1)\cdots(x-\lambda_t)$ where $\lambda_1, \ldots, \lambda_t$ distinct, then T is diagonalizable.

Remark 3.9: More concretely, we may construct

$$V_i := \ker(p_i).$$

Notice V_i T-stable. Since $\gcd(p_1, p_2) = 1$, we have find $a, b \in F[x]$ such that

$$1 = ap_1 + bp_2.$$

Evaluating on T, we find

$$I = a(T)p_1(T) + b(T)p_2(T)$$

and evaluating further on some arbitrary $v \in V$, we find

$$v = \underbrace{a(T)p_1(T)(v)}_{\in V_2} + \underbrace{b(T)p_2(T)(v)}_{\in V_1},$$

i.e. $V_1 \cup V_2 = V$ indeed. Moreover, suppose $v \in V_1 \cap V_2$. Then, $a(T)p_1(T)(v) = 0$ *and* $b(T)p_2(T)(v) = 0$, hence $v = 0$ itself. We conclude $V_1 \cap V_2 = \{0\}$ as desired.

↪**Corollary 3.2**: If F algebraically closed, then V is a direct sum of generalized eigenspaces,

$$V = V_1 \oplus \cdots \oplus V_t,$$

where $V_j = \ker\left(\left(T-\lambda_j\right)^{e_j}\right)$ for some $\lambda_j \in F, e_j \geq 1$.

↪**Definition 3.14** (Generalized Eigenspace): Given $T : V \to V, \lambda \in F$, the eigenspace of T attached to λ is

$$V_\lambda := \{v \in V : Tv = \lambda v\} = \ker(T - \lambda).$$

The *generalized eigenspace* attached to λ is

$$V_{(\lambda)} := \{v \in V : (T-\lambda)^m(v) = 0 \text{ some } m \geq 1\} = \bigcup_{m \geq 1} \ker\big((T-\lambda)^m\big).$$

↪**Theorem 3.12** (Jordan Canonical Form): There is a basis for $V_{(\lambda)}$ for which the matrix of T is of the form

$$\begin{pmatrix} J_{1,\lambda} & 0 & 0 & 0 \\ 0 & J_{2,\lambda} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_{m,\lambda} \end{pmatrix},$$

where

$$J_{1,\lambda} = \begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \ddots & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix},$$

a $d_1 \times d_1$ matrix with diagonals λ and upper-diagonal 1's, and $d_1 + d_2 + \ldots + d_m = \dim\big(V_{(\lambda)}\big)$.

§3.5 Modules over Principal Ideal Domains

↪**Definition 3.15** (Finitely Generated): A module M over a ring R is *finitely generated* if it has a finite spanning set.

↪**Theorem 3.13**: Let M be a finitely generated module over a PID R. Then, there exists elements $a_1|a_2|\cdots|a_t$ and an integer $m \geq 0$ such that

$$M \cong R/(a_1) \oplus R/(a_2) \oplus \cdots \oplus R/(a_t) \oplus R^m.$$

$a_1, \ldots, a_t$ are called the *elementary divisors* of M, and m is called the *rank* of M over R.

Remark 3.10: M free $\Leftrightarrow t = 0$.

Remark 3.11: If G is a finitely-generated abelian group, then

$$G \cong \mathbb{Z}/n_1 \oplus \cdots \oplus \mathbb{Z}/n_t \oplus \mathbb{Z}^m.$$

G finite $\Leftrightarrow m = 0$.

Remark 3.12: If V a generalized eigenspace for T with eigenvalue λ, then

$$V = F[x]/\big((x-\lambda)^{n_1}\big) \oplus \cdots \oplus F[x]/\big((x-\lambda)^{n_t}\big) \oplus F[x]^m,$$

where actually $m = 0$ (since V finite dimensional) and $n_1 \leq n_2 \leq \cdots \leq n_t$.

↪**Theorem 3.14**: If M a finitely generated R-module, then it is a quotient of a free R-module.

Proof. Let $m_1, ..., m_t$ be a system of R-module generators for M. Then consider

$$\varphi : R^t \to M, \qquad (\lambda_1, ..., \lambda_t) \mapsto \lambda_1 m_1 + \cdots + \lambda_t m_t.$$

Since $m_1, ..., m_t$ generate M, this is a surjective module homomorphism. This gives

$$M \cong R^t/(\ker(\varphi)).$$

■

↪**Definition 3.16** (Cyclic R-module): An R-module is said to be *cyclic* if it is isomorphic to R/I for some ideal $I \subset R$. In particular, R is cyclic if it can be generated by a single element, namely $1 + I$.

↪**Proposition 3.7**: If N an R-submodule of a free R-module of rank n, then N is also free, of rank $\leq n$.

Proof. We prove by induction on n. If $n = 1$ and $N \subseteq R$ an R-submodule, then in particular N is an ideal. Since R a PID, there is some $a \in R$ such that $N = (a)$. If a is zero, then N is the zero module so free. Else, consider the map $R \to N, \lambda \mapsto \lambda a$. It is clearly surjective, and its kernel is trivial because a is not a zero divisor.
Suppose the case for n and take $N \subseteq R^{n+1}$. Consider the R-module homomorphism $\varphi : R^{n+1} \to R, (\lambda_1, ..., \lambda_{n+1}) \mapsto \lambda_{n+1}$. $\varphi(N)$ is an ideal of R so $\varphi(N) = (a)$ for some $a \in R$. Let $m_{n+1} \in N$ be such that $\varphi(m_{n+1}) = a$. Consider $N \cap \ker(\varphi)$. The kernel is given by all elements of the form $(\lambda_1, ..., \lambda_n, 0)$ for $\lambda_i \in R$, which is isomorphic to R^n.
If $a = 0$, then $N \subset \ker(\varphi) \cong R^n$, so we may directly apply the inductive hypothesis and find that N is free of rank $\leq n < n + 1$.
Else if $a \neq 0$, then by the induction hypothesis, we know that $N \cap \ker(\varphi)$ is free of rank $\leq n$, being a submodule of R^n. On the other hand, we claim $N \cong (N \cap \ker(\varphi)) \oplus R$. Consider the map

$$\eta : (N \cap \ker(\varphi)) \oplus R \to N, \qquad (n_0, \lambda) \mapsto n_0 + \lambda m_{n+1}.$$

Given $n \in N$, note that $\varphi(n) \in (a)$ hence there is some $\lambda \in R$ such that $\varphi(n) = \lambda a$. Then, taking $n_0 := n - \lambda m_{n+1} \in \ker(\varphi)$, and so

$$\eta(n_0, \lambda) = n_0 + \lambda m_{n+1} = n,$$

so η surjective. For injectivity, suppose $\varphi(n_0, \lambda) = 0 \Rightarrow n_0 + \lambda m_{n+1} = 0$. Applying φ to both sides, we find $\varphi(n_0) + \lambda\varphi(m_{n+1}) = 0$. $n_0 \in \ker(\varphi)$, so we find $\lambda a = 0$, but $a \neq 0$ hence $\lambda = 0$. It follows that $n_0 = 0$, and thus $(n_0, \lambda) = (0, 0)$.
So, we find that $N \cong N_0 \oplus R$, where $N_0 \subseteq R^n$. Applying the inductive hypothesis to N_0, we find N_0 free of rank $m \leq n$, ie $N_0 \cong R^m$, and thus $N \cong R^{m+1}$ of rank $m + 1 \leq n + 1$, completing the proof. ■

We now have a fairly simple representation of our module as R^n/R^m. To ultimately obtain the proof we seek, we wish to simplify the structure of R^m even further. We approach this by considering the image of our module under invertible matrices.

Let $e_1, ..., e_n$ be a basis of R^n. Let $v_1, ..., v_m$ a basis for N, where

$$v_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}, \qquad i = 1, ..., m,$$

and let $v_j = 0$ for $j = m + 1, ..., n$, and so consider the matrix with columns v_i, $A = \left(a_{ij}\right)_{1 \leq i,j \leq n} \in M_n(R)$. Then, our finitely generated R-module $\Omega = R^n/AR^n$. We have the following:

1. If $Q \in \mathrm{GL}_n(R)$, then $R^n/AQ^{-1}R^n = R^n/AR^n$, since Q^{-1} induces an isomorphism on R^n, hence $Q^{-1}R^n = R^n$.
2. If $P \in \mathrm{GL}_n(R)$, then $R^n/PAR^n \cong R^n/AR^n$, by considering the map $R^n/AR^n \to PR^n/PAR^n = R^n/PAR^n, v \mapsto Pv$.

In short, we have an action $\mathrm{GL}_n(R) \times \mathrm{GL}_n(R)$ acting on $X = M_n(R)$ by $(P, Q)(A) = PAQ^{-1}$.

We claim, then, that for any $A \in M_n(R)$, the orbit of A contains a diagonal matrix with entries $d_1|d_2|\cdots|d_n$, where the d_i's may be 0.

We wish to study $\mathrm{GL}_n(R) \setminus M_n(R)/\mathrm{GL}_n(R)$, that is, the orbits of $M_n(R)$ under this action. This is difficult, so we instead consider the restricted orbits $\mathrm{SL}_n(R) \setminus M_n(R)/\mathrm{SL}_n(R)$.

↪**Theorem 3.15**: Let $M = R^n$ and let $N \subseteq M$ a (free) R-submodule. Then, there exists a basis for M $m_1, ..., m_n$ such that N is spanned by $d_1m_1, d_2m_2, ..., d_nm_n$ with $d_1|d_2|\cdots|d_n$.

Remark 3.13: If $v = \lambda_1 m_1 + \cdots + \lambda_n m_n \in M$, then $v \in N \Leftrightarrow d_1|\lambda_1, d_2|\lambda_2, ..., d_n|\lambda_n$. Hence,

$$M/N \cong R^n/\begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix} R^n \cong (R/d_1R) \times (R/d_2R) \times \cdots \times (R/d_nR).$$

Remark 3.14: If $d_1, ..., d_t$ are non-zero, then $d_1m_1, ..., d_tm_t$ are linearly independent; however, in general, there may be some d_i's equal to zero.

Proof. We prove by induction on n. If $n = 1$, $M = R$ and $N \subseteq R$. Let $m_1 = 1$. N an ideal of R, so $N = (d_1) = d_1R$ for some $d_1 \in R$, then N spanned by d_1m_1.

Suppose the claim for n. Let $M = R^{n+1}$. Given $\varphi \in \mathrm{Hom}(M, R)$, $I(\varphi) \coloneqq \mathrm{im}(\varphi|_N) = \{\varphi(n) : n \in N\} \subseteq R$. This is an R-submodule of R, hence an ideal, so we may write $I(\varphi) = \left(d_\varphi\right)$. Let

$$\Sigma \coloneqq \{I(\varphi) : \varphi \in \mathrm{Hom}(M, R)\}.$$

Σ is partially ordered by inclusion (or equivalently divisibility of generators), and so satisfies the maximal chain property. Let $(d_1) = I(\varphi_1)$ be a maximal element of Σ, i.e. d_1 minimal for divisibility. Let $n_1 \in N$ be such that $\varphi_1(n_1) = d_1$. We claim that n_1 is divisible d_1, i.e. there is an $m_1 \in M$ such that $n_1 = d_1m_1$.

Let $\eta_1, ..., \eta_{n+1}$ be the natural projections $M \to R$, $\eta_j(\lambda_1, ..., \lambda_{n+1}) \coloneqq \lambda_j$. So, if $n_1 = (x_1, ..., x_{n+1})$, we need to show that $d_1|x_j = \eta_j(n_1)$ for each $j = 1, ..., n + 1$. Let $d = \eta_j(n_1)$. Let $d_0 = \gcd(d_1, d)$, which we can write $d_0 = rd_1 + sd$ for some $r, s \in R$. We have, unpacking definitions,

$$\begin{aligned} d_0 &= r\varphi_1(n_1) + s\eta_j(n_1) \\ &= \left(r\varphi_1 + s\eta_j\right)(n_1). \end{aligned}$$

The map $r\varphi_1 + s\eta_j \in \mathrm{Hom}(M, R)$, hence $(d_0) \in \Sigma$. We have too that $d_0|d_1$, and by maximality of d_1, $d_1|d_0$ hence it must be $(d_0) = (d_1)$. In addition, $d_0|d$ and thus $d_1|d$. This holds for all $d = \eta_j(n_1)$'s, hence it follows that $d_1|n_1$.

Let m_1 then be such that $d_1m_1 = n_1$. Recall $\varphi_1(n_1) = d_1$. Then,

$$\begin{aligned} \varphi_1(d_1m_1) = \varphi_1(n_1) = d_1 \\ \Rightarrow d_1\varphi_1(m_1) = d_1 \Rightarrow \varphi_1(m_1) = 1. \end{aligned}$$

We claim, then, $M \cong Rm_1 \oplus \ker(\varphi_1)$. Consider the map

$$M \to Rm_1 \oplus \ker(\varphi_1), \qquad m \mapsto (\varphi_1(m)m_1, m - \varphi_1(m)m_1),$$

noticing that

$$\varphi_1(m - \varphi_1(m)m_1) = \varphi_1(m) - \varphi_1(\varphi_1(m)m_1) = \varphi_1(m) - \varphi_1(m)\underbrace{(\varphi_1(m_1))}_{=1} = 0.$$

Let $M_2 \coloneqq \ker(\varphi_1)$, noting then $M_2 \cong R^n$. We can write then $N = Rn_1 \oplus (\ker(\varphi_1) \cap N)$; given $n \in N$, we have, recalling $\varphi_1(N) = (d_1)$,

$$n = \left(\underbrace{\frac{\varphi_1(n)}{d_1}}_{\text{since } \varphi_1(n) \in (d_1)} n_1, n - \frac{\varphi_1(n)}{d_1} n_1 \right).$$

Let $N_2 := (\ker(\varphi_1) \cap N)$. N_2 a submodule of $M_2 \cong R^n$. By the induction hypothesis, there is a basis $m_2, \ldots, m_{n+1}$ for M_2 and $d_2, \ldots, d_{n+1} \in R$ with $d_2 | d_3 | \cdots | d_{n+1}$ such that $n_2 := d_2 m_2, \ldots, n_{n+1} := d_{n+1} m_{n+1}$ spans N_2. Then, $m_1, \ldots, m_{n+1}$ a basis for all of M, and $d_1 m_1, \ldots, d_{n+1} m_{n+1}$ spans N, so it remains to show that $d_1 | d_2$.

Consider η_j be the jth coordinate homomorphism relative to our basis $m_1, \ldots, m_{n+1}$, i.e. if $m = \lambda_1 m_1 + \cdots + \lambda_{n+1} m_{n+1}$, $\eta_j(m) = \lambda_j$. Then, remark since $n_1 = m_1 d_1$

$$\eta_1(n_1) = d_1, \qquad \eta_2(n_1) = 0$$

and since $n_2 \in M_2$,

$$\eta_1(n_2) = 0, \qquad \eta_2(n_2) = d_2.$$

Let $d_0 = \gcd(d_1, d_2) = r d_1 + s d_2$ for some $r, s \in R$. Let $\eta = r\eta_1 + s\eta_2 \in \mathrm{Hom}(M, R)$. Then,

$$\eta(n_1 + n_2) = d_0.$$

Hence $(d_0) \in \Sigma$, since $n_1 + n_2 \in N$. By maximiality of d_1, $d_1 | d_0$, but also $d_0 | d_2$ hence $d_1 | d_2$, as we needed to show. ■

↪**Corollary 3.3**: $M/N = R^n/(d_1 R \oplus \cdots \oplus d_n R) = (R/d_1 R) \oplus \cdots \oplus (R/d_n R)$.

⊛ **Example 3.5**: Let A be a finitely generated abelian group; A then just a fg $\mathbb{Z}$-module. Then, $A \cong \mathbb{Z}/d_1 \oplus \cdots \oplus \mathbb{Z}/d_t \oplus \mathbb{Z}^m$. In particular, if A is finite, then $m = 0$. Then, $d_1 \cdots d_t = \#A$, and $d_t a = 0$ for any $a \in A$, and indeed the smallest such integer with this property (called the *exponent* of A).

Note that A is *not* characterized by its exponent and cardinality; if two groups have the same exponent and cardinality, they need not be isomorphic. For instance, consider

$$\mathbb{Z}/3 \times \mathbb{Z}/3 \times \mathbb{Z}/9, \qquad \mathbb{Z}/9 \times \mathbb{Z}/9,$$

which are not isomorphic but have the same cardinality, 81, and exponent, 9.

⊛ **Example 3.6**: Let $R = \mathbb{F}[x]$ for some field $\mathbb{F}$, and

$$M = \mathbb{F}[x]/(d_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(d_t(x)) \oplus \mathbb{F}[x]^r,$$

where $r = 0 \Leftrightarrow \dim_{\mathbb{F}}(M) < \infty$ as a vector space. In particular,

$$\dim_{\mathbb{F}}(M) = \deg d_1 + \cdots + \deg d_t.$$

In particular, we have a correspondence

$$\begin{aligned} \text{fg } \mathbb{F}[x] \text{ module}, r = 0 &\longleftrightarrow (V, T), \dim_{\mathbb{F}}(V) < \infty \text{ and } T : V \to V \\ M = V, f(x) \cdot v := f(T)(v) &\rightsquigarrow (V, T) \\ V \rightsquigarrow V, T(v) &:= x \cdot v. \end{aligned}$$

The maximal divisor $d_t(x)$ is the minimal polynomial of T.

↪**Proposition 3.8**: $d_1(x) \cdots d_t(x)$ the characteristic polynomial of T.

Proof. We let p_T the minimal polynomial, f_T the characteristic polynomial. Recall that $p_T | f_T$, $\deg f_T = \dim V$, and if $V = V_1 \oplus V_2$ as a $\mathbb{F}[x]$-module (namely, V_1, V_2 respectively stable under the action of $\mathbb{F}[x]$), then $f_T = f_T|_{V_1} \cdot f_T|_{V_2}$.

If $t = 1$, $V = \mathbb{F}[x]/(p(x))$, with $p_T(x) = p(x)$. Since $p_T | f_T$ and $\deg f_T = \dim_{\mathbb{F}} V = \deg p_T$ so $p_T = f_T$.

For general t, $V = \mathbb{F}[x]/(p_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(p_t(x)) =: V_1 \oplus \cdots \oplus V_t$, then $f_T(x)$ is just $f_T|_{V_1} \cdots f_T|_{V_t}$, and since $f_T|_{V_i} = p_i$ as per the $t = 1$ case, the proof follows. ■

We summarize our interpretations of the structures that arise from the theorem:

	$\mathbb{Z}$	$\mathbb{F}[x]$
fg modules	fg abelian groups	fg $F[x]$-modules
$r = 0$	finite abelian groups	finite dimensional (V, T)
d_t	exponent	minimal polynomial
$d_1 \cdots d_t$	cardinality	characteristic polynomial

↪**Proposition 3.9**: Given two matrices $M_1, M_2 \in M_n(\mathbb{F})$, M_1 is conjugate to M_2 if and only if the associated $\mathbb{F}[x]$-modules have the same elementary divisors.

Proof. With $M_1, M_2 : V \to V$, V can be viewed as an $\mathbb{F}[x]$-module in two different ways. We denote these modules V_i, where $f(x)v = f(M_i)(v)$ for $i = 1, 2$. Suppose that V_1, V_2 have the same elementary divisors, $d_1, \ldots, d_t$. This implies

$$V_1 \cong \mathbb{F}[x]/(d_1) \oplus \cdots \oplus \mathbb{F}[x]/(d_t) \cong V_2,$$

as $\mathbb{F}[x]$*-modules*. Hence, there exists some isomorphism, $j : V_1 \to V_2$ of $\mathbb{F}[x]$-modules, so in particular

$$j(xv) = x \cdot j(v)$$

for all $v \in V_1$, but recalling that the action of $\mathbb{F}[x]$ is, we have

$$j(M_1 v) = M_2 j(v),$$

hence,

$$j \circ M_1 = M_2 \circ j,$$

and so M_1, M_2 indeed conjugate.

More concretely, j restricts to an isomorphism of $\mathbb{F}$-vector spaces, which as we know can simply be realized as multiplication by an invertible matrix, hence $JM_1 = M_2 J$ where J the matrix realization of such an isomorphism. ■

⊛ **Example 3.7**: Let $G = \mathrm{GL}_3(\mathbb{F}_2)$. Recall $\#G = 168$. Conjugacy classes in G are, by the previous theorem, in bijection with possible sequences of elementary divisors $(d_1, ..., d_t)$.

t=1: $d_1(x)$ a polynomial of degree three with coefficients in $\mathbb{F}_2$. Since T must be invertible, $x \nmid d_1$, so we have

$$d_1(x) = x^3 + ?\,x^2 + ?\,x + 1$$

where $? \in \{0, 1\}$. We go through all possibilities:

	$d_1(x)$	C
(a)	$x^3 + 1 = (x+1)\left(x^2 + x + 1\right)$	3A
(b)	$x^3 + x + 1$	7A
(c)	$x^3 + x^2 + 1$	7B
(d)	$x^3 + x^2 + x + 1$	4A

- By the PDT, T with minimal polynomial (a) splits $V = V_1 \oplus V_2$ where T acts on V_1 as the identity. Let $T_2 = T|_{V_2}$. Then, since $x^2 + x + 1$ the minimal polynomial of T_2, it must be that $T_2^3 = 1$. Hence, T is an element of order 3.
- For (b), (c), we've seen that these polynomials are irreducible over $\mathbb{F}_2$. In particular, if T has such a minimal polynomial, then T of order 7.
- For (d), $x^3 + x^2 + x + 1 = (x+1)\left(x^2+1\right) = (x+1)^3$. Then, $T^3 = T^2 + T + 1$ so $T^4 = T^3 + T^2 + T = T^2 + T + T^2 + T + 1 = 1$, hence T of order 4.

$t = 2$: consider (d_1, d_2). It must be that $d_1(x) = x + 1$, and so $d_2(x) = (x+1)^2$ is the only possibility. Then, $\mathbb{F}_2^3 = \mathbb{F}_2[x]/(x+1) \oplus \mathbb{F}_2[x]/\left((x+1)^2\right) = \mathbb{F}_2 \oplus \mathbb{F}_2[\varepsilon]/\left(\varepsilon^2\right)$. Then, $T \leftrightarrow (1, 1+\varepsilon)$, $T^2 \leftrightarrow \left(1, 1 + 2\varepsilon + \varepsilon^2\right) = (1, 1)$, so T of order 2. We have

	$(d_1(x), d_2(x))$	C
(e)	$\left(x+1, (x+1)^2\right)$	2A

$t = 3$: it must be $(d_1, d_2, d_3) = (x+1, x+1, x+1)$. Such a transformation must be the identity.

	$(d_1(x), d_2(x), d_3(x))$	C
(f)	$(x+1, x+1, x+1)$	1A

↪**Theorem 3.16**: If $t = 1$, then $Z_G(A) = (\mathbb{F}[x]/(d_1(x)))^\times$.

Proof. We know $V \cong \mathbb{F}[x]/(d_1(x))$ as an $\mathbb{F}[x]$-module. Then, $Z_G(A) = \mathrm{Aut}_{\mathbb{F}[x]}(V) = \mathrm{Aut}_{\mathbb{F}[x]}(\mathbb{F}[x]/(d_1(x))) = (\mathbb{F}[x]/(d_1))^\times$.

We have that $M \in Z_G(A) \Leftrightarrow MA = AM \Leftrightarrow MAv = AMv$ for every $v \in V$. Thinking of these as $\mathbb{F}[x]$-modules, this is equivalent to aking $M(xv) = xMv$, or equivalently M an automorphism of V as an $\mathbb{F}[x]$-module, i.e. respects $\mathbb{F}[x]$ scalar multiplication. ■

⊛ **Example 3.8**: We compute the size of the conjugacy classes of $\mathrm{GL}_3(\mathbb{F}_2)$ found above.

3A: By the previous theorem,

$$Z_G(3\mathrm{A}) = \left(\mathbb{F}_2[x]/\left(x^3+1\right)\right)^\times = \left(\mathbb{F}_2 \times \mathbb{F}_2[x]/\left(x^2+x+1\right)\right)^\times = \mathbb{F}_2^\times \times \mathbb{F}_4^\times = 1 \times \mathbb{Z}/3\mathbb{Z}.$$

7A: Similarly,

$$\begin{aligned} Z_G(7\mathrm{A}) &= \left(\mathbb{F}_2[x]/\left(x^3+x+1\right)\right)^\times \\ &= \left(\mathbb{F}_8\right)^\times = \mathbb{Z}/7\mathbb{Z}. \end{aligned}$$

7B:

$$Z_G(7\mathrm{B}) = \left(\mathbb{F}_2[x]/\left(x^3+x^2+1\right)\right)^\times = \left(\mathbb{F}_8\right)^\times = \mathbb{Z}/7\mathbb{Z}.$$

4A:

$$Z_G(4\mathrm{A}) = \left(\mathbb{F}_2[x]/\left((x+1)^3\right)\right)^\times = \left(\mathbb{F}_2[\varepsilon]/\left(\varepsilon^3\right)\right)^\times.$$

The ring $\mathbb{F}_2[\varepsilon]/\left(\varepsilon^3\right) = \left\{a + b\varepsilon + c\varepsilon^2 : a, b, c \in \mathbb{F}_2\right\}$. An element is invertible if and only if $a = 1$, so $\left(\mathbb{F}_2[\varepsilon]/\left(\varepsilon^3\right)\right)^\times = \left\{1 + b\varepsilon + c\varepsilon^2 : b, c \in \mathbb{F}_2\right\}$, with

$$\left(1 + b\varepsilon + c\varepsilon^2\right)^{-1} = 1 + \left(b\varepsilon + c\varepsilon^2\right) + \left(b\varepsilon + c\varepsilon^2\right)^2.$$

So, we have that $\#Z_G(4\mathrm{A}) = 4$, namely only the powers of the element itself.

§4 Midterm Review

§4.1 A_5 has no normal subgroups

↪**Proposition 4.1**: A_5, the group of even permutations on 5 letters, contains no normal subgroups.

Normal subgroups are always unions of conjugacy classes, so we begin by analyzing these. Remark that for any $x \in A_5$, the conjugacy class $C_A(x)$ of $x \in A_5$ is a subset of $C_S(x)$ of $x \in S_5$. However, we cannot simply assume they are the same, since while two elements may be conjugate in S_5, the element needed to conjugate between them may not be in A_5.

Let $x \in A_5$. Then, by the orbit-stabilizer theorem,

$$\#C_A(x) = \frac{\#A_5}{\#\mathrm{Stab}_A(x)}$$

since A_5 acts on $C_A(x)$ transitively by conjugation. Similarly,

$$\#C_S(x) = \frac{\#S_5}{\#\mathrm{Stab}_S(x)}.$$

Note that $\mathrm{Stab}_S(x) \supseteq \mathrm{Stab}_A(x)$ a subgroup, hence $\#\mathrm{Stab}_S(x) = k \cdot \mathrm{Stab}_A(x)$ for some $k \in \mathbb{N}$. Moreover, since $\#S_5 = 2 \cdot \#A_5$, we may combine the expressions above and find

$$\#C_S(x) = \frac{2}{k}\#C_A(x) \Rightarrow k = 1, 2.$$

So, in particular, $\#C_A(x)$ is either equal to or half of $\#C_S(x)$. Since we know $C_A(x) \subset C_S(x)$, then if the two are of the same size they are therefore equal.

We can now specialize to particular elements in A_5.

- $(ab)(cd)$: there are 15 such elements in A_5, hence $\#C_S((ab)) = 15$. This isn't divisble by 2, hence it must be that $C_S((ab)) = C_A((ab))$. (We can also see this by noting that (ab) stabilizes $(ab)(cd)$ but isn't contained in A_5, so the formula above gives the same result).
- (abc): there are 20 3-cycles in A_5, so we need to do a little more work here. Notice that by our work above

  $$C_S(x) = C_A(x) \Leftrightarrow \#\mathrm{Stab}_S(x) = 2 \cdot \#\mathrm{Stab}_A(x) \Leftrightarrow \mathrm{Stab}_A(x) \subsetneq \mathrm{Stab}_S(x),$$

  so to show the conjugacy classes are equal, it suffices to show that the stabilizers *aren't* equal. Remark that, for instance, $(12) \in \mathrm{Stab}_S((345))$, but $(12) \notin A_5$ so certainly $(12) \notin \mathrm{Stab}_A((345))$. It follows that the two subgroups are not equal, and thus $C_A((345)) = C_S((345)) = C_S((abc))$.
- $(abcde)$: there are 24 such elements in A_5; but remark that $24 \nmid \#A_5 = 60$, hence it can't be that A_5 acts transitively on the set of 5-cycles. It follows then by our work above that there must be precisely two distinct conjugacy classes, each of size 12, of 5-cycles in A_5, which we can represent, for instance, by $C_A((12345)), C_A((12354))$.

We can see this more explicitly another way. Put $\sigma := (12345)$ and consider again $\mathrm{Stab}_S(\sigma)$. Clearly, $\sigma^t \in \mathrm{Stab}_S(\sigma)$ for $t = 0, \ldots, 4$, so $\mathrm{Stab}_S(\sigma) \geq 5$; remark that each of these elements in A_5 as well. Suppose $g \in \mathrm{Stab}_S(\sigma)$. Then, for every $k \in \{1, \ldots, 5\}$,

$$\begin{aligned} \sigma = g^{-1}\sigma g &\Leftrightarrow \sigma(k) = g^{-1}\sigma g(k) \\ &\Leftrightarrow g(k+1) = g(k) + 1, \end{aligned}$$

since σ just "shifts" elements (mod 5). In particular, such g's are uniquely determined by their effect on a single element, since then we can apply this recursive relation to find its affects on the others. So, we have 5 choices for, say, $g(1)$, and so in particular $\#\mathrm{Stab}_S(\sigma) \leq 5$, and thus equals 5. Hence, since every element in the stabilizer also in A_5, we conclude that $\mathrm{Stab}_A(\sigma) = \mathrm{Stab}_S(\sigma)$, and thus $\#C_A(\sigma) = \frac{1}{2}\#C_S(\sigma)$.

In summary, we have the following table:

Conj. Class	#
()	1
(12)(34)	15
(123)	20
(12345)	12
(12354)	12
	60

We can now use the fact that normal subgroups are always unions of conjugacy classes and that the order of a subgroup always divides the order of the group to conclude that A_5 has no normal subgroups. Indeed, the possible orders of subgroups of A_5 would be the divisors of 60, namely,

$$1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30,$$

none of which cannot be achieved by adding cardinalities of conjugacy classes.

§4.2 Sylow 2-subgroups of S_{n-1}, S_n

↪**Proposition 4.2**: Let n odd. Then, S_{n-1} and S_n have the same Sylow 2-subgroup, and the number of Sylow 2-subgroups in S_n is precisely n times that in S_{n-1}.

We have the natural inclusion $S_{n-1} \subset S_n$ by fixing an element, hence any Sylow 2-subgroups of S_{n-1} are necessarily contained in S_n. Moreover, we have that

$$\frac{\#S_n}{\#S_{n-1}} = \frac{n!}{(n-1)!} = n,$$

by assumption odd, hence the powers of two in $n!, (n-1)!$ are the same, and so the Sylow 2-subgroups of the two must be the same as well.

To show the second claim, let

$$X_n := \{\text{Sylow 2-subgroups of } S_n\}, \quad X_{n-1} := \{\text{Sylow 2-subgroups of } S_{n-1}\}.$$

Fix some $P \in X_{n-1}$, noting that $P \in X_n$ as well. Then, we have that

$$\#X_n = \frac{\#S_n}{\#\mathrm{Stab}_{S_n}(P)}, \qquad \#X_{n-1} = \frac{\#S_{n-1}}{\#\mathrm{Stab}_{S_{n-1}}(P)},$$

Since X_n, X_{n-1} are transitive S_n, S_{n-1} sets respectively. Clearly, $\mathrm{Stab}_{S_{n-1}}(P) \subset \mathrm{Stab}_{S_n}(P)$, so $\#\mathrm{Stab}_{S_n}(P) = k \cdot \#\mathrm{Stab}_{S_{n-1}}(P)$ for some $k \in \mathbb{N}$. This implies then that

$$\#X_n = \frac{n \cdot \#S_{n-1}}{k \cdot \#\mathrm{Stab}_{S_{n-1}}(P)} = \frac{n}{k}\#X_{n-1},$$

so in particular, $\#X_n \leq n \cdot \#X_{n-1}$. I claim that $k = 1$, namely that $\mathrm{Stab}_{S_n}(P) = \mathrm{Stab}_{S_{n-1}}(P)$. Clearly, we have $\mathrm{Stab}_{S_n}(P) \supseteq \mathrm{Stab}_{S_{n-1}}(P)$. Suppose there existed some $\sigma \in \mathrm{Stab}_{S_n}(P) - \mathrm{Stab}_{S_{n-1}}(P)$; namely, then, $\sigma \in S_n - S_{n-1}$ i.e. σ doesn't fix n. Let $p \in P$, then remark that

$$\sigma^{-1}p\sigma(n) = n \Leftrightarrow p\sigma(n) = \sigma(n) \Leftrightarrow p \text{ fixes } \sigma(n)$$

$\sigma(n) \neq n$ by assumption, so this means that p fixes some non-n element. I claim this is impossible. I claim that P acts upon $\{1, \ldots, n-1\}$ without fixed points. Suppose towards a contradiction that there exists some $x \in \{1, \ldots, n-1\}$ (wlog, $x = n-1$) such that $px = x$ for every $p \in P$. Then, this implies we can embed $P \subset S_{n-2}$. However, $\#P = 2^t$ and $\#S_{n-2} = \frac{\#S_n}{n(n-1)} = 2^t\frac{m}{n(n-1)}$ so $\#P \nmid \#S_{n-2}$ so this is impossible. Hence, p acts upon $\{1, \ldots, n-1\}$ without fixed points, and thus such a σ cannot exist. We conclude $\mathrm{Stab}_{S_n}(P) = \mathrm{Stab}_{S_{n-1}}(P)$ indeed. The proof follows.

§4.3 Midterm Questions

↪**Proposition 4.3**: Describe two non-abelian groups of cardinality 8 and show that they are not isomorphic.

Proof. Consider D_8 (symmetry group of the square) and $\mathbb{H}$ (the quaternions). Argue on the order of elements. ■

↪**Proposition 4.4**: Write down the class equation for the symmetric group S_4 on 4 elements and use this to give a complete list of the normal subgroups of S_4.

Proof.

$$S_4 = \{1\} \sqcup \{\text{transpositions}\} \sqcup \{\text{3-cycles}\} \sqcup \{2, \text{2-cycles}\} \sqcup \{\text{4-cycles}\}$$

$$\#S_4 = 1 + \binom{4}{2} + 2 \cdot \binom{4}{3} + 3 + 3!$$

A subgroup is normal if it a union of conjugacy classes, so it suffices to check the possible unions of conjugacy classes. Should give $\{1\}, S_4, A_4, K_4$. ■

↪**Proposition 4.5**: Give a formula for the number of distinct ways of coloring the 8 corners (i.e. vertices) of a cube with t distinct colors. (Note that the class equation computed in Question 2 can and should be used to assist you with this question.)

Proof. ■

↪**Proposition 4.6**: Let p be a prime number. State the Sylow theorem for p. Starting with the fact (which was proven in class, and which you may assume for this question) that every finite group of cardinality not a power of p can be made to act transitively on a set whose cardinality is neither 1 nor divisible by p, show that eery finite group contains a Sylow p-subgroup.

↪**Proposition 4.7**: Show that S_5 can be made to act transitively on a set X of size 6, and describe how the elements of order 3 and 6 in S_5 act on X. (I.e. describe their cycle shapes.)

§5 Final Review

↪**Proposition 5.1**: Let T be a linear transformation over a field F having $(x-\lambda)^2$ as minimal polynomial, for some $\lambda \in F$, and let $g(x)$ be a polynomial in $F[x]$. Show that

$$g(T) = g(\lambda)I + g'(\lambda)(T - \lambda I)$$

where I is the identity transformation. Can you generalize this formula to the case where the minimal polynomial is $(x-\lambda)^k$?

Proof. Since $p_T(x) = (x-\lambda)^2$, it follows that

$$0 = T^2 - 2\lambda T + \lambda^2.$$

If $g(x) = g_m x^m + \cdots + g_1 x + g_0 \in \mathbb{F}[x]$, then

$$g(\lambda)I + g'(\lambda)(T - \lambda I) = (g_m\lambda^m + \cdots + g_0) + \left(mg_m\lambda^{m-1} + \cdots + g_1\right)(T - \lambda I).$$

Let us compare coefficients of g. ■