Conjugacy Classes of Centralizers in Groups

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Bhargavi Parthasarathy

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2019

Supervisor: Dr. Anupam Kumar Singh

Certificate

This is to certify that this dissertation entitled Conjugacy Classes of Centralizers in Groups towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Bhargavi Parthasarathy at Indian Institute of Science Education and Research under the supervision of Dr. Anupam Kumar Singh, Associate Professor, Department of Mathematics, during the academic year 2018-2019.

Dr. Anupam Kumar Singh

Committee:

Dr. Anupam Kumar Singh

Dr. Vivek Mohan Mallick

This thesis is dedicated to Amma, Appa and Ambreesh.
Your existence makes life worthwhile.

Declaration

I hereby declare that the matter embodied in the report entitled Conjugacy Classes of Centralizers in Groups are the results of the work carried out by me at the Department of Mathematics, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Anupam Kumar Singh and the same has not been submitted elsewhere for any other degree.

Bhargavi

Bhargavi Parthasarathy

Acknowledgments

First and foremost, I would like express my heartfelt gratitude to my supervisor, Dr. Anupam Kumar Singh, for being the mentor one could only wish for. It is under his able guidance and with his constant support that I was able to navigate my way through this project and bring it to fruition. He often had more belief in my capabilties that I did, whether it be by boosting my confidence when I'd be stuck or giving me the independence to make my mistakes and learn from them. At the end of a year of honest discussions, I can say I am more confident about my mathematical abilities than when I began this endeavour, and for that I will always be grateful.

I would also like to thank the expert on my Thesis Advisory Committee, Dr. Vivek Mohan Mallick for the valuable insight I gained through our discussions and whose patient and constant encouragment always bolstered me.

I am immensely thankful to Saikat for engaging me in insightful conversations and putting up with me and Uday for providing me with important feedback.

I would finally like to thank my friends and family, who put up with me through this ardous journey. I cannot imagine completing this without their unquestioning love and inspiring support. They've always been there, whether I needed them or not, standing in support and solidarity and I'm forever indebted.

Abstract

Two elements in a group G are said to be in the same z-class or z-equivalent if their centralizers are conjugate in G. This is an equivalence relation on G and provides a partition of G into disjoint equivalence classes. The structure of centralizers and their conjugacy classes provides important insight into the group structure. Although z-equivalence is a weaker relation than conjugacy, it is interesting to note that there are infinite groups which have infinitely many conjugacy classes but fintely many z-classes. In fact, the finiteness of z-classes in algebraic groups and Lie groups is an interesting problem. We have studied the structure of z-classes for symmetric groups S_n, general linear groups $GL_n(\mathbb{F})$ and general affine groups $GA_n(\mathbb{F})$ and have proven that there are finitely many z-classes, for $n \geq 5$ in S_n and when $\mathbb{F}$ has finitely many extensions, in the latter cases. We also investigate the idea that there is a relation between the finiteness of z-classes and and the intuitive understanding of the finiteness of "dynamical types" of transformations in geometry through group actions.

Contents

Introduction

Let G be a group. Two elements $x, y \in G$ are said to be z-equivalent if their centralizers $\mathcal{Z}_G(x), \mathcal{Z}_G(y)$ are conjugate in G. The problem of characterizing z-classes and explicitly counting them has been studied for various groups. Since they require deep understanding of the structures of conjugacy classes in addition to the structure of centralizers, these are important problems to study, giving us great insight into the group structure. We further note that the non-triviality of these strutures is what makes these problems interesting, and in that spirit, we persevere to understand the conditions under which such structures occur. The existence of instances where an infinite group has finitely many z-classes has been used by Kulkarni to provide the rudimentary notion of "dynamical types" and their relation to z-classes [SK07]. Expanding in this direction, the z-classes of real hyperbolic isometries have been classified and counted by Gongopadhyay and Kulkarni [GK09a].
Another perspective on this problem was provided through Steinberg's theorem [Ste74] that proved the finiteness of z-classes in reductive algebraic groups over a field of good characterstic, allowing one to explore the assumptions on a field $\mathbb{F}$ under which a group defined over this field has finitely many z-classes. This has motivated further exploration of ideas of finiteness of z-classes in various other groups.
Bhunia, Kaur and Singh have determined the number of z-classes in symmetric groups and alternating groups and their relation to restricted partitions in [BKS17]. In the case of general linear maps and affine maps, Kulkarni [Kul07] has shown the conditions the underlying field must comply to for these groups to have finitely many z-classes. Using the fact that the number of z-classes is invariant for a family of isoclinic groups, Kulkarni, Kitture and Jadhav [KKJ16] have obtained bounds for the number of z-classes in certain families of groups, following which Dattatreya and Jadhav [DJ14] have determined the number of z-classes in p-groups of order $\leq p^5$. Gongopadhyay and Kulkarni [GK09b] have parametrized z-classes in the group of isometries and determined that they are finite when the undelying field is perfect and has finitely many extensions. [Gon13] provides a unified approach to the deter-

mination of the conjugacy classes of centralizers in hyperbolic geometries. Gouraige [Gou06] studied z-classes in central simple algebras, Singh [Sin08] in G_2, Bhunia [BS19] in unitary groups and Bhunia and Singh [Bhu19] in upper triangular matrices .

Chapter 1

Definitions and Examples

Let G be a group. We begin by defining when two elements in G are in the same z-class and show that z-equivalence is a weaker relation than conjugacy. We will also look at some examples in this chapter.

1.1 Definitions

Given a group G, two elements $g_1, g_2 \in G$ are *conjugate* in G, if

$$\exists\ t \in G \text{ such that } tg_1t^{-1} = g_2.$$

Proposition 1.1.1. *Conjugacy is an equivalence relation.*

Proof. It is symmetric due to the existence of inverses in G, reflexive as the identity belongs to G and closure of G ensures transitivity. $\square$

For an element $g \in G$, *centralizer* of g in G, denoted as $\mathcal{Z}_G(g)$ is all those elements in G that commute with g.

$$\mathcal{Z}_G(g) := \{x \in G \mid xg = gx\}.$$

For a group G, two elements g_1, g_2 are *z-equivalent* if their centralizers are conjugate subgroup within G, i.e., if

$$\exists\ t \in G \text{ such that } t\mathcal{Z}_G(g_1)t^{-1} = \mathcal{Z}_G(g_2).$$

z-equivalence is an equivalence relation on G and represents the conjugacy classes of centralizer subgroups of G. These equivalence classes are called *z-classes* or *centralizer classes*. The proof for equivalence is similar to the proof for Proposition 1.1.1.

Proposition 1.1.2. *If two elements are conjugates, then they are z-equivalent.*

Proof. Given $\exists\ t \in G$ such that $g_2 = tg_1t^{-1}$, we need to show that $\exists\ k \in G$ such that

$$\mathcal{Z}_G(g_2) = k\mathcal{Z}_G(g_1)k^{-1}$$

Let $x \in \mathcal{Z}_G(g_2) \Rightarrow xg_2 = g_2x$

$$\begin{aligned}
&\Rightarrow xtg_1t^{-1} = tg_1t^{-1}x \\
&\Rightarrow t^{-1}xtg_1 = g_1t^{-1}xt \\
&\Rightarrow t^{-1}xt \in \mathcal{Z}_G(g_1) \\
&\Rightarrow x \in t\mathcal{Z}_G(g_1)t^{-1} \\
&\Rightarrow \mathcal{Z}_G(g_2) \subseteq t\mathcal{Z}_G(g_1)t^{-1}
\end{aligned}$$

Let $y \in \mathcal{Z}_G(g_1) \Rightarrow g_1y = yg_1$

$$\begin{aligned}
g_2(tyt^{-1}) = tg_1t^{-1}tyt^{-1} = tyg_1t^{-1} = tyt^{-1}(tg_1t^{-1}) = tyt^{-1}g_2 \\
\Rightarrow tyt^{-1} \in \mathcal{Z}_G(g_2) \\
\Rightarrow t\mathcal{Z}_G(g_1)t^{-1} \subseteq \mathcal{Z}_G(g_2)
\end{aligned}$$

$\therefore \mathcal{Z}_G(g_2) = t\mathcal{Z}_G(g_1)t^{-1}$ $\square$

However, the converse need not true. In fact, z-equivalence is a weaker relation than conjugacy.

1.2 Examples

Example 1.2.1. *Abelian group $A = \{a_1, \ldots, a_n\}$*
There are n conjugacy clases in A:

$$\{a_1\}, \ldots, \{a_n\}$$

$$\mathcal{Z}_A(a_i) = A\ \forall a_i \in A$$

There is one z-class in A:

$$\{A\}$$

This example shows us that z-equivalence is a weaker relation. Further, since Abelian groups have only one z-class, we focus on studying the z-classes of non-Abelian groups.

Example 1.2.2. $S_3 = \{1, (12), (13), (23), (123), (132)\}$
There are three conjugacy classes in S_3:

$$\{1\}, \{(12), (13), (23)\}, \{(123), (132)\}$$

$$\mathcal{Z}_{S_3}((12)) = \{1, (12)\}$$
$$\mathcal{Z}_{S_3}((123)) = \{1, (123), (132)\}$$

There are three z-classes in S_3:

$$\{1\}, \{(12), (13), (23)\}, \{(123), (132)\}$$

Remark. *For the symmetric groups S_3 and S_4, the z-classes are same in number as the number of conjugacy classes. However this is not true for $n \geq 5$. For $n = 5$, the elements $(12)(345)$ and (345) are z-equivalent but they are not conjugates. We will elaborate on this in Chapter 2.*

Example 1.2.3. $Q_8 = \{\pm 1, \pm i, \pm j, \pm k \mid i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j\}$
There are five conjugacy classes in Q_8:

$$\{1\}, \{-1\}, \{\pm i\}, \{\pm j\}, \{\pm k\}$$

$$\mathcal{Z}_{Q_8}(-1) = Q_8$$
$$\mathcal{Z}_{Q_8}(\pm i) = \{\pm 1, \pm i\}$$
$$\mathcal{Z}_{Q_8}(\pm j) = \{\pm 1, \pm j\}$$
$$\mathcal{Z}_{Q_8}(\pm k) = \{\pm 1, \pm k\}$$

There are four z-classes in Q_8:

$$\{\pm 1\}, \{\pm i\}, \{\pm j\}, \{\pm k\}$$

Example 1.2.4. $D_{2n} = \langle r, s | r^n = s^2 = 1, rs = sr^{-1} \rangle$

The number of conjugacy classes and z-classes is dependent on whether n is odd or even. The conjugacy classes in D_{2n} are as follows:

1. *If n is odd, there are $\frac{n+3}{2}$ conjugacy classes.*
   - *the identity element: $\{1\}$,*
   - *$\frac{n-1}{2}$ conjugacy classes of size 2: $\{r^{\pm i}\}$ for $i = 1, \ldots, \frac{n-1}{2}$,*
   - *all the reflections : $\{sr^i \mid 0 \leq i \leq n-1\}$.*
2. *If n is even, there are $\frac{n+6}{2}$ conjugacy classes.*
   - *two conjugacy classes of size 1: $\{1\}, \{r^{\frac{n}{2}}\}$,*
   - *$\frac{n}{2} - 1$ conjugacy classes of size 2: $\{r^{\pm i}\}$ for $i = 1, \ldots, \frac{n}{2} - 1$,*
   - *the reflections fall into two conjugacy classes: $\{sr^{2i} \mid 0 \leq i \leq \frac{n}{2} - 1\}$ and $\{sr^{2i+1} \mid 0 \leq i \leq \frac{n}{2} - 1\}$.*

The z-classes in D_{2n} are as follows:

1. *If n is odd,*

$$\mathcal{Z}_{D_{2n}}(r^j) = \{r^i \mid i = 0, \ldots, n-1\}, \text{ for any } j = 0, \ldots, n-1$$
$$\mathcal{Z}_{D_{2n}}(sr^j) = \{1, sr^j\}, \text{ for any } j = 0, \ldots, n-1$$

   There are three z-classes:

$$\{1\}, \{r^j \mid j = 1, \ldots, n-1\}, \{sr^j \mid j = 0, \ldots, n-1\}$$

2. *If n is even*

$$\mathcal{Z}_{D_{2n}}(r^{\frac{n}{2}}) = D_{2n}$$
$$\mathcal{Z}_{D_{2n}}(r^j) = \{r^i \mid i = 0, \ldots, n-1\}, \text{ for any } j = 0, \ldots, n-1, j \neq \tfrac{n}{2}$$
$$\mathcal{Z}_{D_{2n}}(sr^j) = \{1, sr^j, r^{\frac{n}{2}}, sr^{j+\frac{n}{2}}\}, \text{ for any } j = 0, \ldots, n-1$$

   (a) *When n is divisible by 4*
   There are four z-classes:

$$\{1, r^{\frac{n}{2}}\}, \{r^j \mid j = 1, \ldots, n-1, j \neq \tfrac{n}{2}\},$$
$$\{sr^{2j} \mid j = 0, \ldots, \tfrac{n}{2}\}, \{sr^{2j+1} \mid j = 0, \ldots, \tfrac{n}{2}\}$$

(b) *When n is not divisible by 4*

There are three z-classes:

$$\{1, r^{\frac{n}{2}}\}, \{r^j \mid j = 1, \ldots, n-1, j \neq \tfrac{n}{2}\}, \{sr^j \mid j = 0, \ldots, n-1\}$$

Chapter 2

Conjugacy classes of $GL_n(\mathbb{F})$

The *general linear group* $GL_n(\mathbb{F})$ is the set of $n \times n$ invertible matrices together with the operation of matrix multiplication. In this chapter, we will discuss conjugacy classes of any matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ and since $GL_n(\mathbb{F})$ is a subset, the same will apply for $GL_n(\mathbb{F})$. Further, we will discuss some results we had assumed in Section 4.3., as discussed in [BR02]
Two $n \times n$ matrices A and B are *conjugates* when they are similar, i.e,

$$\exists \text{ invertible matrix } P \in \mathrm{Mat}_n(\mathbb{F}) \text{ such that } PB = AP$$

Thus, similarity is the same as conjugacy in $GL_n(\mathbb{F})$. We will discuss various results that will help us determine simple representatives in each conjugacy class.
We also know that, given a vector space V over the field $\mathbb{F}$, we can associate a matrix to every linear map $T : V \to V$. Suppose V is n-dimensional, then it has a basis $\{v_1, v_2, \ldots, v_n\}$. We can determine the action of T by expressing each $T(v_i)$ as a linear combination of the basis elements and the scalars associated with each of these actions help determine the matrix A associated with this linear transformation.
Finally, if V is a vector space over a field $\mathbb{F}$ and if $T : V \to V$ is linear, then a subspace W of V is said to be T-*invariant* if $T(W) \subseteq W$, i.e.,

$$\text{if } x \in W \Rightarrow T(x) \in W$$

2.1 Primary Decomposition Theorem

Theorem 2.1.1. *If $T : V \to V$ is linear, then for every polynomial $p \in \mathbb{F}[x]$, the subspaces Im $p(T)$ and Ker $p(T)$ are T-invariant.*

Proof. For every polynomial p, we have $T \circ p(T) = p(T) \circ T$.
If $x = p(T)(y)$, then $T(x) = (T \circ p(T))(y) = p(T)[T(y)]$. We can clearly see that Im $p(T)$ is T-invariant.
Now, if $p(T)(x) = 0_V$, then $T(p(T)(x)) = T(0_V) = 0_V \Rightarrow T(p(T)) \in \text{Ker } p(T)$. $\square$

Suppose V is n-dimensional, $T : V \to V$ is linear and the subspace W of V is T-invariant. Then T induces a linear mapping $\overline{T}$

$$\begin{aligned} \overline{T} : W &\to W \\ w &\mapsto \overline{T}(w) = T(w) \in W \end{aligned}$$

Choose a basis $\{w_1, \ldots, w_r\}$ of W and extend it to the following basis of V,

$$\{w_1, \ldots, w_r, v_1, \ldots, v_{n-r}\}.$$

Let us consider the matrix of T relative to this basis. Since W is T-invariant, $T(w_i) \in W$ and so $\forall i$,

$$T(w_i) = \lambda_{i1}w_1 + \cdots + \lambda_{ir}w_r + 0v_1 + \cdots + 0v_{n-r}$$

This matrix is of the form

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

where A is the $r \times r$ matrix that represents the mapping induced on W by T.
If $V = \bigoplus_{i=1}^{k} V_k$, where each V_i is T-invariant and if B_i is a basis of V_i $\forall$ i, then the matrix relative to the basis $B = \bigcup_{i=1}^{k} B_i$ is of the block diagonal form

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

in which A_i is the matrix representing the mapping induced on V_i by T, so A_i is of size dim $V_i \times$ dim V_i.

Theorem 2.1.2. *[Primary Decomposition Theorem] Let V be a non-zero finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear map. Let the characteristic and* *minimal polynomials of T be*

$$c_T = p_1^{d_1} p_2^{d_2} \dots p_k^{d_k},\ m_T = p_1^{e_1} p_2^{e_2} \dots p_k^{d_k}$$

repectively, where $p_1(x), \dots, p_k(x)$ are distinct irreducible polynomials in $\mathbb{F}[x]$. Then each subspace $V_i = Ker\ p_i^{e_i}(T)$ is T-invariant and $V = \bigoplus_{i=1}^k V_i$.
Further, if $T_i : V_i \to V_i$ is the linear mapping that is induced on V_i by T then the characteristic polynomial of T_i is $p_i^{d_i}$ and the minimal polynomial of T_i is $p_i^{e_i}$.

Proof. If $k = 1$, this is trivial.
Suppose $k \geq 2$. For $i = 1, \dots, k$ let

$$q_i = m_T / p_i^{e_i} = \prod_{j \neq i} p_j^{e_j}$$

No irreducible factor exists such that it is common to each $q_1, \dots, q_k$ and so there exist $a_1, \dots, a_k \in \mathbb{F}[x]$ such that

$$a_1 q_1 + a_2 q_2 + \dots + a_k q_k = 1$$

Let us write $q_i a_i = t_i\ \forall\ i$, we get,

$$t_1(T) + t_2(T) + \dots + t_k(T) = \text{ id}_V \tag{2.1}$$

By definition of q_i. we have that if $i \neq j$ then m_T divides $q_i q_j$. As a consequence, $q_i(T)q_j(T) = 0$ for $i \neq j$ and

$$t_i(T)t_j(T) = 0 \text{ for } i \neq j \tag{2.2}$$

By Equations 2.1 and 2.2, we can see that $t_i(T)$ is a projection and

$$V = \bigoplus_{i=1}^{k} \text{ Im } t_i(T)$$

By Theorem 2.3.1 each of the subspaces Im $t_i(T)$ is T-invariant. We need to now show that Im $t_i(T) =$ Ker $p_i^{e_i}(T)$.
Since $p_i^{e_i} q_i = m_T \Rightarrow p^{e_i}(T) q_i(T) = m_T(T) = 0 \Rightarrow p_i^{e_i}(T) q_i(T) a_i(T) = 0 \Rightarrow$ Im $t_i(T) \subseteq$ Ker $p_i^{e_i}(T)$.

Now, for the reverse inclusion,

$$\begin{aligned}
t_j(T) = a_j(T)q_j(T) = a_j(T)\prod_{i\neq j} p_i^{e_i}(T) \ \forall\ j \\
\Rightarrow \text{Ker } p_i^{e_i}(T) &\subseteq \bigcap_{j\neq i} \text{ Ker } t_j(T) \\
&\subseteq \sum_{j\neq i} \text{ Ker } t_j(T) \\
&= \text{ Ker } (\text{id}_V - t_i(T)) \text{ by Equation 2.3} \\
&= \text{Im } t_i(T)
\end{aligned}$$

For the induced mapping $T_i : V_i \to V_i$, let m_i be its minimal polynomial. If $p_i^{e_i}(T)$ is the zero map on V_i, then $p_i^{e_i}(T_i)$ is the zero map as well. Then $m_{T_i}|p_i^{e_i} \Rightarrow m_{T_i}|m_T$ and m_{T_i} are relatively prime. Let $g \in \mathbb{F}[x]$ be a multiple of $m_{T_i}\ \forall\ i$. Then $g(T_i)$ is the zero map on V_i. For every $x = \sum_{i=1}^k v_i \in \bigoplus_{i=1}^k = V_i$, we have

$$g(T)(x) = \sum_{i=1}^k g(T)(v_i) = \sum_{i=1}^k g(T_i)(v_i) = 0_V$$

So, $g(T) = 0 \Rightarrow m_T|g$. We can see that m_T is the least common multiple of $m_{T_1}, \ldots, m_{T_k}$. Since these polynomials are relatively prime, $m_T = \prod_{i=1}^k m_{T_i}$. We know that $m_t = \prod_{i=1}^k p_i^{e_i}$ and $m_{T_i}|p_i^{e_i}$. Given all these polynomials are monic, then $m_{T_i} = p_i^{e_i}$ for $i = 1, \ldots, k$.
We can put together a basis for V using the bases of subspaces V_i and the matrix of T with respect to such a basis is of the block diagonal form

$$M = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

From the theory of determinants

$$\det\ (xI - M) = \prod_{i=1}^k \det\ (xI - A_i)$$

We see that $c_T = \prod_{i=1}^k c_{T_i}$. We know that $m_{T_i} = p_i^{e_i}$ and since m_{T_i} and c_{T_i} have the same zeros, then $c_{T_i} = p_i^{r_i}$ some $r_i \geq e_i$. Therefore $\prod_{i=1}^k p_i^{r_i} = c_T = \prod_{i=1}^k p_i^{d_i}$ from which it follows that $r_i = d_i$ for $i = 1, \ldots, k$. $\square$

Corollary 2.1.3. *dim* $V_i = d_i$ *deg* p_i

Corollary 2.1.4. *If all the eigenvalues of* T *lie in* $\mathbb{F}$*, such that*

$$c_T = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \ldots (x - \lambda_k)^{d_k}$$
$$m_T = (x - \lambda_1)^{e_1}(x - \lambda_2)^{e_2} \ldots (x - \lambda_k)^{e_k}$$

then $V_i = (Ker\ T - \lambda_i id_V)^{e_i}$ *is* T*-invariant, of dimension* d_i *and* $V = \bigoplus_{i=1}^{k} V_i$

A special case of the Primary Decomposition Theorem is when each of the irreducible factors p_i of m_T is linear and $e_i = 1$, i.e.,

$$m_T = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k)$$

In this case, $T : V \to V$ is said to be *diagonalisable* if there is a basis of V consisting of eigenvectors of T, equivalently, if there is a basis of V with respect to which the matrix of T is diagonal.

Theorem 2.1.5. *Let* V *be a non-zero finite-dimensional vector space and let* $T : V \to V$ *be linear. Then the following statements are equivalent:*

1. *the minimal polynomial* m_T *of* T *is a product of distinct linear factors*

2. T *is diagonalisable*

Proof. Suppose

$$m_T = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k)$$

where $\lambda_1, \ldots, \lambda_k$ are distinct elements in $\mathbb{F}$. By Theorem 2.1.2, V is the direct sum of T-invariant subspaces $V_i = \text{Ker } (T - \lambda_i \text{id}_V)$.
$\forall\ x \in V_i$, we have $(T - \lambda_i \text{id}_V)(x) = 0_V \Rightarrow T(x) = \lambda_i x$. Each λ_i is an eigenvalue of T, every non-zero element of V_i is an eigenvector of T associated with λ_i. We can then put together bases of $V_1, \ldots, V_k$ to form a basis for V. Then V has a basis consisting of eigenvectors of T and T is diagonalisable.
Conversely, let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of T. The mapping $p(T)$ where

$$p = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_k)$$

maps every basis vector to 0_V and as a consequence, $p(T) = 0$. The minimal polynomial m_T must divide p and thus coincides with p since every eigenvalue of T is a zero of m_T. $\square$

Remark. *Although we will not prove this, it is interesting to note that two diagonalisable linear mappings $f, g : V \to V$ are simultaneously diagonalisable iff $f \circ g = g \circ f$.*

2.2 Triangular Form

We will now look at a more general situation where the minimal polynomial of T factorises as a product of linear factor that are not necessarily distinct, i.e.,

$$m_T = \prod_{i=1}^{k} (x - \lambda_i)^{e_i}$$

where each $e_i \geq 1$. This is always true when the underlying field is algebraically closed.
By Corollary 2.1.4, we can write V as a direct sum of the T-invariant subspaces $V_i = \text{Ker } (T - \lambda_i \text{id}_{V_i})^{e_i}$. Let $T_i : V_i \to V_i$ be the linear mapping induced on V_i by T and consider the mapping $T_i - \lambda_i \text{id}_{V_i} : V_i \to V_i$. Then $(T_i - \lambda_i \text{id}_{V_i})^{e_i}$ is the zero map on V_i and so $T_i - \lambda_i \text{id}_{V_i}$ is nilpotent.

Theorem 2.2.1. *Let V be a non-zero finite-dimensional vector space and let $T : V \to V$ be a nilpotent linear mapping. Then there is an ordered basis $\{v_1, \ldots, v_n\}$ of V such that*

$$\begin{aligned} T(v_1) &= 0_V \\ T(v_2) &\in \mathit{Span}\ \{v_1\} \\ T(v_3) &\in \mathit{Span}\ \{v_1, v_2\} \\ &\vdots \\ T(v_n) &\in \mathit{Span}\ \{v_1, \ldots, v_{n-1}\} \end{aligned}$$

Proof. Since T is nilpotent, there is a positive integer m such that $T^m = 0$. Let us assume $T \neq 0$. Then there is a smallest positive integer k such that $T^k = 0$. Then $T^i \neq 0$ for $1 \leq i \leq k-1$. Since $T^{k-1} \neq 0$, there exists $v \in V$ such that $T^{k-1}(v) \neq 0_V$. Let $v_1 = T^{k-1}(v)$. Then $T(v_1) = 0_V$. If we proceed recursively, we will be able to find $v_1, \ldots, v_r$ satisfying the conditions. Now, consider the subspace $W = \text{Span}\{v_1, \ldots, v_r\}$ and let $W \neq V$.
There now are two possibilities depending on whether $\text{Im } T \subseteq W$ or $\text{Im } T \nsubseteq W$, which we

need to consider to show that $\{v_1, \ldots, v_r\}$ is a basis.
In the first case, let v_{r+1} be any element of $V \backslash W$. In the latter case, since we have the chain

$$\{0_V\} = \text{Im } T^k \subseteq \text{Im } T^{k-1} \subseteq \cdots \subseteq \text{Im } T^2 \subseteq \text{Im } T$$

there is a positive integer j such that Im $T^j \nsubseteq W$ and Im $T^{j+1} \subseteq W$. Choose $v_{r+1} \in \text{Im } T^j$ with $v_{r+1} \notin W$. Then $\{v_1, \ldots, v_{r+1}\}$ is linearly independent, with $T(v_{r+1}) \in W$. □

Corollary 2.2.2. *If $T : V \to V$ is nilpotent then there is an ordered basis of V with respect to which the matrix of T is upper triangular with all diagonal entries 0.*

Let $g_i = T_i - \lambda_i \text{id}_{V_i}$. Then

$$\begin{aligned}
\text{Mat } T_i &= \text{Mat } g_i + \lambda_i \text{Mat id}_{V_i} \\
&= \begin{pmatrix} \lambda_i & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & \lambda_i & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1,n} \\ 0 & 0 & 0 & \ldots & \lambda_i \end{pmatrix}
\end{aligned}$$

Theorem 2.2.3 (Triangular Form)**.** *Let V be a non-zero finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear mapping whose characteristic and minimal polynomials are*

$$c_T = \prod_{i=1}^{k} (x - \lambda_i)^{d_i}, \; m_T = \prod_{i=1}^{k} (x - \lambda_i)^{e_i}$$

for distinct $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ and $e_i \leq d_i$. Then there is an ordered basis of V with respect to which the matrix of T is upper triangular; more specifically, is a block diagonal matrix

$$M = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

in which A_i is a $d_i \times d_i$ upper triangular matrix

$$\begin{pmatrix} \lambda_i & \star & \dots & \star \\ 0 & \lambda_i & \dots & \star \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

in which the entries marked $\star$ are elements of $\mathbb{F}$.

Now, if $T : V \to V$ is linear and every eigenvalue of T lies in $\mathbb{F}$, which is the ground field of V, then each induced mapping on the T_i on the T-invariant subspace V_i =Ker $(T - \lambda_i \mathrm{id}_V)^{e_i}$ can be written in the form $T_i = g_i + \lambda_i \mathrm{id}_{V_i}$, where g_i is nilpotent.

Theorem 2.2.4 (Jordan Decomposition)**.** *Let V be a non-zero finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear mapping all of whose eigenvalues belong to $\mathbb{F}$. Then there is a diagonalisable linear mapping $\delta : V \to V$ and a nilpotent linear mapping $\eta : V \to V$ such that $T = \delta + \eta$ and $\delta \circ \eta = \eta \circ \delta$. Moreover, there are polynomials $p, q \in \mathbb{F}[x]$ such that $\delta = p(T)$ and $\eta = q(T)$. Furthermore, δ and η are uniquely determined.*

Proof. The minimal polynomial of T is $m_T = \prod_{i=1}^{k} (x - \lambda_i)^{e_i}$ where $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ are distinct. Further, $V = \bigoplus_{i=1}^{k} V_i$ where V_i =(Ker $T - \lambda_i \mathrm{id}_V)^{e_i}$.
Let $\delta : V \to V$ be given by $\delta = \sum_{i=1}^{k} \lambda_i p_i$ where $p_i : V \to V$ is the projection on V_i perpendicular to $\sum_{j \neq i} V_j$. Then for every $v_i \in V_i$, we have $\delta(v_i) = \big(\sum_{j=1}^{k} \lambda_j p_j\big)(v_i) = \lambda_i(v_i)$ and V has a basis consisting of eigenvectors of $\delta \Rightarrow \delta$ is diagonalisable.
Let $\eta = T - \delta$. Then $\forall\ v_i \in V_i$,

$$\begin{aligned} \eta(v_i) &= T(v_i) - \delta(v_i) = (T - \lambda_i \mathrm{id}_V)(v_i) \\ \Rightarrow \eta^{e_i}(v_i) &= (T - \lambda_i \mathrm{id}_V)^{e_i}(v_i) = 0_V \end{aligned}$$

For some r, Ker η^r contains a basis of V, so $\eta^r = 0$ and hence η is nilpotent.
Since $V = \bigoplus_{i=1}^{k} V_i$, every $v \in V$ can be uniquely written in the form $v = v_1 + \cdots + v_k$ with

$v_i \in V_i$. Since each V_i is T-invariant, we have

$$\begin{aligned} p_i[T(v)] &= p_i[T(v_1) + \cdots + T(v_k)] = T(v_i) = T[p_i(v)] \\ \Rightarrow p_i \circ T &= T \circ p_i \ \forall \ i \\ \delta \circ T &= \sum_{i=1}^{k} \lambda_i p_i \circ T = \sum_{i=1}^{k} \lambda_i T \circ p_i \\ &= T \circ \sum_{i=1}^{k} \lambda_i p_i = T \circ \delta \end{aligned}$$

Now, by Theorem 2.1.2, we know that $p_i = t_i(T)$. Then by definition, we have $\delta = p(T)$, where $p = \sum_{i=1}^{k} \lambda_i t_i$. Since $\eta = T - \delta, \exists \ q \in \mathbb{F}[x]$ such that $\eta = q(T)$.
Suppose $\delta', \eta' : V \to V$ are diagonalisable and nilpotent respectively, with $T = \delta' + \eta'$ and $\delta' \circ \eta' = \eta' \circ \delta'$. Now, we have just shown that, there are polynomials p, q such that $\delta = p(T)$ and $\eta = q(T) \Rightarrow \delta' \circ \delta = \delta \circ \delta'$ and $\eta' \circ \eta = \eta \circ \eta'$.
$\delta + \eta = \delta' + \eta' \Rightarrow \delta - \delta' = \eta' - \eta$. Further, $\eta' - \eta$ is nilpotent and can be represented by a nilpotent matrix N. Also, δ', δ commute and there is a basis of V consisting of eigenvectors of both δ and δ'. Then each such eigenvector is an eigenvector $\delta - \delta'$ is represented by a diagonal matrix D.
Now, N and D are similar and then the only possibility $N = D = 0 \Rightarrow \delta - \delta' = \eta' - \eta = 0 \Rightarrow \delta = \delta'$ and $\eta = \eta'$. □

2.3 Jordan Canonical Form

The aim now is to find better bases for the subspaces that appear as the direct summands in the Primary Decomposition Theorem.
If the linear mapping $T : V \to V$ is nilpotent then the smallest positive integer k such that $T^k = 0$ is called the *index* of T.

Theorem 2.3.1. *If $T : V \to V$ is linear then, for every positive integer i,*

1. *Ker $T^i \subseteq$Ker T^{i+1}*
2. *if $x \in$ Ker T^{i+1}, then $T(x) \in$ Ker T^i*

Proof. 1. If $x \in \text{Ker } T^i$ then $T^i(x) = 0_V \Rightarrow T^{i+1}(x) = T[T^i(x)] = T(0_V) = 0_V$
$\therefore x \in \text{Ker } T^{i+1}$

2. If $x \in \text{Ker } T^{i+1}$ then $T^i[T(x)] = T^{i+1}(x) = 0_V \Rightarrow T(x) \in \text{Ker } T^i$.

$\square$

Theorem 2.3.2. *Let V be a non-zero vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear mapping that is nilpotent of index k. Then there is the chain of distinct subspaces*

$$\{0_V\} \subset \ Ker\ T \subset \ Ker\ T^2 \subset \cdots \subset \ Ker\ T^{k-1} \subset \ Ker\ T^k = V.$$

Proof. First Ker $T \neq \{0_V\}$, otherwise we would have $T^{k-1}(x) = 0_V\ \forall\ x$ which is a contradiction.
As a result of Theorem 2.3.1, we just need to show that

$$\text{Ker } T^i \neq \text{Ker } T^{i+1} \text{ for } i = 1, \ldots, k-1$$

Suppose $\exists\ i\ \in \{1, \ldots, k-1\}$ such that Ker T^i =Ker T^{i+1} .
Then, $\forall\ x \in V$, we have

$$0_V = T^k(x) = T^{i+1}[T^{k-(i+1)}(x)]$$

where $T^{k-(i+1)}(x) \in \text{Ker } T^{i+1} = \text{Ker } T^i$ and so

$$0_V = T^i[T^{k-(i+1)}(x)] = T^{k-1}(x)$$

This leads to a contradiction. $\square$

A square matrix of the form

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

in which all the diagonal entries are λ and all the entries immediately above the diagonal entries are 1, and all other entries are 0 is called *elementary Jordan matrix associated with* $\lambda \in \mathbb{F}$.

A matrix of the form

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

where each J_i is an elementary Jordan matrix associated with λ and all other entries are 0 is called a *Jordan block matrix*.

Theorem 2.3.3. *Let V be a non-zero finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be linear and nilpotent of index k. Then there is a basis of V with respect to which the matrix of T is a Jordan block matrix associated with the eigenvalue 0.*

Proof. For $i = 0, \ldots, k$, let $W_i = \text{Ker } T^i$. We have the following chain,

$$\{0_V\} = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_{k-1} \subset W_k = V$$

Now, choose a basis B_1 of W_1 and extend this to a basis $B_2 = B_1 \cup T_2$ of W_2 where $T_2 \subseteq W_2 \backslash W_1$ and so on. Then $B_k = B_1 \cup T_2 \cup \cdots \cup T_k$ is a basis of V.
Let $T_k = \{x_1, \ldots, x_\alpha\}$. Then Theorem 2.3.2 gives us a linearly independent subset $\{T(x_1), \ldots, T(x_\alpha)\}$ of W_{k-1}. This set is disjoint from W_{k-2}. Consider the set

$$B_{k-2} \cup \{T(x_1), \ldots, T(x_\alpha)\}$$

This is linearly independent in W_{k-1} and can be extended to the following basis of W_{k-1}

$$B_{k-2} \cup \{T(x_1), \ldots, T(x_\alpha)\} \cup \{y_1, \ldots, y_\beta\}$$

where $y_i \in W_{k-1}\ W_{k-2}\ \forall\ i$. We have replaced T_{k-1} in the basis B_k by

$$T^*_{k-1} = \{T(x_1), \ldots, T(x_\alpha)\} \cup \{y_1, \ldots, y_\beta\}$$

We can similarly repeat this argument to construct a basis of W_{k-2} of the form

$$B_{k-3} \cup \{T^2(x_1), \ldots, T^2(x_\alpha)\} \cup \{T(y_1), \ldots, T(y_\beta)\} \cup \{z_1, \ldots, z_\gamma\}$$

where $z_i \in W_{k-2} \backslash W_{k-3}\ \forall\ i$. We have replaced T_{k-2}.
Continuing in this way, we can replace the basis B_k of V by the basis described in the

following array:

$$\begin{array}{llllllllll}
T_k: & x_1, & \ldots, & x_\alpha, & & & & & & \\
T_{k-1} \rightsquigarrow & T(x_1), & \ldots, & T(x_\alpha), & y_1, & \ldots, & y_\beta, & & & \\
T_{k-2} \rightsquigarrow & T^2(x_1), & \ldots, & T^2(x_\alpha), & T(y_1), & \ldots, & T(y_\beta), & z_1, & \ldots, & z_\gamma \\
\vdots & & & & & & & & & \\
B_1 \rightsquigarrow & T^{k-1}(x_1), & \ldots, & T^{k-1}(x_\alpha), & T^{k-2}(y_1), & \ldots, & T^{k-2}(y_\beta), & q_1 & \ldots\ldots, & q_\omega.
\end{array}$$

Take the first column from the bottom, then the second column from the bottom and so on, to order this basis of V. This matrix of T relative to the ordered basis is a Jordan block matrix associated with the eigenvalue 0. □

An assumption we make for Theorem 2.3.3 is that T is nilpotent. We will now attempt to extend the scope of this theorem to when T is not nilpotent. We will assume that all the eigenvalues of T lie in $\mathbb{F}$.

Theorem 2.3.4 (Jordan Form). *Let V be a non-zero finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be linear. If $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T and if each λ_i belongs to $\mathbb{F}$ then there is an ordered basis of V with respect to which the matrix of T is a block diagonal matrix*

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

in which J_i is a Jordan block matrix associated with λ_i.

The proof for this theorem is similar to the proof of Theorem 2.3.3.

2.4 Rational Canonical Form

Jordan forms can be used only when all the irreducible polynomials in the minimal polynomials are linear, which happens when the eigenvalues belong to the underlying field $\mathbb{F}$. We know that this occurs when $\mathbb{F}$ is algebraically closed. We will now look for a canonical representation for any general case.

The additive group V/W along with the operation of multiplication by scalars which makes the natural surjection map linear becomes a vector space over $\mathbb{F}$ is called the *quotient space* *of V by W.* The natural surjection map is defined as follows:

$$\begin{aligned} \natural_W : V &\to V/W \\ x &\mapsto x + W \end{aligned}$$

Theorem 2.4.1. *Let V be a finite-dimensional vector space and let W be a subspace of V. Then the quotient space V/W is also finite-dimensional. Moreover, if $\{v_1, \dots, v_m\}$ is a basis of W and $\{x_1 + W, \dots, x_k + W\}$ is a basis of V/W then $\{v_1, \dots, v_m, x_1, \dots, x_k\}$ is a basis of V.*

Proof. Let $\{x_1 + W, \dots, x_p + W\}$ be any linearly independent subset of V/W. Then the set $\{x_1, \dots, x_p\}$ of coset representatives is a linearly independent subset of V. Suppose $\sum_{i=1}^{p} \lambda_i x_i = 0_V$. Then,

$$0_{V/W} = \natural_W(0_V) = \natural_W(\sum_{i=1}^{p} \lambda_i x_i) = \sum_{i=1}^{p} \lambda_i \natural_W(x_i) = \sum_{i=1}^{p} \lambda_i (x_i + W)$$

and so $\lambda_i = 0 \ \forall \ i \Rightarrow p \leq$dim V. Then, every linearly independent subset of V/W has at most dim V elements. Hence V/W is of finite dimensions.

Suppose $\{v_1, \dots, v_m\}$ is a basis of W and that $\{x_1 + W, \dots, x_k + W\}$ is a basis of V/W. Consider the set $B = \{v_1, \dots, v_m, x_1, \dots, x_k\}$. Applying $\natural_W$ to any linear combination of elements of B, we see that it is linearly independent. Now for every $x \in V$ we have $x + W \in V/W$ and so there exists scalars λ_i such that

$$x + W = \sum_{i=1}^{k} \lambda_i (x_i + W) = (\sum_{i=1}^{k} \lambda_i x_i) + W$$

$\Rightarrow x - \sum_{i=1}^{k} \lambda_i x_i \in W$. Then $\exists \ \mu_j$ such that $x - \sum_{i=1}^{k} \lambda_i x_i = \sum_{i=1}^{m} \mu_j v_j$.

As a consequence, x is a linear combination of the elements of B. Then, the linearly independent set B is also a spanning set and therefore a basis of V. $\square$

Theorem 2.4.2. *Let V be a finite-dimensional vector space and let $T : V \to V$ be linear. If W is an T-invariant subspace of V then the prescription*

$$T^{+}(x + W) = T(x) + W$$

defines a linear mapping $T^{+} : V/W \to V/W$, the minimal polynomial of which divides the *minimum polynomial of T.*

Proof. If $x+W=y+W$ then $x-y\in W$. Since W is T-invariant, $T(x)-T(y)=T(x-y)\in W\Rightarrow T(x)+W=T(y)+W$. The above prescription defines a mapping from V/W to itself.

$$\begin{aligned}T^+[(x+W)+(y+W)] &= T^+[(x+y)+W]\\ &= T(x+y)+W\\ &= [T(x)+T(y)]+W\\ &= [T(x)+W)+[T(y)+W]\\ &= T^+(x+W)+T^+(y+W)\end{aligned}$$

$$\begin{aligned}T^+[\lambda(x+W)] &= T^+[\lambda x+W]\\ &= T(\lambda x)+W\\ &= \lambda T(x)+W\\ &= \lambda[T(x)+W]\\ &= \lambda T^+(x+W)\end{aligned}$$

By induction we will show that

$$(T^+)^n=(T^n)^+$$

This is clearly true for $n=1$. Now, suppose that $(T^+)^n=(T^n)^+$. Then for every $x\in W$

$$\begin{aligned}(T^+)^{n+1}(x+W) &= T^+[(T^+)^n(x+W)]\\ &= T^+[(T^n)^+(x+W)]\\ &= T^+[T^n(x)+W]\\ &= T[T^n(x)]+W\\ &= T^{n+1}(x)+W\\ &= (T^{n+1})^+(x+W)\end{aligned}$$

Then, for every polynomial $p=\sum_{i=0}^m a_iX^i$, we have

$$p(T^+)=\sum_{i=0}^m a_i(T^+)^i=\sum_{i=0}^m a_i(T^i)^+=[p(T)]^+$$

in particular, $p = m_T$, we obtain $m_T(T^+) = 0 \Rightarrow m_{T^+}|m_T$ $\square$

The intersection of any family of T-invariant subspaces of V is also an T-invariant subspace of V. We will denote the intersection of all the T-invariant subspaces that contain X by Z_X^T. In case $X = x$, Z_X^T is simply Z_x.

Theorem 2.4.3. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be linear. Then for every $x \in V$,*

$$Z_x = \{p(T)(x)|p \in \mathbb{F}[x]\}$$

Proof. The set $W = \{p(T)(x)|p \in \mathbb{F}[x]\}$ is a subspace of V that contains x. This subspace is T-invariant.
Suppose now that U is a T-invariant subspace that contains x. Then $T^k(x) \subseteq U \ \forall \ k$. U also contains $p(T)(x)$ for every polynomial $p \in \mathbb{F}[x]$. Thus $W \subseteq U$. Hence W is the smallest T-invariant subspace that contains x and coincides with Z_x. $\square$

We will now construct a basis for the subspace Z_x. Consider the sequence

$$x, T(x), T^2(x), \ldots, T^r(x), \ldots$$

of elements of Z_x. Then there exists a smallest positive integer k such that $T^k(x)$ is a linear combination of the elements that precede it

$$T^k(x) = \lambda_0 x + \lambda_1 T(x) + \cdots + \lambda_{k_1} T^{k-1}(x)$$

and $\{x, T(x), \ldots, T^{k-t}(x)\}$ is then a linearly independent subset of Z_x.
Writing $a_i = -\lambda_i$ for $i = 0, \ldots, k-1$ we deduce that the polynomial

$$m_x = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1} + x^k$$

is the monic polynomial of least degree such that $m_x(T)(x) = 0_V$. We will call m_x the *T-annihilator* of x.

Theorem 2.4.4. *Let V be a finite-dimensional vector space and let $T : V \to V$ be linear. If $x \in V$ has T-annihilator*

$$m_x = a_0 + a_1 x + \cdots + a_{k-1} x^{k-1} + x^k$$

then the set

$$B_x = \{x, T(x), \ldots, T^{k-1}(x)\}$$

is a basis of Z_x *and therefore dim* Z_x = *deg* m_x. *Moreover, if* $T_x : Z_x \to Z_x$ *is the induced linear mapping on the* T*-invariant subspace* Z_x *then the matrix of* T_x *relative to the ordered basis* B_x *is*

$$C_{m_x} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -a_{k-1} \end{pmatrix} \tag{2.3}$$

Finally, the minimal polynomial of T_x *is* m_x.

Proof. B_x is linearly independent and $T^k(x) \in$ Span B_x.
By induction, we prove that $T^n(x) \in$ Span B_x for every n. Suppose that $n > k$ and that $T^{n-1}(x) \in$ Span B_X. Then $T^{n-1}(x)$ is a linear combination of $x, T(x), \ldots, T^{k-1}(x)$ and so $T^n(x)$ is a linear combination of $T(x), T^2(x), \ldots, T^k(x) \Rightarrow T^n(x) \in$ Span $B_x \Rightarrow p(T)(x) \in$ Span B_x for every polynomial p. Thus $Z_x \subseteq$ Span B_x. The reverse inclusion is trivial.
Now, B_x is a basis of Z_x. Now, since

$$\begin{aligned} T_x(x) &= T(x) \\ T_x[T(x)] &= T^2(x) \\ &\vdots \\ T_x[T^{k-2}(x)] &= T^{k-1}(x) \\ T_x[T^{k-1}(x)] &= T^k(x) = -a_0 x - a_1 T(x) - \cdots - a_{k-1} T^{k-1}(x) \end{aligned}$$

The matrix of T_x relative to the basis B_x is the matrix 2.3.
Now, let the minimal polynomial of T_x be

$$m_{T_x} = b_0 + b_1 x + \cdots + b_{r-1} T^{r-1}(x) + T^r(x)$$

Then

$$0_V = m_{T_x}(T)(x) = b_0 x + b_1 T(x) + \cdots + b_{r-1} T^{r-1}(x) + T^r(x)$$

We can see that $T^r(x)$ is a linear combination of $x, T(x), \ldots, T^{r-1}(x)$ and $k \leq r$. But m_{T_x} is the zero map on Z_x and so $m_x(T_x)$ is also a zero map. As a result, we have m_{T_x} divides

m_x and so $r \leq k$. $\therefore r = k$. $\Rightarrow m_{T_x} = m_x$. $\square$

A subspace W of V is called *T-cyclic* if it is T-invariant and has a basis of the form $\{x, T(x), \ldots, T^r(x)\}$. Such a basis is called a *cyclic basis*, and x is called a *cyclic vector* for W.

Theorem 2.4.4 shows that x is a cyclic vector for the subspace Z_x with cyclic basis B_x. The subspace Z_x is called the *T-cyclic subspace spanned by* $\{x\}$. The matrix C_{m_x} that we defined in Theorem 2.4.4 is called the *companion matrix* of the T-annihilator m_x.

Theorem 2.4.5. *Let W be an T-invariant subspace of V. Then $\forall\ x \in V$ both the T-annihilator of T and the T^+ -annihilator of $x + W$ divide the minimal polynomial of T.*

Proof. The proof is a consequence of Theorems 2.4.2 and 2.4.4. $\square$

Theorem 2.4.6 (Cyclic Decomposition)**.** *Let V be a non-zero vector space of finite dimension over a field $\mathbb{F}$ and let $T : V \to V$ be linear with minimal polynomial $m_T = p^t$ where p is irreducible over $\mathbb{F}$. Then there are T-cyclic vectors $x_1, \ldots, x_k$ and positive integers $n_1, \ldots, n_k$ with each $n_i \leq t$ such that*

(i) $V = \bigoplus_{i=1}^{k} Z_x$

(ii) the T-annihilator of x_i is p^{n_i}.

Proof. We will prove this by induction on dim V. When dim $V = 1$, the result is trivial. Suppose the result holds for all vector spaces of dimension less than $n = \dim V$.

Since $m_T = p^t$, $\exists\ x_1 (\neq 0) \in V$ with $p^{t-1}(f)(x_1) \neq 0_V$. Then m_{x_1} is the T-annihilator of x_1. Let $W = Z_x$ and $T^+ : V/W \to V/W$ be the induced mapping. By Theorem 2.4.2, the minimal polynomial of T^+ divides $m_T = p^t$.

$$\Rightarrow V/W = \bigoplus_{i=2}^{k} Z_{y_i+W}$$

where $Z_{y_2+W}, \ldots, Z_{y_k+W}$ are T^+-cyclic subspaces of V/W. Further, for $2 \leq i \leq k$, the T^+-annihilator of $y_i + W$ is $p^{n_i}, n_i \leq t$.

Now there exists a polynomial h such that

$$p(T)^{n_i}(y_i) = h(T)(x_1)$$

for some $x_1 \in y_1 + W$.

$$\Rightarrow 0_V = p(T)^t(y_i) = p(T)^{t-n_i} h(T)(x_1)$$

Now, as p^t is the T-annihilator of x_1, so $p^t | p^{t-n_i} h \Rightarrow h = p^{n_i} q$ for some polynomial q.
Let $x_i = y_i - q(T)(x_1)$. Then

$$y_i = x_i = q(T)(x_1) \in W \Rightarrow x_i \in y_i + W$$

The T^+-annihilator of $y_i + W$ divides the T-annihilator of x_i. But,

$$p(T)^{n_i}(x_i) = p(T)^{n_i}[y_i - q(T)(x_1)] = p(T)^{n_i}(y_i) - h(T)(x_1) = 0_V$$

Clearly, p^{n_i} is the T-annihilator of x_i.
Now, let deg $p = d$. Then deg $p^{n_i} = dn_i$. Since p^{n_i} is the T-annihilator of x_i and the T^+-annihilator of $x_i + W$, by Theorem 2.4.2, there is a basisA_i for Z_{x_i}

$$A_i = \{x_i, f(x_i), \ldots, f^{dn_i - 1}(x_i)\}$$

and a basis B_i for $Z_{x_i + W}$

$$B_i = \{x_i + W, T^+(x_i + W), \ldots, (T^+)^{dn_i - 1}(x_i + W)\}$$

Since

$$V/W = \bigoplus_{i=2}^{k} Z_{y_i + W} = \bigoplus_{i=2}^{k} Z_{x_i + W}$$

Then $\bigcup_{i=2}^{k} B_i$ is a basis of V/W. Then, $\bigcup_{i=2}^{k} A_i$ is a basis of $V \Rightarrow V = \bigoplus_{i=1}^{k} Z_x$. □

Corollary 2.4.7. *With the above notation, relative to the basis $\bigcup_{i=1}^{k} A_i$ the matrix of f is of the form*

$$\bigoplus_{i=1}^{k} C_i = \begin{pmatrix} C_1 & & & \\ & C_2 & & \\ & & \ddots & \\ & & & C_k \end{pmatrix}$$

Corollary 2.4.8. *dim $V = (n_1 + \cdots + n_k)$ deg p*

We may assume that the T-cyclic vectors $x_1, \ldots, x_k$ are arranged in descending order, i.e.,

$$t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$$

Theorem 2.4.9. *The integers $n_1, \ldots, n_k$ are uniquely determined by T.*

Proof. We know

$$\dim Z_{x_i} = \deg m_{x_i} = \deg p^{n_i} = dn_i \ \forall\ i$$

Also, $\forall\ j$, the image of Z_{x_i} under $p(T)^j$ is the T-cyclic subspace $Z_{p(T)^j(x_i)}$. As the T-annihilator of x_i is p^{n_i},

$$\dim Z_{p(T)^j(x_i)} = \begin{cases} 0 & \text{if } j \geq n_i \\ d(n_i - j) & \text{if } j < n_i \end{cases}$$

We can uniquely write every $x \in V$ as

$$x = v_1 + \cdots + v_k$$

where $v_i \in Z_{x_i}$ So, every element of Im $p(T)^j$ can be written uniquely as

$$p(T)^j(x) = p(T)^j(v_1) + \cdots + p(T)^j(v_k)$$

Now, if $r \in \mathbb{Z}$ such that $n_1, \ldots, n_r > j$ and $n_{r+1} \leq j$, then

$$\text{Im } p(T)^j = \bigoplus_{i=1}^{r} Z_{p(T)^j(x_i)}$$

$$\Rightarrow \dim \text{Im } p(T)^j = d\sum_{i=1}^{r}(n_i - j) = d\sum_{n_i > j}(n_i - j)$$

Then

$$\begin{aligned} \dim \text{Im } p(T)^{j-1} - \dim \text{Im } p(T)^j &= d\big(\sum_{n_i > j-1}(n_i - j + 1) - \sum_{n_i > j}(n_i - j)\big) \\ &= d(\{\# \text{ of } n_i \geq j\}) \end{aligned}$$

This determines the sequence

$$t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$$

completely. □

When the minimal polynomial of T is of the form p^t, where p is irreducible then, the polynomials $p^{n_1}, p^{n_2}, \ldots, p^{n_t}$ determined uniquely by $t = n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$ are called the *elementary divisors* of T.

Let us apply the above results to the case when the characteristic and minimal polynomials of a linear mapping $f : V \to V$ are

$$c_T = p_1^{d_1} p_2^{d_2} \dots p_k^{d_k} \text{ and } m_T = p_1^{e_1} p_2^{e_2} \dots p_k^{d_k}$$

where $p_1, \dots, p_k$ are distinct irreducible polynomials.

We know that by the Primary Decomposition Theorem, there is an ordered basis of V with respect to which the matrix of T is a block diagonal matrix

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$

where each A_i is the matrix representing the induced mapping T_i on $V_i = \text{ Ker } p_i(T)^{e_i}$. Now, the minimal polynomial of T_i is $p_i^{e_i}$. By Cyclic Decomposition Theorem, there is a basis of V_i with respect to which A_i is the block diagonal matrix

$$\begin{pmatrix} C_{i1} & & & \\ & C_{i2} & & \\ & & \ddots & \\ & & & C_{it} \end{pmatrix}$$

where C_{ij} are the companion matrices associated with the elementary divisors of T_i. This block diagonal form where each A_i is a block diagonal of companion matrices is unique and is called the *rational canonical matrix* of T.

Chapter 3

Symmetric groups

In this chapter, we will compute and classify the conjugacy classes and z-classes of S_n. We will show that centralizers of symmetric groups are a product of generalized symmetric groups and provide a brief introduction to wreath products. One can view generalized symmetric products as a particular case of wreath products. For classification of z-classes, we will be discussing the results from [BKS17]. By the end of this chapter, we would have gained enough understanding to prove the following theorem.

Theorem 3.0.1. *[BKS17] Suppose $n \geq 3$. Let ν be a restricted partition of $n-2$ in which 1 and 2 do not appear as its part. Let $\lambda = 1^2\nu$ and $\mu = 2^1\nu$ be partitions of n obtained by extending ν. Then, the conjugacy classes of λ and μ belong to the same z-class in S_n. Further, the converse is also true.*

Corollary 3.0.2. *[BKS17] The number of z-classes in S_n is $p(n) - \tilde{p}(n-2)$. Thus, the number of z-classes in S_n is equal to $p(n) - p(n-2) + p(n-3) + p(n-4) - p(n-5)$.*

3.1 Conjugacy classes and centralizer of elements in S_n

We have followed Conrad's notes [Con] for this section.
To understand the conjugacy classes in S_n, we begin by computing the conjugates of a k-cycle.

Theorem 3.1.1. *For any cycle $(i_1 i_2 \dots i_k) \in S_n$ and any $\sigma \in S_n$*

$$\sigma(i_1 i_2 \dots i_k)\sigma^{-1} = (\sigma(i_1)\sigma(i_2)\dots\sigma(i_k)).$$

Proof. Let $\pi = \sigma(i_1 i_2 \dots i_k)\sigma^{-1}$. We need to show

- π sends $\sigma(i_1)$ to $\sigma(i_2)$, $\sigma(i_2)$ to $\sigma(i_3), \dots$ and $\sigma(i_k)$ to $\sigma(i_1)$.
- π does not move any number other than $\sigma(i_1), \sigma(i_2), \dots, \sigma(i_k)$

For any r, $1 \leq r \leq k-1$,

$$\pi(\sigma(i_r)) = \sigma(i_1 i_2 \dots i_k)\sigma^{-1}(\sigma(i_r)) = \sigma(i_1 i_2 \dots i_k)(i_r) = \sigma(i_{r+1})$$

For $r = k$,

$$\pi(\sigma(i_r)) = \sigma(i_1 i_2 \dots i_k)(i_k) = \sigma(i_1)$$

Let $a \notin \{i_1, i_2, \dots, i_k\}$. We need to show that $\pi(a) = a$.
Since $a \neq \sigma(i_r)$, for any $1 \leq r \leq k \Rightarrow \sigma^{-1}(a) \neq i_r$
$\Rightarrow (i_1 i_2 \dots i_k)(\sigma^{-1}(a)) = \sigma^{-1}(a)$

$$\therefore\ \pi(a) = \sigma(i_1 i_2 \dots i_k)\sigma^{-1}(a) = \sigma\sigma^{-1}(a) = a$$

□

This theorem shows that conjugate of any k-cycle is a cycle of the same length. We will now prove that the converse is also true.

Theorem 3.1.2. *All cycles of the same length in S_n are conjugate.*

Proof. Let $(a_1 \dots a_k)$ and $(b_1 \dots b_k)$ be two k-cycles in S_n. We can choose $\sigma \in S_n$ to be a bijection such that $\sigma(a_i) = b_i, 1 \leq i \leq k$ and the complement of $\{a_1, \dots, a_k\}$ is arbitrarily mapped to the complement of $\{b_1, \dots, b_k\}$. By using Theorem 3.1.1, we see that conjugation by σ carries the first k-cycle to the second one. □

All elements of S_n are not k-cycles, so we ensure that each pemutation is written as a product of disjoint cycles arranged in ascending order of cycle length, including 1-cycles if there are any fixed points. This length is called the *cycle type* of the permutation.
Eg: (123)(46) in S_6 is (5)(46)(123) and has the cycle type (1,2,3).
The cycle type of a permutation in S_n is just a set of positive integers that add up to n, which is called a *partition* of n. Let us set notation clearly by defining a partition λ of n as $\lambda = \lambda_1^{e_1} \dots \lambda_r^{e_r}$ where $1 \leq \lambda_1 < \dots < \lambda_r \leq n, e_i \geq 1\ \forall\ i$ and $n = \sum_{i=1}^{r} e_i \lambda_i$. There are 11

partitions of 6.

$$6^1,\ 1^1 5^1,\ 2^1 4^1,\ 3^2,\ 1^2 4^1,\ 1^1 2^1 3^1,\ 2^3,\ 1^3 3^1,$$
$$1^2 2^2,\ 1^4 2^1,\ 1^6$$

Thus, the permutations of S_6 have 11 cycle types and the cycle type (1,2,3) is denoted by the partition $\lambda = 1^1 2^1 3^1$.

Lemma 3.1.3. *If π_1 and π_2 are disjoint permutations in S_n, then $\sigma\pi_1\sigma^{-1}$ and $\sigma\pi_2\sigma^{-1}$ are disjoint permutations for any $\sigma \in S_n$.*

Proof. $\nexists\ i$ such that $\pi_1(i) = i$ and $\pi_2(i) = i$. Suppose $\sigma\pi_1\sigma^{-1}$ and $\sigma\pi_2\sigma^{-1}$ are not disjoint.

$$\Rightarrow \exists\ i \text{ such that } \sigma\pi_1\sigma^{-1}(i) = i \text{ and } \sigma\pi_2\sigma^{-1}(i) = i$$
$$\Rightarrow \pi_1(\sigma^{-1}(i)) = \sigma^{-1}(i) \text{ and } \pi_2(\sigma^{-1}(i)) = \sigma^{-1}(i)$$

This is a contradiction. □

Theorem 3.1.4. *Two permutations in S_n are conjugate iff they have the same cycle type.*

Proof. Let $\pi \in S_n$ be written as a product of disjoint cycles. By Lemma 3.1.3, $\sigma\pi\sigma^{-1}$ will be a product of the σ-conjugates of the disjoint cycles for π and these σ-conjugates are disjoint cycles with the same respective lengths.
$\therefore\ \sigma\pi\sigma^{-1}$ has the same cycle type as π. Conversely, suppose the cycle type is $(m_1, m_2, \dots)$. Then

$$\pi_1 = (a_1 a_2 \dots a_{m_1})(a_{m_1+1} \dots a_{m_1+m_2}) \dots$$

and

$$\pi_2 = (b_1 b_2 \dots b_{m_1})(b_{m_1+1} \dots b_{m_1+m_2}) \dots$$

where the cycles are disjoint. Now define a permutation $\sigma \in S_n$ from π_1 to π_2 as $\sigma(a_i) = b_i$. Then, by Theorem 3.1.2, $\sigma\pi_1\sigma^{-1} = \pi_2$. □

Since the conjugacy class of a permutation in S_n is determined by its cycle type, which is a certain partition of n, the number of conjugacy classes in S_n is the number of partitions of n. Let $p(n)$ denote the number of partitions of n.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$p(n)$	1	2	3	5	7	11	15	22	30	42	56	77	101	135

We know that, the number of partitions of n is equal to the coefficient of x^n in the product $(1+x+\ldots)(1+x^2+\ldots)(1+x^3+\ldots)\ldots$, the generating function for $p(m)$ is

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i}$$

Further, the size of the conjugacy class represented by the partition $\lambda = \lambda_1^{e_1}\ldots\lambda_r^{e_r}$ (let us henceforth call this element σ_λ) is

$$\frac{n!}{\prod_{i=1}^{r} \lambda_i^{e_i} e_i!} \tag{3.1}$$

Let $\lambda = {\lambda_1}^{e_1}\ldots{\lambda_r}^{e_r}$ be a partition of n, $n_i = \sum_{j=1}^{i} \lambda_j e_j$ and $n_0 = 0$. We can represent the element of S_n corresponding to λ as a product of cycles. Now, we choose a representative of the class denoted by $\sigma_\lambda = \sigma_{\lambda_1}\ldots\sigma_{\lambda_i}\ldots\sigma_{\lambda_r}$ where

$$\sigma_{\lambda_i} = \underbrace{(n_{i-1}+1,\ldots,n_{i-1}+\lambda_i)\ldots(n_{i-1}+(e_i-1)\lambda_i+1,\ldots,n_{i-1}+e_i\lambda_i)}_{e_i}$$

is a product of e_i many disjoint cycles, each of length λ_i.

Now, let $\pi \in \mathcal{Z}_{S_n}(\sigma_\lambda) \Rightarrow \pi\sigma_\lambda = \sigma_\lambda\pi \Rightarrow \pi\sigma_\lambda\pi^{-1} = \sigma_\lambda$.
By Theorem 3.1.1 and Lemma 3.1.3, we know that $\pi\sigma_\lambda\pi^{-1} = \pi\sigma_{\lambda_1}\pi^{-1}\ldots\pi\sigma_{\lambda_2}\pi^{-1}$. Theorem 3.1.4 further tells us that $\pi\sigma_{\lambda_i}\pi^{-1}$ must have the same cycle type as σ_{λ_i}, $\forall\ i$.
$\Rightarrow \pi\sigma_{\lambda_i}\pi^{-1} = \sigma_{\lambda_i}\ \forall\ i$. But we can allow permutations between various e_i disjoint cycles which constitute σ_{λ_i}, i.e., if we were to describe σ_{λ_i} as $(1\ldots e_i) \in S_{e_i}$ where each element in this permutation corresponding to a disjoint cycle of length λ_i, then

$$\pi\sigma_{\lambda_i}\pi^{-1} = (\pi(1)\ldots\pi(e_i)) = (1\ldots e_i)$$

$$\therefore\ \pi(i) \equiv (i+k) \bmod e_i \tag{3.2}$$

We will describe the centralizer in more detail in Section 3.3, but we can compute the size of $\mathcal{Z}_{S_n}(\sigma_\lambda)$.
We know that, for each $g \in G$, its conjugacy class has the same size as the index of its centralizer,

$$|\{xgx^{-1}|\ x \in G\}| = [G : \mathcal{Z}_G(g)]$$

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\tilde{p}(m)$	0	0	1	1	1	2	2	3	4	5	6	9	10	13

In this case, $G = S_n$ and since S_n is a finite group,

$$[G : \mathcal{Z}_G(g)] = \frac{|G|}{|\mathcal{Z}_G(g)|}$$

Therefore, the size of the centralizer of the element represented by the partition λ is

$$|\mathcal{Z}_{S_n}(\sigma_\lambda)| = \prod_{i=1}^{r} \lambda_i^{e_i} e_i! \tag{3.3}$$

3.2 Restricted partitions

Partitions are significant due to their one-one correspondence with conjugacy classes of the symmetric groups S_m. The partitions which are obtained by putting some conditions are called *restricted partitions*.
A partition of m is $\lambda = m_1^{e_1} \dots m_r^{e_r}$ where $1 \leq m_1 < \cdots < m_r \leq m$, $e_i \geq 1 \ \forall \ i$ and $m = \sum_{i=1}^{r} e_i m_i$. This is also denoted as $\lambda \vdash m$ or $m_1^{e_1} \dots m_r^{e_r} \vdash m$.
Let $\tilde{p}(m)$ be the number of partitions of m in which 1 and 2 do not appear as its part. Then,

$$\tilde{p}(m) = |\{\lambda = m_1^{e_1} \dots m_r^{e_r} \vdash m \mid m_1 \geq 3\}|$$

The generating function for $\tilde{p}(m)$ is

$$\prod_{i\geq 3} \frac{1}{1-x^i}$$

Now,

$$p(m) = \prod_{i\geq 1} \frac{1}{1-x^i} = \left(\frac{1}{1-x}\right)\left(\frac{1}{1-x^2}\right)\tilde{p}(m)$$

$$\Rightarrow (1-x)(1-x^2)p(m) = (1-x-x^2+x^3)p(m) = p(m) - xp(m) - x^2p(m) + x^3p(m)$$

We have the formula to compute $\tilde{p}(m)$ in terms of partition function $p(m)$ is

$$\tilde{p}(m) = p(m) - p(m-1) - p(m-2) + p(m-3)$$

3.3 Wreath products

The theory for wreath products that we use can be found in [JK84]. Let G be a group and H a subgroup of S_n. We denote by G^n the set of all mappings from $\mathbf{n} = \{1, \ldots, n\}$ into G:

$$G^{\mathbf{n}} := \{f|\ f : \mathbf{n} \to G\}$$

For $f \in G^{\mathbf{n}}$ and $\pi \in H$, we define $f_\pi \in G^{\mathbf{n}}$ by

$$f_\pi := f \circ \pi^{-1}$$

The wreath product of G by H is a group defined by

$$G \wr H := G^n \times H = \{(f;\pi)|\ f : \mathbf{n} \to G \text{ and } \pi \in H\}$$

together with the composition map defined by

$$(f;\pi)(f';\pi') := (f f'_\pi; \pi\pi')$$

The order is

$$|G \wr H| = |G|^n |H|$$

If we define $e \in G^{\mathbf{n}}$ by

$$e(i) := 1_G,\ i \in \mathbf{n}$$

and for $f \in G^{\mathbf{n}}$ the mapping $f^{-1} \in G^{\mathbf{n}}$ by

$$f^{-1}(i) := f(i)^{-1},\ i \in \mathbf{n},$$

then we obtain for the identity element in $G \wr H$ and for the inverse of $(f;\pi) \in G \wr H$:

$$1_{G \wr H} = (e; 1_H) \text{ and } (f;\pi)^{-1} = (f^{-1}_{\pi^{-1}}; \pi^{-1}),$$

where

$$f^{-1}_{\pi^{-1}} := (f_{\pi^{-1}})^{-1} = (f^{-1})_{\pi^{-1}}.$$

Let us define G^*, called the *base group* of $G \wr H$, which is a normal subgroup of $G \wr H$ as

$$G^* := \{(f; 1_H)|\ f \in G^{\mathbf{n}}\}$$

It is the direct product of n copies G_i of G, where

$$G_i := \{(f; 1_H) | \; \forall\, j \neq i, f(j) = 1_G\} \cong G$$

The subgroup

$$H' := \{(e; \pi) | \pi \in H\} \cong H$$

is a complement of G^*, so we have

(i) $G \wr H = G^* \cdot H'$,

(ii) $G^* = \prod_i G_i$,

(iii) $G^* \cap H' = \{1_{G \wr H}\} = \{(e; 1_H)\}$.

If G is a permutation group of finite degree, say $G \leq S_m$, then we obtain a permutation representation ψ of $G \wr H$ as follows:

$$\begin{aligned} \psi : G \wr H &\to S_{mn} \\ (f; \pi) &\mapsto \begin{pmatrix} (j-1)m + i \\ (\pi(j) - 1)m + f(\pi(j))(i) \end{pmatrix}_{1 \leq i \leq m, 1 \leq j \leq n} \end{aligned}$$

Here, $\psi[G_1]$ acts on $\{1, \ldots, m\} \subseteq \{1, \ldots, mn\}$ the same way G acts on $\{1, \ldots, m\}$, and the restriction of $\psi[G_1]$ to $\{1, \ldots, m\}$ is just G. Similarly, $\psi[G_2]$ acts on $\{m+1, \ldots, 2m\}$ the same way G acts on $\mathbf{m}$ and so on. Also $\psi[H']$ permutes the subsets $\{1, \ldots, m\}, \{m + 1, \ldots, 2m\}, \ldots, \{(n-1)m + 1, \ldots, mn\}$ in the same way as H acts on $\{1, \ldots, n\}$.

Now if we consider $C_m := \langle (1 \ldots m) \rangle \leq S_m$, then $\psi[C_m \wr S_n]$ is the centralizer of the permutation

$$(1, \ldots, m)(m + 1, \ldots, 2m) \ldots ((n - 1)m + 1, \ldots, nm) \in S_{nm}$$

as $\psi[C_m \wr S_n]$ would permute both various elements in a disjoint cycle and elements across disjoint cycles and we can see from Theorem 3.1.1 and Eqns. 3.2 and 3.3. This group $S(m, n) = C_m \wr S_n$ is called a *generalised symmetric group*. Here C_m is a cyclic group and S_n is a symmetric group.

We can also see from Theorem 3.1.2 and Eqns. 3.2 and 3.3 that the centralizer of a general element $\sigma_\lambda \in S_n$ is

$$\mathcal{Z}_{S_n}(\lambda) := \mathcal{Z}_{S_n}(\sigma_\lambda) \cong \prod_{i=1}^{r} C_{\lambda_i} \wr S_{e_i} \tag{3.4}$$

3.4 z-classes in S_n

For $n = 3$ and 4, the conjugacy classes and z-classes are same. So, we may assume $n \geq 5$ in this section. Eqn. 3.4 gives the formula for the centralizer of an element σ_λ. Thus, the center of $\mathcal{Z}_{S_n}(\sigma_\lambda)$ is

$$Z_\lambda = \mathcal{Z}(\mathcal{Z}_{S_n}(\sigma_\lambda)) = \begin{cases} \prod_{i=1}^{r} \langle \sigma_{\lambda_i} \rangle & \text{if } \lambda_i^{e_i} \neq 1^2 \\ \langle (1,2) \rangle \times \prod_{i=2}^{r} \langle \sigma_{\lambda_i} \rangle & \text{when } \lambda_i^{e_i} = 1^2 \end{cases}$$

This is because $C_1 \wr S_2 \cong S_2 = \{e, (1,2)\}$. Therefore, it permutes representatives 1 and 2 as a 2-cycle.

Lemma 3.4.1. *Let $\lambda = {\lambda_1}^{e_1} \ldots {\lambda_r}^{e_r}$ be a partition of n. Then $\mathcal{Z}_{S_n}(\lambda)$ determines r uniquely.*

Proof. The natural action of $\mathcal{Z}_{S_n}(\lambda) = \prod_{i=1}^{r} C_{\lambda_i} \wr S_{e_i}$ on the set $\{1, \ldots, n\}$ as a subgroup of S_n and each $C_{\lambda_i} \wr S_{e_i}$ permutes just the set $\{n_{i-1}+1, \ldots, n_{i-1}+\lambda_i\} \ldots \{n_{i-1}+(e_i-1)\lambda_i + 1, \ldots, n_{i-1}+e_i\lambda_i\}$ and the elements within this set but not between cycles of different cycles. Thus, we obtain the following orbits: $\{\{1, \ldots, n_1\}, \{n_1+1, \ldots, n_2\}, \ldots\}$ and there are r such orbits. $\square$

Lemma 3.4.2. *Let $\lambda = {\lambda_1}^{e_1} \ldots {\lambda_r}^{e_r}$ be a partition of n and ${\lambda_1}^{e_1} = 1^2$. Let Z_λ be the center of $\mathcal{Z}_{S_n}(\lambda)$. Then Z_λ determines the partition λ uniquely.*

Proof. When we consider the action of Z_λ on the set $\{1, \ldots, n\}$, Z_λ acts on each of the r orbits obtained in 3.4.1 and determines elements in all e_i cycles of each σ_{λ_i}. Thus, the orbits will be of size λ_i and each of them occurs e_i times which determines the partition λ. $\square$

Proposition 3.4.3. *Let $\lambda = {\lambda_1}^{e_1} \ldots {\lambda_r}^{e_r}$ and $\mu = {\mu_1}^{f_1} {\mu_2}^{f_2} \ldots {\mu_s}^{f_s}$ be a partitions of n. Then $\mathcal{Z}_{S_n}(\lambda)$ is conjugate to $\mathcal{Z}_{S_n}(\mu)$ iff*

1. *$r = s$,*
2. *for all $i \geq 2, \lambda_i$ and μ_i are ≥ 3 and ${\lambda_i}^{e_i}$ and ${\mu_i}^{f_i}$,*
3. *${\lambda_1}^{e_1} = 1^2$ and ${\mu_1}^{f_1} = 2^1$ or vice versa.*

Proof. Suppose $\lambda = 1^2\nu$ and $\mu = 2^1\nu$ where $\nu = \nu_1^{l_1} \ldots \nu_k^{l_k}$ is a partition of $n-2$ with $\nu_1 > 2$. Then the representative elements of the conjugacy classes are $\sigma_\mu = (12)\sigma_\nu$ and $\sigma_\lambda = \sigma_\nu$, where σ_μ has cycles of length > 2. Then, the centralizers for these two elements are the

same.
For the converse, let us choose representative elements σ_λ and σ_μ such that $\mathcal{Z}_{S_n}(\sigma_\lambda)$ and $\mathcal{Z}_{S_n}(\sigma_\mu)$ are conjugates. Lemma 3.4.1 implies $r = s$. Now, let us take the center of both these groups Z_λ and Z_μ and make them act on the set $\{1, \ldots, n\}$. If $\lambda_1^{e_1}$ and $\mu_1^{f_1}$ both are not 1^2, then Lemma 3.4.2 determine the partitions λ and μ. But, since λ and μ are different partitions, the only possibility would be ${\lambda_1}^{e_1} = 1^2$ and ${\mu_1}^{f_1} = 2^1$ or vice versa, which proves our result. □

This proves Theorem 3.0.1.

Chapter 4

Dynamical types and z-classes in groups

The aim of the chapter is to discuss the results published in the paper titled "Dynamical types and conjugacy classes in groups" [SK07]. In this paper, Kulkarni attempts to relate "dynamical types" to z-classes, without explicitly defining them. We will similarly use our understanding of "dynamical types", which is derived from human perception. We know that these "dynamical types" are associated with transformations in classical geometries. There are infinitely many transformations, equipped with natural spatial and numerical invariants and their "dynamical types" are finite in number.

4.1 The α- and σ- fibrations

Let G be a group acting on a set X. For $x \in X$, let $G(x)$ be the G-orbit of x and G_x the stabilizer subgroup of G at x.

$$
\begin{aligned}
G(x) &= \{y \in X | \; y = g \cdot x\} \\
G_x &= \{g \in G | \; g \cdot x = x\}
\end{aligned}
$$

Since, $G(x) = G(y)$ or $G(x) \cap G(y) = \emptyset$, $\forall \; x, y \in G$ we get the first partition

$$X = \bigcup_{x \in X} G(x)$$

If $y \in G(x) \Rightarrow y = g \cdot x$, for some $g \in G$ and $G_y = gG_xg^{-1}$. Thus, the conjugacy class $[G_x]$ of the point-stabilizers in a G-orbit is well defined. Two elements $x, y \in X$ are in the same *orbit-class* if $G_x \sim G_y$ and we shall denote this equivalence relation $x \sim_o y$. Let $R(x)$ be the equivalence class of x with respect to $\sim_o$. We, thus, obtain the second partition of X.

$$X = \bigcup_{G_x \sim G_y} R(x)$$

We will now provide a description of $R(x)$ in two different ways.
Let F_x be the set of fixed point of G_x

$$F_x = \{y \in X |\ G_y \supset G_x\},$$

and F'_x be the set of "generic" elements in X

$$F'_x = \{y \in X |\ G_y = G_x\}.$$

Let N_x be the normalizer of G_x in G,

$$N_x = \{g \in G |\ gG_xg^{-1} = G_x\}.$$

Finally, let W_x be the Weyl group at x,

$$W_x = N_x/G_x.$$

We now define a canonical free action of W_x on G/G_x. Let $[n]$ denote the class of $n \in N_x$ in W_x.

$$\begin{aligned} W_x \times G/G_x &\to G/G_x \\ ([n], gG_x) &\mapsto [n] \bullet (gG_x) = gG_xn^{-1} \end{aligned}$$

Since $n \in N_x \Rightarrow n^{-1}G_x = G_xn^{-1} \Rightarrow g(G_xn^{-1}) = g(n^{-1}G_x)$.

Proposition 4.1.1. *This action is free, i.e.,*

$$\forall\ n \in N_x,\ \ \textit{if}\ [n] \bullet (gG_x) = gG_x \Rightarrow n \in G_x$$

Proof. $[n] \bullet (gG_x) = gn^{-1}G_x = gG_x \Rightarrow n^{-1}G_x = G_x \Rightarrow n^{-1} \in G_x$

$\Rightarrow n \in G_x$. $\square$

Let us now define a second canonical free action of W_x on F'_x.

(i) $gF_x = F_{gx}$.

$$\begin{aligned} F_{gx} &= \{y \in X|\ G_{gx}y = y\} = \{y \in X|\ gG_xg^{-1}y = y\} \\ &= \{y \in X|\ G_xg^{-1}y = g^{-1}y\} = \{y \in X|\ g^{-1}y \in F_x\} \\ &= gF_x \end{aligned}$$

(ii) N_x leaves F_x invariant. Let $n \in N_x$, then

$$\begin{aligned} nF_x = F_{nx} &= \{y \in X|\ G_{nx}y = y\} \\ &= \{y \in X|\ nG_xn^{-1}y = y\} \\ &= \{y \in X|\ G_xy = y\} = F_x \end{aligned}$$

(iii) N_x leaves F'_x invariant.
Let $y \in F'_x \Rightarrow G_y = G_x$.
For $n \in N_x$, $G_{ny} = nG_yn^{-1} = nG_xn^{-1} = G_x \Rightarrow nF'_x = F'_x$.

(iv) N_x acts on F_x via W_x
For $n \in N_x$ and $[n]$, its class in W_x, $[n]F_x = nG_xF_x = nF_x$

(v) W_x acts freely on F'_x. If $n \in N_x$, $y \in F'_x$, and $[n]y = ny = y$, then $n \in G_y = G_x$.

We will now discuss the main result here.

Theorem 4.1.2. *[Kul07] Let G act on X. Consider the diagonal action of W_x on $G/G_x \times F'_x$. Then the map*

$$\begin{aligned} \phi : G/G_x \times F'_x &\to R(x) \\ (gG_x, y) &\mapsto gy \end{aligned}$$

is well-defined, and induces a bijection,

$$\bar{\phi} : \{G/G_x \times F'_x\}/W_x \to R(x)$$

Proof. (i) ϕ is well defined.

Let $g \in G,\ y \in F'_x$. We need to show that $gy \in R(x)$.

$$G_{gy} = gG_y g^{-1} = gG_x g^{-1} \Rightarrow gy \sim_o x \Rightarrow gy \in R(x)$$

We also need to show that for $g \in G$, $u \in G_x$ and $y \in F'_x$, $guy = gy$ and this is true because $G_y = G_x = uG_x$

(ii) ϕ is surjective. Let $y \in R(x)$.
$\exists\ g \in G$ such that $G_y = gG_x g^{-1} = G_{gx}$.
So, $y \in F'_{gx} = gF'_x \Rightarrow g^{-1}y \in F'_x$.

$$\phi(gG_x, g^{-1}y) = y \Rightarrow y \in \text{ im } (\phi)$$

(iii) ϕ is constant on the W_x-orbits on $G/G_x \times F'_x$.
Let $[n] \in W_x$.

$$\phi([n]gG_x, [n]y) = \phi(gn^{-1}G_x, ny) = gy = \phi(gG_x, y)$$

So, we have an induced surjective map $\bar{\phi}$.

(iv) $\bar{\phi}$ is injective.
Let $y, z \in F'_x$ and we have $\phi(gG_x, y) = \phi(hG_x, z) \Rightarrow gy = hz$.
Let $u = h^{-1}g \Rightarrow z = uy \Rightarrow G_z = G_x = G_{uy} = uG_y u^{-1} = uG_x u^{-1}$
$u \in N_x$
So, $[u] \in W_x$ and we have

$$[u](gG_x, y) = (gu^{-1}G_x, uy) = (hG_x, z)$$

$\Rightarrow (gG_x, y)$ and (hG_x, z) are in the same W_x-orbits.

$\square$

By projecting the second factor, the map $\bar{\phi}$ induces the map

$$\alpha : R(x) \to F'_x/W_x$$

whose "fiber" is G/G_x. Now, $G/G_x \cong G(x)$ up to a natural equivalence of G-actions, so this is just a representation of the orbit-class $R(x)$ as the union of orbits. We can consider this map α as a one-one parametrization of the orbits by F'_x/W_x. Further, W_x acts freely on F'_x.

So F'_x/W_x can be identified with a subset of F'_x which picks up one point in each W_x-orbit. So, this map α provides a "numerical" invariant for elements in $R(x)$.
Similarly, $\{G/G_x\}/W_x \cong G/N_x$. So, $\bar{\phi}$ induces a map

$$\sigma : R(x) \to G/N_x$$

whose "fiber" is F'_x. So, this map σ provides a "spatial" invariant for elements in $R(x)$.

4.2 The critical abelian subgroups and z-classes in a group

We will now consider a particular action of an arbitrary group G, conjugation on itself. Orbits in this case are the conjugacy classes of elements in G.

$$\text{Point stabilizer, } G_x = \{g \in G |\ g \cdot x = gxg^{-1} = x\}$$

This is exactly the centralizer $\mathcal{Z}_G(x)$. The fixed point set of $\mathcal{Z}_G(x)$ is the center of $\mathcal{Z}_G(x)$, denoted by $S(x)$.

$$S(x) = F_x = \{k \in \mathcal{Z}(x) |\ g \cdot k = k \Rightarrow gkg^{-1} = k \ \ \forall\ g \in G\}$$

Let us consider the set $\mathcal{S}$ of all subgroups of G and the set $\mathcal{A}$ of all abelian subgroups of G and the two canonical maps

$$\phi : \mathcal{S} \to \mathcal{A} \text{ and } \psi : \mathcal{A} \to \mathcal{S} \tag{4.1}$$

where, ϕ associates a subgroup H to its center and ψ associates an abelian group A to its centralizer.

$$\text{im}(\phi \circ \psi) := \mathcal{A}_o \text{ and } \text{im}(\psi \circ \phi) := \mathcal{S}_o$$

We will call the elements of $\mathcal{A}_o$ the *critical* abelian subgroups of G.

Proposition 4.2.1. *[SK07]*

(i) The maps ϕ and ψ restricted to $\mathcal{S}_o$ and $\mathcal{A}_o$ respectively are bijections onto $\mathcal{A}_o$ and $\mathcal{S}_o$ respectively and they are inverses of each other.

(ii) Let A be a critical abelian subgroup of a group G. Let $\mathcal{Z}_G(A)$ and $N_G(A)$ be the centralizer and normalizer of A in G. Then the normalizer of $\mathcal{Z}_G(A)$ in G equals N.

Proof. (i) Let A be an abelian subgroup of G, $\mathcal{Z}_G(A)$ be the centralizer of A and S be the center of $\mathcal{Z}_G(A)$. $\mathcal{Z}_G(A)$ consists of elements in G which commute with all elements of A. Since A is abelian $\Rightarrow A \subseteq \mathcal{Z}_G(A)$ and $A \subseteq S$. Further, the centralizer of S is contained in the centralizer of A, i.e., $\mathcal{Z}_G(A) \subseteq \mathcal{Z}_G(S)$. Let $g \in \mathcal{Z}_G(A)$

$$\begin{aligned}
&\Rightarrow ga = ag \ \forall \ a \in A \\
&\Rightarrow \ \forall \ s \in S, gs = sg \\
&\Rightarrow g \in \mathcal{Z}_G(S) \\
&\Rightarrow \mathcal{Z}_G(S) \subseteq \mathcal{Z}_G(A)
\end{aligned}$$

We now show that if S is the center of $\mathcal{Z}_G(A)$, then $\mathcal{Z}_G(A)$ is the centralizer of S. Let the centralizer of S be Z'.

$$Z' = \{g \in G| \ gs = sg \ \forall \ s \in S\}$$

Let $z \in Z'$.

$$\begin{aligned}
&za = az \ \forall \ a \in S \Rightarrow za = az \ \forall \ a \in A \\
&\Rightarrow z \in \mathcal{Z}_G(A) \Rightarrow Z' \subseteq \mathcal{Z}_G(A)
\end{aligned}$$

Let $g \in \mathcal{Z}_G(A) \Rightarrow ga = ag \ \forall \ a \in A$. Now for any $s \in S$, $sg = gs$.

$$\begin{aligned}
&\Rightarrow g \in Z' \Rightarrow \mathcal{Z}_G(A) \subseteq Z' \\
&\therefore \ \mathcal{Z}_G(A) = Z'
\end{aligned}$$

By construction, $S = \phi \circ \psi(A)$, so S is a critical abelian subgroup of G.

$A \xrightarrow{\psi} \mathcal{Z}_G(A) \xrightarrow{\phi} S$

$S \xrightarrow{\psi} \mathcal{Z}_G(A) \xrightarrow{\phi} S$

$S = \phi \circ \psi(S)$.

Let H be a subgroup, S its center and $\mathcal{Z}_G(A)$ be the centralizer of S.

$H \xrightarrow{\phi} S \xrightarrow{\psi} \mathcal{Z}_G(A)$

$Z \xrightarrow{\phi} S \xrightarrow{\psi} \mathcal{Z}_G(A)$

$\mathcal{Z}_G(A) = \psi \circ \phi(\mathcal{Z}_G(A))$

(ii) If A is a critical abelian subgroup of G, then it is the center of $\mathcal{Z}_G(A)$. Let $N_G(A)$ be the normalizer of A in G and $N(\mathcal{Z})$ be the normalizer of $\mathcal{Z}_G(A)$ in G. We need to show $N_G(A) = N(\mathcal{Z})$.
Let $n \in N(\mathcal{Z}) \Rightarrow n\mathcal{Z}_G(A)n^{-1} = \mathcal{Z}_G(A)$
If i_n denotes conjugation by n, then i_n leaves $\mathcal{Z}_G(A)$ invariant. Let $z \in \mathcal{Z}_G(A) \Rightarrow \exists y \in \mathcal{Z}_G(A)$ such that $z = nyn^{-1} \Rightarrow zn = ny \Rightarrow n^{-1}z = yn^{-1}$.

$$\begin{aligned}
&z(nan^{-1}) = nyan^{-1} = nayn^{-1} = nan^{-1}z \text{ for any } z \in \mathcal{Z}_G(A) \\
&\Rightarrow nan^{-1} \in A \\
&\Rightarrow n \in N_G(A) \Rightarrow N(\mathcal{Z}) \subseteq N_G(A)
\end{aligned}$$

Let $n \in N_G(A)$, $z \in \mathcal{Z}_G(A) \Rightarrow nA = An$, $az = za$

$$\begin{aligned}
nzn^{-1} &= = n(aza^{-1})n^{-1} \\
&= (nan^{-1})(nzn^{-1})(na^{-1}n^{-1})
\end{aligned}$$

We know that, $\exists\ b \in A$ such that $nan^{-1} = b$.

$$\begin{aligned}
&\Rightarrow nzn^{-1} = b(nzn^{-1})b^{-1} \\
&\Rightarrow (nzn^{-1})b = b(nzn^{-1}) \ \forall\ b \in A \\
&nzn^{-1} \in \mathcal{Z}_G(A) \\
&nzn^{-1} \subseteq \mathcal{Z}_G(A) \Rightarrow n \in N(\mathcal{Z}) \\
&N_G(A) \subseteq N(\mathcal{Z})
\end{aligned}$$

$\therefore\ N_G(A) = N(\mathcal{Z})$

□

For an abelian subgroup A of G, let $\mathcal{Z}_G(A)$, $N_G(A)$ be the centralizer and normalizer of A. $\mathcal{Z}_G(A)$ is a normal subgroup of $N_G(A)$. Let us now define Weyl group of A, $W(A)$.

$$W(A) = N_G(A)/\mathcal{Z}_G(A)$$

Weyl group of an abelian group A coincides with the Weyl group of the critical subgroup canonically associated to A.

Let H be the critical subgroup canonically associated to A.

$$A \xrightarrow{\quad\psi\quad} \mathcal{Z}_G(A) \xrightarrow{\quad\phi\quad} S(\mathcal{Z}_G(A)) = H$$

We know that, $N(\mathcal{Z}_G(H)) = N_G(H)$ by Theorem 4.2.1.

$$W(H) = N_G(H)/\mathcal{Z}_G(H) = N_G(\mathcal{Z}_G(H))/\mathcal{Z}_G(H)$$

H is the center of $\mathcal{Z}_G(A)$, then $\mathcal{Z}_G(A)$ is the centralizer of $H \Rightarrow \mathcal{Z}_G(H) = \mathcal{Z}_G(A)$

$$N_G(H)/\mathcal{Z}_G(H) = N(\mathcal{Z}_G(H))/\mathcal{Z}_G(H) = N(\mathcal{Z}_G(A))/\mathcal{Z}_G(A)$$

Since A is abelian, $\mathcal{Z}_G(A) = A \Rightarrow W(H) = N_G(A)/\mathcal{Z}_G(A) = W(A)$.

Given $x, y \in G$, $x \sim y$ means that x, y are conjugates in G and $x \sim_o y$ means that their centralizers are conjugates in G. Let $C(x)$ and $R(x)$ be the centralizer class of x in G and conjugacy class of x in G respectively. Let $\mathcal{Z}(x)$ denote the centralizer of x, $S(x)$ be the center of $\mathcal{Z}(x)$ and $N(x)$ be the normalizer of $\mathcal{Z}(x)$. Then $W(x) = N(x)/\mathcal{Z}(x)$ is the Weyl group at x.
Let $S'(x) = \{y \in S(x) \mid \mathcal{Z}(y) = \mathcal{Z}(x)\}$ denote the "generic" elements of $S(x)$.
$W(x)$ acts freely on $S'(x)$.
Let $n\mathcal{Z}(x) \in W(x)$ and $y \in S'(x)$.

$$n\mathcal{Z}(x) \cdot y = nyn^{-1} = y \Rightarrow ny = yn \Rightarrow n \in \mathcal{Z}(y) = \mathcal{Z}(x)$$

Since $n \in \mathcal{Z}(x) \Rightarrow n\mathcal{Z}(x) \in \mathcal{Z}(x)$.
$\therefore$ $n\mathcal{Z}(x)$ is the identity.
Since $W(x)$ acts freely on $S'(x)$, we can consider $S'(x)/W(x)$ as a subset of $S'(x)$ which picks up one point each in $W(x)$-orbit in $S'(x)$. Using the terminology described in Section 4.1, we have,

$$R(x) = \bigcup_{\mathcal{Z}(y)=\mathcal{Z}(x)} C(y) = \bigcup_{S(y)=S(x)} C(y) \text{ and}$$
$$G = \bigcup_{x \in X} R(x)$$

Theorem 4.1.2 allows us to construct two fibrations α and σ, where α is the numerical invariant and σ, the spatial invariant.

$$\alpha : R(x) \to S'(x)/W(x) \text{ and } \sigma : R(x) \to G/N(x)$$

Thus, z-classes in a group G are in a one-one correspondence with the conjugacy classes of critical abelian subgroups associated to cyclic subgroups.
Let G be a group and H its normal subgroup. For $x \in H$, let $C_G(x)$ and $C_H(x)$ be the conjugacy classes of x, $R_G(x)$, $R_H(x)$ be the z-classes of x and $\mathcal{Z}_G(x)$ and $\mathcal{Z}_H(x)$ be the centralizers of x.

(i) $C_G(x) \subset H$
Let $y \in C_G(x) \Rightarrow \exists\, g \in G$ such that $gyg^{-1} = x$.
For any $p \in G$, $py = pg^{-1}xg = zpg^{-1}g = zp \Rightarrow y \in H$

(ii) $\mathcal{Z}_H(x) = H \cap \mathcal{Z}_G(x)$
$\mathcal{Z}_H(x) = \{y \in H \mid yx = xy\}$. Clearly, $\mathcal{Z}_H(x) \subseteq H$ and since $H \subseteq G$, $y \in \mathcal{Z}_G(x)$.
$\Rightarrow \mathcal{Z}_H(x) \subseteq H \cap \mathcal{Z}_G(x)$.
Let $y \in H \cap \mathcal{Z}_G(x) \Rightarrow y \in H$ and $yx = xy \Rightarrow y \in \mathcal{Z}_H(x)$.

(iii) $C_G(x) \cong G/\mathcal{Z}_G(x)$.

$$\begin{aligned} \phi : G &\to C_G(x) \\ g &\mapsto gxg^{-1} \end{aligned}$$

Clearly, ϕ is surjective.
Now, $\ker\phi = \{y \in G \mid \phi(y) = x\}$.

$$\begin{aligned} \phi(y) = yxy^{-1} \Rightarrow & yx = xy \Rightarrow y \in \mathcal{Z}_G(x) \\ & G/\mathcal{Z}_G(x) \cong C_G(x) \end{aligned}$$

We see that $C_G(x)$ splits into H-conjugacy classes.
Finally, we can see that if x, y are two elements of H which are in the same z-class in G, then they are in the same z-class in H and $R_G(x) \cap H$ is contained in $R_H(x)$.

Let x, y are in the same z-class in G.

$$\begin{aligned}
&\exists\ g \in G \text{ such that } \mathcal{Z}_G(x) = g\mathcal{Z}_G(y)g^{-1} \\
&\mathcal{Z}_H(x) = H \cap \mathcal{Z}_G(x) \text{ and } \mathcal{Z}_H(y) = H \cap \mathcal{Z}_G(y) \\
&\Rightarrow \mathcal{Z}_H(x) = g\mathcal{Z}_G(x)g^{-1}
\end{aligned}$$

x, y are in the same z-class in H.
Let $g \in R_G(x) \cap H \Rightarrow g \in R_G(x)$ and $g \in H$. $g \in R_G(x) \Rightarrow \mathcal{Z}_G(x) = h\mathcal{Z}_G(g)h^{-1}$ for some $h \in G$.
Now, x, g are in the same z-class in G $\Rightarrow x, g$ are in the same z-class in H.
$\Rightarrow R_G(x) \cap H \subset R_H(x)$.

4.3 z-classes of semisimple linear operators

Let $\mathbb{F}$ be a field and V a vector space over $\mathbb{F}$. Let $X = L(V)$, which is the set of all linear maps from V to V. It contains $G = GL(V)$, the subset which consists of the invertibe elements of $L(V)$. Then X is an $\mathbb{F}$-algebra and G is a group.
Let G act on X by conjugation. The G-orbit of A in X is called the *similarity class* of A. If $A \in G$, then its similarity class is precisely its conjugacy class of A in G.
For $A \in X$, let $\mathcal{Z}_G(A)$ be the set of all elements B in X which commute with A. Let $\mathcal{Z}_G^*(A) = \mathcal{Z}_G(A) \cap G$. $\mathcal{Z}_G(A)$ is an $\mathbb{F}$-subalgebra of X and in case $A \in G$, $\mathcal{Z}_G^*(A)$ is the centralizer of A. Let $S(A)$ be the center of $\mathcal{Z}_G(A)$ and $S^*(A) = S(A) \cap G$. Using the definitions in Section 4.2, we can say that $S^*(A)$ is the critical subgroup associated to the cyclic subgroup generated by A.
An important invariant to study the dynamics of a linear operator A is the minimal polynomial $m_A(x)$ (non-zero). We know that $\mathbb{F}[A]$ is a $\mathbb{F}$-subalgebra of X consisting of operators which can be written as polynomials in A. Thus, there is a canonical surjective homomorphism ϕ of the polynomial ring $\mathbb{F}[x]$ onto $\mathbb{F}[A]$.

$$\begin{aligned}
\phi : \mathbb{F}[x] &\to \mathbb{F}[A] \\
p(x) &\mapsto p(A)
\end{aligned}$$

Let us assume ker ϕ is non-zero. ker ϕ is, by definition, generated by $m_A(x)$. When V is finite-dimensional, ker ϕ is automatically non-zero as X is finite-dimensional.

Lemma 4.3.1. *Let $A : V \to V$ be an operator admitting a minimal polynomial $m_A(x)$ of degree at least 1. Suppose that $m_A(x)$ is irreducible. Then either $m_A(x)$ is x and A is zero or A is invertible.*
In case A is invertible, V admits a structure of a vector space over a simple extension field $\mathbb{E}$ of $\mathbb{F}$ with respect to which the action of A is equivalent to the scalar multiplication by a primitive element of $\mathbb{E}$ over $\mathbb{F}$.

Proof. Let $\mathbb{E} = \mathbb{F}/(m_A(x))$. Since $m_A(x)$ is irreducible, then $\mathbb{E}$ is a field. Further, $\mathbb{F}[A] \cong \mathbb{E}$.
Case 1: x divides $m_A(x)$. Then, $m_A(A) = A = 0 \Rightarrow A = 0$ and $\mathbb{E} \cong \mathbb{F}$.
Case 2: x does not divide $m_A(x)$.
Let $\alpha \equiv x \bmod m_A(x) \Rightarrow \alpha - x \equiv 0 \bmod m_A(x)$, $\alpha \in \mathbb{E} \Rightarrow \alpha$ is a root of $m_A(x)$ in $\mathbb{E}$.
$\Rightarrow \mathbb{F}(\alpha) \cong \mathbb{F}[x]/(m_A(x)) = \mathbb{E}$
$\mathbb{E}$ is a simple extension of $\mathbb{F}$ and α is a primitive element of $\mathbb{E}$ over $\mathbb{F}$.
Since x does not divide $m_A(x) \Rightarrow 0$ is not an eigenvalue of the operator $A \Rightarrow A$ is invertible.
$\mathbb{F}(\alpha) \cong \mathbb{F}[A]$ and therefore, action of A is equivalent to scalar multiplication by a primitive element of $\mathbb{E}$ over $\mathbb{F}$. □

Lemma 4.3.2. *Let $A : V \to V$ be an operator. Suppose $f(x) \in \mathbb{F}[x]$ is a monic polynomial such that $f(A) = 0$. Suppose $f(x) = r(x)s(x)$ where gcd $(r(x), s(x)) = 1$. Then ker $r(A)$=im $s(A)$, ker $s(A)$=im $r(A)$ and V is a direct sum of ker $r(A)$ and ker $s(A)$.*

Proof. By Euclidean algorithm, we know $\exists\ a(x), b(x) \in \mathbb{F}[x]$ such that

$$\begin{aligned}
&a(x)r(x) + b(x)s(x) = 1\\
&a(A)r(A) + b(A)s(A) = I\\
&\forall\ \vec{v} \in V,\ \text{we have } a(A)r(A)\vec{v} + b(A)s(A)\vec{v} = \vec{v}\\
\Rightarrow& r(A)a(A)r(A) = r(A) \text{ and } s(A)b(A)s(A) = s(A)\\
\Rightarrow& r(A)(a(A)r(A) - I) = 0 \text{ and } s(A)(b(A)s(A) - I) = 0\\
\Rightarrow& a(A)r(A) - I \in \text{Ker } r(A) \text{ and } b(A)s(A) - I \in \text{Ker } s(A)\\
\Rightarrow& s(A)b(A) \in \text{Ker } r(A) \text{ and } r(A)a(A) \in \text{Ker } s(A)\\
\Rightarrow& \text{im } s(A) \subseteq \text{Ker } r(A) \text{ and im } r(A) \subseteq \text{Ker } s(A)
\end{aligned}$$

Further, im $s(A) +$ im $r(A) = V$ and ker $s(A) \cap$ ker $r(A) = 0$
ker $s(A) \subseteq$ im $r(A)$ and ker $r(A) \subseteq$ im $s(A)$

$$\text{im } s(A) = \text{ker } r(A) \text{ and im } r(A) = \text{ker } s(A)$$

Now, $n - \dim(\ker\ s(A)) + n - \dim(\ker\ r(A)) = n$
$\Rightarrow \dim(\ker\ s(A)) + \dim(\ker\ r(A)) = n$ and $\ker\ s(A) \cap \ker\ r(A) = 0$.
$\Rightarrow V = \ker\ s(A) \oplus \ker\ r(A)$. □

A consequence of Lemma 4.3.2 is that to study the dynamics of A with $m_A(x) \neq 0$, we can reduce the situation to when $m_A(x) = p(x)^d$ where $p(x)$ is an irreducible monic polynomial in $\mathbb{F}[x]$.
The only case we will consider in this chapter is $d = 1$ and $p(x) \neq x$.
Consider the extension field $\mathbb{E} = \mathbb{F}[x]/(p(x))$ and $\alpha = [x](\alpha \equiv x \bmod p(x))$. Then in the $\mathbb{E}$-structure on V, the operator A is just the scalar multiplication by α. (By Lemma 4.3.1) $\Rightarrow \mathbb{E} = \mathbb{F}[A]$. Then, the centralizer $\mathcal{Z}(A)$ is the $\mathbb{E}$-linear operators on V and the center $S(A)$ is $\mathbb{E}$. $S^*(A) = \mathbb{E}^* = \mathbb{E} - \{0\}$ is in fact the critical subgroup in G associated to A.

Theorem 4.3.3. *Let $p(x) \neq x$ be a monic irreducible polynomial in $\mathbb{F}[x]$. Let deg $p(x) = m$. If V is finite-dimensional, assume m divides dim$V$. Then*

(i) There exists A in G such that $m_A(x) = p(x)$.

(ii) An element B in G is conjugate to A iff $m_B(x) = p(x)$.

(iii) An element B in G is z-equivalent to A, iff $m_B(x)$ is irreducible and the fields $\mathbb{F}[x]/(m_A(x))$ and $\mathbb{F}[x]/(m_B(x))$ are isomorphic over $\mathbb{F}$.

Proof. (i) Let V be finite-dimensional and dim $V = n$ and $n = ml$
We need to choose subspaces V_i such that dim V_i=m. Then $V = \bigoplus_{i=1}^{l} V_i$.
Let $\mathbb{E} = \mathbb{F}[x]/(p(x))$ and h_i is $\mathbb{F}$-isomorphism of vector spaces.

$$h_i : \mathbb{E} \to V_i$$
$$a_o + a_1 x + \cdots + a_{m-1}x^{m-1} \mapsto (a_o, a_1, \ldots, a_{m-1})$$

Let $\alpha = [x]$ and consider the operator $\mu_\alpha : y \mapsto \alpha y$ on $\mathbb{E}$.
Now, when we define $A : V \to V$ such that $A \restriction_{V_i} = h_i \circ \mu_\alpha \circ h_i^{-1}$.
$\Rightarrow A \restriction_{V_i}$ is equivalent to scalar multiplication by α. Also, since operators are similar, the characteristic and minimal polynomials are the same.
$\Rightarrow m_A(x) = p(x)$

(ii) If B is conjugate to A, then $m_B(x) = m_A(x) = p(x)$. Conversely, suppose $B \in G$ such that $m_B(x) = p(x)$. We have $\mathbb{F}[B] \cong \mathbb{E}$ and $\mathbb{F}[A] \cong \mathbb{E}$ and σ_A, σ_B equip V as vector

spaces over $\mathbb{E}$ via $\mathbb{F}[A]$ and $\mathbb{F}[B]$.
Let $\{\vec{e}\}, i \in I$ be an $\mathbb{E}$-basis of V with respect to σ_A.
Let $E_i = \{\vec{e_i}, A\vec{e_i}, \ldots, A^{m-1}\vec{e_i}\}$. Let V_i be the $\mathbb{F}$-span of E_i.
Since A^m is a linear combination of $I, \ldots, A^{m-1}$, each V_i is invariant under A and $\bigcup_{i \in I} E_i$ is an $\mathbb{F}$-basis of V.
Similarly, let $\{\vec{f_j}\}, j \in J$ be an $\mathbb{E}$-basis of V with respect to σ_B.
Let $F_j = \{\vec{f_j}, B\vec{f_j}, \ldots, B^{m-1}\vec{f_j}\}$. Then, $\bigcup_{j \in J} F_j$ is also an $\mathbb{F}$-basis of V.
Since cardinality of $\mathbb{F}$-basis is well-defined, there exists a bijection between the index sets I and J. (in fact, we can take $I = J$). Let

$$\begin{aligned} h : V &\to V \\ A^k\vec{e_i} &\mapsto B^k\vec{f_i} \end{aligned}$$

$\Rightarrow$ h is well defined, $m_A(x) = m_B(x) = p(x)$ and h provides the required conjugacy between A and B.

(iii) We know that, $\mathcal{Z}_G(A)$ is $\mathbb{E}$-linear maps of V, $\mathbb{E} = \mathbb{F}[A]$ and $S(A) = \mathbb{E}$ which is a field. Let $B \in G$. By definition, A, B are z-equivalent iff $\mathcal{Z}_G^*(A)$ and $\mathcal{Z}_G^*(B)$ are conjugates. Replacing B by a suitable conjugate , we can assume $\mathcal{Z}_G^*(A) = \mathcal{Z}_G^*(B) \Rightarrow S^*(A) = S^*(B)$.
We know that, $S^*(A) \cup \{0\}$ is closed under addition and is the field $\mathbb{F}[A]$
$\Rightarrow S^*(B) \cup \{0\}$ since $S^*(B) \cup \{0\} \subseteq S(B)$.
Also, $B \in S^*(B) = S^*(A) \Rightarrow B \in \mathbb{F}[A] \Rightarrow \mathbb{F}[B] \subseteq \mathbb{F}[A]$. In particular, $\mathbb{F}[B]$ is an integral domain.
But if $m_b(x) = q(x)$, then $\mathbb{F}[B] \cong \mathbb{F}[x]/(q(x)) \Rightarrow q(x)$ must be irreducible and so $\mathbb{F}[B]$ is a field.
Suppose, if possible that $\mathbb{F}[B] = \mathbb{E}_1 \neq \mathbb{F}[A]$. Then V also has the structure of a vector space over $\mathbb{E}_1$. Then $\mathcal{Z}_G(B)$ is isomorphic to $\mathbb{E}_1$-endomorphisms of V.
$\Rightarrow S(B)$ must coincide with $\mathbb{E}_1 \Rightarrow S^*(B) = \mathbb{E}_1^* \subseteq \mathbb{E}^*$.
This contradicts $S^*(A) = S^*(B) = \mathbb{E}^*$.

$$\mathbb{F}[x]/(m_A(x)) \cong \mathbb{F}[x]/(m_B(x))$$

$\square$

We will now use this information to describe the z-classes of semi-simple elements in G.

An operator A in G is called *semisimple* iff every A-invariant subspace of V has an A-invariant complement.
Consider V to be finite-dimensional and let dim $V = n$. Then, A is semisimple iff $m_A(x)$ factors into pairwise distinct irreducible factors.
Let $m_A(x) = \prod_{i=1}^{r} p_i(x)$, where each $p_i(x) \in \mathbb{F}[x]$ is monic and irreducible.
Let $V_i = \ker\, p_i(A)$ and dim $V_i = n_i$. Then $V = \bigoplus_{i=1}^{r} V_i, V_i \cap V_j = \emptyset$.
If W is an A-invariant subspace, then $W = \bigoplus_{i=1}^{r} W \cap V_i$.
Let $\chi_A(x) = \prod_{i=1}^{r} p_i(x)^{d_i}$, which is the characteristic polynomial, is a complete invariant of the conjugacy class of A. Let $\chi_A(x) = \prod_{i=1}^{r} p_i(x)^{d_i}$.
Let deg $p_i(x) = m_i \Rightarrow n = \sum_{i=1}^{r} m_i d_i$ and let $n_i = m_i d_i$.
Let $A_i = A\restriction_{V_i}$.
$\mathcal{Z}_G(A)$ is canonically isomorphic to $\prod_{i=1}^{r} \mathcal{Z}_G(A_i)$. This is because if an element commutes with V, it must commute with each individual V_i and would therefore belong to $\mathcal{Z}_G(A_i)$ and vice versa.
Then, $S(A)$ is canonically isomorphic to $\prod_{i=1}^{r} S(A_i)$.

$$\mathcal{Z}_G^*(A) \cong \prod_{i=1}^{r} \mathcal{Z}_G^*(A_i) \text{ and } S^*(A) \cong \prod_{i=1}^{r} S^*(A_i)$$

Let $\mathbb{E}_i = \mathbb{F}[x]/(p_i(x))$ be the corresponding extension fields. Then, by Lemma 4.3.1, $\mathbb{F}[A_i] \cong \mathbb{E}_i$ equips V_i with a structure of a vector space over $\mathbb{E}_i$. Let $\alpha_i = [x]$ be the root of $p_i(x)$ in $\mathbb{E}_i$. It is a primitive element of $\mathbb{E}_i$ over $\mathbb{F}$.

We, thus, have the following invariant attached to A.

(i) The partition of an integer n, π: $n = \sum_{i=1}^{r} n_i$.

(ii) A decomposition D_π of V using π: $v = \bigoplus_{i=1}^{r} V_i$ where dim $V_i = n_i$.

(iii) Irreducible polynomials $p_i(x)$ in $\mathbb{F}[x]$, deg $p_i(x) = m_i$ and the corresponding extension fields $\mathbb{E}_i = \mathbb{F}[x]/(p(x))$ of $\mathbb{F}$, so that m_i divides n_i where $n_i = d_i m_i$.

(iv) A structure $\sigma_{\mathbb{E}_i}$ on V_i as a vector space over $\mathbb{E}_i$ which extends as a vector space over $\mathbb{F}$.

(v) A primitive element α_i of $\mathbb{E}_i$ over $\mathbb{F}$ such that $p_i(x)$ is the irreducible polynomial of α_i. We assume $p_i(x)$'s to be pairwise distinct.

We will denote this data as $(\pi, D_\pi, \varepsilon_\pi, \sigma_\pi, \alpha_\pi)$ and using this we can uniquely determine a semisimple operator $A : V \to V$ by setting $A\restriction_{V_i}: V_i \to V_i$ to be the scalar multiplication by

α_i in the $\mathbb{E}_i$ structure.

To understand the numerical and spatial invariants, we must determine $N^*(A)$, the normalizer of $\mathcal{Z}_G^*(A)$. $\mathcal{Z}_G^*(A_i)$ acts irreducibly on V_i because if $B \in \mathcal{Z}_G^*(A_i)$, then it is conjugate of A_i and the minimal polynomial of B is the same. So an element in $N^*(A)$ must permute V_i's.

Let $h : N^*(A) \to S_r$ be the corresponding homomorphism and $N_1^*(A) = \ker h$. An element of $N_1^*(A)$ leaves each V_i invariant. This is because the kernel will consist of all those elements that do not permute V_i's, i.e., that take each V_i to itself. Therefore, $N_1^*(A)$ normalizes $\mathbb{E}_i = \mathbb{F}[A_i]$.

Consider the homomorphism $h_1 : N_1^*(A) \to \prod_{i=1}^r \mathrm{Aut}(\mathbb{E}_i/\mathbb{F})$. The kernel of h_1 is exactly $\mathcal{Z}_G^*(A)$. Through this we arrive at the following conclusion.

Proposition 4.3.4. *[SK07] Let V be a vector space of dimension n over a field $\mathbb{F}$, and $A : V \to V$ a semisimple operator. Then the Weyl group $W(A)$ is a finite group.*

Theorem 4.3.5. *[SK07] Let V be a vector space of dimension n over a field $\mathbb{F}$. Then*

(i) The semisimple operators on V are in the one-one correspondence with the symbols $(\pi, D_\pi, \varepsilon_\pi, \sigma_\pi, \alpha_\pi)$ as constructed above.

(ii) The conjugacy classes of semisimple operators are in one-one correspondence with the symbols (π, α_π). Equivalently they are in one-one correspondence with the monic polynomials of degree n in $\mathbb{F}[x]$.

(iii) The z-classes of semisimple operators are in one-one correspondence with the symbols (π, ϵ_π).

Corollary 4.3.6. *If there are only finitely many field extensions of degree $\leq n$ of a given field $\mathbb{F}$ then there are only finitely many z-classes of semisimple operators on an n-dimensional vector space V.*

Chapter 5

z-classes of linear and affine maps

In this chapter, we will extend the theory of "dynamical types" we studied in Chapter 3 to linear and affine maps and investigate their z-classes. This chapter is based on the results of Kulkarni's paper titled "Dynamics of linear and affine maps". [Kul07] We had concluded Chapter 3 with a theorem that showed given an n-dimensional vector space, if there are finitely many simple extensions of atmost degree n over the underlying field, then there are finitely many z-classes of the semisimple operators on V. The work we will do in this chapter would result in an extension of this theorem to any linear and affine map.
From here on, $\mathbb{F}$ would refer to a field, V to a vector space over $\mathbb{F}$ and $p(x)$ to a monic irreducible polynomial in $\mathbb{F}[x]$, unless explicitly mentioned. Further, $L(V)$ is the set of all linear maps from V to V. The vector space $\mathbb{A}$ which is underlying V, has no distinguished base-point and is called the *affine space*. An *affine map* of $\mathbb{A}$ is a map $(A, v) : V \to V$ of the form $(A, v)(x) = Ax + v$, $A \in L(V)$ and $x \in V$. So, $A(V)$ is the set of all affine maps from V to V. Let $GL(V)$ and $GA(V)$ be the subsets of $L(V)$ and $A(V)$ respectively, consisting of invertible elements. They form groups and they act on $L(V)$ and $A(V)$ respectively by conjugation. Then, the stabilizer subgroups at T in $GL(V)$ and $GA(V)$ are the centralizers of T in $GL(V)$ and $GA(V)$ respectively and we denote this by $\mathcal{Z}_L^*(T)$ and $\mathcal{Z}_A^*(T)$ respectively.

5.1 Classical Theory for $L(V)$

Let $T \in L(V)$ and $m_T(x)$ be the minimal polynomial of T. We know that $\mathbb{F}[T] \cong \mathbb{F}[x]/(m_T(x))$. Now, let $m_T(x) = \prod_{i=1}^{r} p_i(x)^{d_i}$ such that $p_i(x) \in \mathbb{F}[x]$ and $p_i(x)$ are pairwise distinct.
The Primary Decomposition Theorem (2.1.2) provides a decomposition $V = \bigoplus_{i=1}^{r} V_i$, where

$V_i = \text{Ker } p_i(T)^{d_i}$ are T-invariant subspaces.

$$\text{Let } x \in \text{Ker } p_i(T)^{d_i} \text{ and } S \in \mathcal{Z}_L(T) \Rightarrow ST = TS$$
$$p_i(T)^{d_i}\big(S(x)\big) = S\big(p_i(T)^{d_i}(x)\big) = S(0) = 0 \Rightarrow S(V_i) \subseteq V_i$$

This shows that the decomposition is invariant under $\mathcal{Z}_L(T)$.
Let $T_i = T \restriction_{V_i}$. Then $m_{T_i}(x) = p_i(x)^{d_i}$. We also have a canonical decomposition

$$\mathcal{Z}_L(T) = \prod_{i=1}^{r} \mathcal{Z}_L(T_i)$$

We have now reduced the problem to $m_T(x) = p(x)^d$, where $p(x) \in \mathbb{F}[x]$.
Let us now consider the example where $V = \mathbb{F}[x]/(p(x)^d)$. For $u(x) \in \mathbb{F}[x]$, let $[u(x)]$ be the class of $u(x)$ in $\mathbb{F}[x]/(p(x)^d)$. Let T be the operator $\mu_x : [u(x)] \to [xu(x)]$ and $V_i = \{[f(x)p(x)^i] \mid f(x) \in \mathbb{F}[x]\}$. We have the following flag of subspaces

$$0 = V_d \subset V_{d-1} \subset \cdots \subset V_1 \subset V_0 = V$$

Suppose, W is a T-invariant subspace of V. If $[f(x)p(x)^i] \in W \Rightarrow [g(x)f(x)p(x)^i] \in W$, $\forall$ $g(x) \in \mathbb{F}[x]$. Let i be the smallest non-negative integer such that for $[f(x)p(x)^i] \in W$, $p(x)$ does not divide $f(x)$. Then $[f(x)]$ is a unit in $V \Rightarrow [p(x)^i] \in W \Rightarrow W = V_i$. This shows that that V_i's are the only T-invariant subspaces. These V_i's do not have a complementary T-invariant subspace and so, (V, T) is an indecomposable pair.
We will now show that the converse of this example is also true.

Theorem 5.1.1. *Let (V, T) be a pair such that $m_T(x) = p(x)^d$ where $p(x) \in \mathbb{F}[x]$ and deg $p(x) = m$. Then (V, T) is a direct sum of T-invariant indecomposable subspaces, each dynamically equivalent to $(\mathbb{F}[x_i]/(p(x_i)^{d_i}), \mu_{x_i})$. Here $d_i \leq d$, and $d_i = d$ for at least one i and dim $V = m\sum_i d_i$*

We have already proven the above theorem for the case $d = 1$ as a consequence of Lemma 4.3.2. The consequences of this proof were that $\mathcal{Z}_L(T)$ is the set of linear operators which are linear in $\mathbb{E} = \mathbb{F}[x]/(p(x))$. Thus, $\mathcal{Z}_L(T) \cong L_{\mathbb{E}}(V)$, $\mathcal{Z}_L^*(T) \cong GL_{\mathbb{E}}(V)$ and the orbits of $\mathcal{Z}_L^*(T)$ are V and $\{0\}$. We will call the T-action *dynamically semi-simple.*

5.2 Orbits of $\mathcal{Z}_L^*(T)$ and a Canonical Maximal $\mathcal{Z}_L(T)$-Invariant Flag

Let $T \in L(V)$, $m_T(x) = p(x)^d$, where $p(x)$ in $\mathbb{F}[x]$ and deg $p(x) = m$. Then, $\mathbb{E} = \mathbb{F}/(p(x))$ is a simple field extension of $\mathbb{F}$ and $\dim_{\mathbb{E}}\mathbb{F} = m$. We may assune $d \geq 2$. Let $N = p(T)$ and $V_i =$ ker $N^i, i = 0, 1, \dots, d$. Since N is nilpotent, we have a $Z_L(T)$-invariant flag of subspaces

$$0 = V_0 \subset V_1 \subset \cdots \subset V_d = V$$

Now, let $\bar{T}_i$ be the operator induced by T on V_i/V_{i-1}, for $i = 1, 2, \dots, d$. Then $m_{\bar{T}_i}(x) = p(x)$. Then, V_i/V_{i-1} has a canonical $\mathbb{E}$-structure. Then $\dim_{\mathbb{F}}V_i/V_{i-1}$ and consequently $\dim_{\mathbb{F}}V$ are divisible by m. Let $\dim_{\mathbb{F}}V = ml = n$.

The aim now is to obtain a canonical, maximal $\mathcal{Z}_L(T)$-invariant refinement of this flag. We shall introduce a double subscript notation $V_{i,j}$ for convenience, where $V_i = V_{i,0}$. If we insert $k-1$ terms between V_i and V_{i+1} then $V_{i+1} = V_{i,k}$.

From the theory of elementary divisors, we know that for T, elementary divisors are of the form $p(x)^{s_i}, i = 1, 2, \dots, r$, where $1 \leq s_1 < \cdots < s_r = d$ are the exponents and σ_i is the multiplicity of $p(x)^{s_i}$. Then, for n =dim $V = ml$, $l = \sum_{i=1}^r s_i\sigma_i$.

Lemma 5.2.1. *1. For $i > 0$, $N = p(T)$ maps V_i into V_{i-1}*

2. For $i > 1$, the map induced by N on $V_i/V_{i-1} \to V_{i-1}/V_{i-2}$ is injective.

Proof. 1. Let $x \in V_i$ =Ker N^i. We know that, $p(T)(x) \in$ Ker $N^{i-1} = V_{i-1}$.

2. Let $x + V_{i-1}$. Then $N(x + V_{i-1}) = p(T)(x + V_{i-1}) = p(T)x + V_{i-2}$.

□

Let $(e_1, \dots, e_k)$ be elements in V_d such that their images $(\bar{e_1}, \dots, \bar{e_k})$ form an $\mathbb{E}$-basis of V_d/V_{d-1}. Since, $T^j(e_i), 1 \leq j \leq m-1, 1 \leq i \leq k$ are independent over $\mathbb{F}$ mod V_{d-1} , they are linearly independent over $\mathbb{F}$. This is because if $(\bar{e_1}, \dots, \bar{e_k})$ are a basis, then they must be non-zero. Since $\bar{e_i} = \bar{T}^d(e_i) \neq 0 \Rightarrow e_i \notin V_{d-1}$. Then, by T-invariance, $T^j(e_i) \in V_d$ and as T is a linear map, $T^j(e_i) \notin V_{d-1}$.

Let W_d be the $\mathbb{F}$-span of $T^j(e_i)$. We can construct $V_d = W_d \oplus V_{d-1}$. Here, V_{d-1} is T-invariant but W_d is not. But our construction is T-invariant mod V_{d-1}, so, we will call subspaces like W_d, *almost T-invariant subspace.*

Now, by Lemma 5.2.1, $W_d \subseteq V_d \Rightarrow N(W_d) \subseteq V_{d-1}$ and since $N(W_d) \cong V_d/V_{d-1}$, then N maps

W_d injectively into V_{d-1}/V_{d-2} as a subspace which is complementary to V_{d-2}. The subspace $V_{d-2}+N(W_d)$ is T-invariant and $\mathcal{Z}_L(T)$-invariant subspace of V_{d-1}. Now, if $V_{d-2}+N(W_d)$ is a proper subspace of V_{d-1}, we will insert an additional subspace in the flag between V_{d-2} and V_{d-1}.

Suppose $V_{d-2}+N(W_d)$ is a proper subspace of V_{d-1}. In fact, $(V_{d-2}+N(W_d))/V_{d-2}$ is an $\mathbb{E}$-subspace of V_{d-1}/V_{d-2}. We will denote e_i by $e_{d,i}$ and $k=\dim_{\mathbb{E}} V_d/V_{d-1}$ by k_d. Now, let $k_{d-1}=\dim_{\mathbb{E}} V_{d-1}/V_{d-2} - \dim_{\mathbb{E}}(V_{d-2}+N(W_d))/V_{d-2}$.

If $k_{d-1}\neq 0$, we will choose $e_{d-1,i}, 1\leq i\leq k_{d-1}$ in V_{d-1}/V_{d-2} such that mod V_{d-2}, they form an $\mathbb{E}$-basis of the subspace of V_{d-1}/V_{d-2} complementary to $(V_{d-2}+N(W_d))/V_{d-2}$. Then $T^j(e_{d-1,i}), 1\leq j\leq m-1, 1\leq i\leq k_{d-1}$ are linearly independent over $\mathbb{F}$. Let W_{d-1} be the $\mathbb{F}$-span of $T^j(e_{d-1,i})$. Then W_{d-1} is an almost T-invariant subspace of V_{d-1}. Then N injectively maps W_{d-1} to a subspace complementary to $V_{d-3}+N^2(W_d)$. If $V_{d-3}+N(W_{d-1})+N^2(W_d)$ is a proper subspace, then we will insert it as an additional subspace in the flag between $V_{d-3}+N^2(W_d)$ and V_{d-2}. In case $V_{d-2}+N(W_d)$ is not a proper subspace of V_{d-1} , then just take $W_{d-1}=0$.

If we proceed in this way, we obtain the following refined flag of non-increasing dimensions

$$\begin{aligned}
0=V_0\subset N^{d-1}(W_d)\subset N^{d-1}(W_d)+N^{d-2}(W_{d-1})\subset\dots\\
N^{d-1}(W_d)+N^{d-2}(W_{d-1})+\dots+N(W_2)+W_1=V_1\subset\\
V_1+N^{d-2}(W_d)\subset V_1+N^{d-2}(W_d)+N^{d-3}(W_{d-1})\subset\dots\\
V_1+N^{d-2}(W_d)+N^{d-3}(W_{d-1})+\dots+N(W_3)+W_2=V_2\subset\dots\\
\dots\dots\\
V_{d-3}\subset V_{d-3}+N^2(W_d)\subset V_{d-3}+N^2(W_d)+N(W_{d-1})\subset\\
V_{d-3}+N^2(W_d)+N(W_{d-1})+W_{d-2}=V_{d-2}\subset\\
V_{d-2}+N(W_d)\subset V_{d-2}+N(W_d)+W_{d-1}=V_{d-1}\subset V_{d-1}+W_d\subset V_d
\end{aligned}$$

Now, the sum $\bigoplus_{j=0}^{d-1}N^j(W_d)$ forms a T-invariant subspace dynamically equivalent to k_d copies of $\mathbb{F}[x]/(p(x)^d)$. If $W_s=0$, then those terms do no occur. We know that $\mathbb{F}$-span of $T^j(e_{s,1}),\dots,T^j(e_{s,k_s}), 0\leq j\leq m-1$. Then,
$N^uT^j(e_{s,1}),\dots,N^uT^j(e_{s,k_s}), 0\leq j\leq m-1, 0\leq u\leq s-1$ is a basis of $\bigoplus_{j=0}^{s-1}N^j(W_s)$.

For a flag that has strictly increasing dimensions, let $1\leq s_1<\dots<s_r=d$ be integers such that $W_{s_i}\neq 0$. Let $m\sigma_i=$dim $W_{s_i}, 1\leq i\leq r$, $V_i=V_{i,0}$ and for $0\leq i\leq s_{r-j+1}, 1\leq j\leq r$, set

$$V_{i,j}=V_i+N^{s_r-i}(W_{s_r})+N^{s_{r-1}-i}(W_{s_{r-1}})+\dots+N^{s_{r-j+1}-i}(W_{s_{r-j+1}})$$

In order to avoid ambiguities, consider W'_s to be another choice of almost T-invariant subspace complementary to the subspace before V_{s+1} in the refined flag. We can construct W'_s in the same way by starting with $e'_{s,1}, e'_{s,2}, \ldots, e'_{s,k_s}$ such that $N^uT^j(e'_{s,1}), \ldots, N^uT^j(e'_{s,k_s}), 0 \leq j \leq m-1, 0 \leq u \leq s-1$ is the corresponding basis of $\bigoplus_{j=0}^{s-1} N^j(W'_s)$. If we define a $\mathbb{F}$-linear map that sends $N^uT^j(e_{s,m})$ to $N^uT^j(e'_{s,m})$ and is identity on the remaining $\bigoplus_{j=0}^{t-1} N^j(W_t)$ for $t \neq s$. This map is invertible, commutes with T and carries W_s into W'_s.
For two successive terms $V_{i,j}$ and $V_{i,j+1}$, we observe that $Z_L^*(T)$ is transitive on $V_{i,j+1} - V_{i,j}$ and we conclude this section by summarizing this result in the following theorem.

Theorem 5.2.2. *Let T be in $L(V)$, $m_T(x) = p(x)^d$, where $p(x) \in \mathbb{F}[x]$. Then V admits a canonical, maximal $\mathcal{Z}_L(T)$-invariant flag. A complement of each term appearing in the flag in its succeeding term is an orbit of $\mathcal{Z}_L^*(T)$. In particular, the quotient of each term appearing in the flag by its preceding term is an irreducible module over the group $\mathcal{Z}_L(T)^*$.*

5.3 Strongly Commuting Operators

Let $T \in L(V)$. An operator $S \in L(V)$ *strongly commutes* with T if S commutes with T and leaves every T-invariant subspace of V invariant.

Theorem 5.3.1. *Let $T \in L(V)$. An operator $S \in \mathcal{Z}_L(T)$ strongly commutes with T iff $S \in \mathbb{F}[T]$.*

Proof. Let $S \in \mathbb{F}[T]$. Then, $S = a_0 + a_1T + a_2T^2 + \ldots$. It is fairly obvious that $S \circ T = T \circ S$ and S also leaves every T-invariant subspace invariant.
Conversely, suppose $S \in \mathcal{Z}_L(T)$ strongly commutes with T.
Case 1: Suppose (V, T) is dynamically equivalent to $(\mathbb{F}[x]/(p(x)^d), \mu_x)$, where $p(x) \in \mathbb{F}[x]$. Let $S \in \mathcal{Z}_L(T)$ and $S(1) = [f(x)]$. Then $S = f(T)$ and $\mathcal{Z}_L(T) = \mathbb{F}[T]$
Case 2: $m_T(x) = p(x)^d \in \mathbb{F}[x]$. Then V is a direct sum of T-invariant subspaces W_i that are dynamically equivalent to $(\mathbb{F}[x_i]/(p(x_i)^{d_i}, \mu_x)$. Let $e_i, 1 \leq i \leq k$ be a T-module generator in W_i.
Now, say $S \restriction_{W_1} = q_1(T)$ where $q_1(x)$ is a unique polynomial of degree at most dm. For $j \geq 2$, let $q_j(x)$ be the polynomial of degree at most d_jm, such that $S \restriction_{W_j} = q_j(T)e_j \Rightarrow S(e_1 + e_j) = q_1(T)e_1 + q_j(T)e_j$. Also, $S(e_1 + e_j) = u(T)(e_1 + e_j)$ for some polynomial $u(x)$ of degree at most dm. It follows that $(q_1(T) - u(T))e_1 = -(q_j(T) - u(T))e_j$. As $W_1 \cup W_j = 0 \Rightarrow (q_1(T) - u(T)) \equiv (q_j(T) - u(T)) \equiv 0(\text{mod } p(x)^{d_j})$. So $q_1(T) \equiv q_j(T)(\text{mod } p(x)^{d_j})$.

Case 3: Consider the general case. If we consider the primary decomposition of V, then S leaves each of those V_i's invariant. Now, when $S\restriction_{V_i} = q_i(T)$ is determined uniquely mod $p_i(x)^{d_i}$. Then, by Chinese Remainder Theorem, there will exist a unique polynomial $q(x)$ mod $m_T(x)$ which is congruent to $q_i(x)$ mod $p_i(x)^{d_i}$. □

5.4 "$S + N$"-decomposition

Let $T \in L(V)$ and $\mathbb{E}$ be an extension of $\mathbb{F}$. An $\mathbb{E}$-*structure* on V is an $\mathbb{F}$-algebra homomorphism $\sigma_{\mathbb{E}} : \mathbb{E} \to L(V)$, which is injective and this enables us to look at V as a vector space over $\mathbb{E}$. An $\mathbb{E}$-structure $\sigma_{\mathbb{E}}$ is said to be T-invariant if the image of $\sigma_{\mathbb{E}}$ lies in $\mathcal{Z}_L(T)$.
We investigate when V admits a T-invariant $\mathbb{E}$-structure. For $d = 1$, T induces a canonical $\mathbb{E}$-structure, $\mathbb{F}[T] \cong \mathbb{E}$ and the inclusion map of $\mathbb{F}[T]$ in $\mathcal{Z}_L(T)$ is a T-invariant flag. We shall assume $d \geq 2$. Then V_i/V_{i-1} admits a canonical T-invariant $\mathbb{E}$-structure as the minimal polynomial of the induced operator is $p(x)$. We will see if these canonical $\mathbb{E}$-structures on V_i/V_{i-1}'s can be lifted to a canonical $\mathbb{E}$-structure on V, and by that we mean:

1. Each T-invariant subspace is an $\mathbb{E}$-subspace

2. For each $i = 1, \dots, d$, the induced $\mathbb{E}$-structure on V_i/V_{i-1} coincides with the one induced by T.

For $f(x) \in \mathbb{F}[x]$, let $f'(x)$ be the formal derivative.

Theorem 5.4.1. *Let $T \in L(V)$, $m_T(x) = p(x)^d$, where $p(x) \in \mathbb{F}[x]$ and $\mathbb{E} = \mathbb{F}[x]/(p(x))$. Then V admits a T-invariant $\mathbb{E}$-structure iff either $d = 1$ or $p'(x)$ is not identically zero.*

Proof. We will assume $d \geq 2$ and deg $p(x) \geq 2$. The proof for the omitted cases is trivial.
We know that (V, T) is dynamically equivalent to a direct sum of pairs of the form $(\mathbb{F}[x]/(p(x)^e), \mu_x)$ where $e \leq d$ and $\mu_x([u(x)]) \mapsto [xu(x)]$. For $y \in \mathbb{F}[x]$, let $[y]$ denote its class in $\mathbb{F}[x]/(p(x)^e)$.
We need to show that there exists a polynomial $z = u(x) \in \mathbb{F}[x]$ such that the minimal polynomial of the corresponding operator is $p(x)$.
Now, $p(x)$ and $p'(x)$ are relatively prime and by Euclidean Algorithm $\exists\ a(x), b(x) \in \mathbb{F}[x]$ such that $a(x)p(x) + b(x)p'(x) = 1$. Let $y = x - b(x)p(x)$. We can take $\epsilon = -b(x)p(x)$ and use the Taylor's theorem to get

$$\begin{aligned} p(y) = p(x+\epsilon) &= p(x) + \epsilon p'(x) + \frac{\epsilon^2}{2}p''(x) + \dots \\ \equiv p(x)(1 - b(x)p'(x)) + \dots &\equiv p(x)(a(x)p(x)) + \dots \equiv 0(\text{ mod } p(x)^2) \end{aligned}$$

So, $p(y)^r = 0$ for a suitable $r < e$. Then, μ_y has a minimal polynomial of the form $p(x)^r \Rightarrow \mathbb{F}[[y]] = \mathbb{F}[x]/(p(x)^r)$ and $\mathbb{F}[[y]] \subset \mathbb{F}[[x]]$. We can prove the existence of the aforementioned polynomial by induction.
Conversely, suppose we have a pair (V, T), $m_T(x) = p(x)^d, d \geq 2$, $\mathbb{E} = \mathbb{F}[x]/(p(x))$ and V admits a T-invariant $\mathbb{E}$-structure. Then $\exists\ S \in \mathcal{Z}_L(T)$ such that $m_S(x) = p(x)$ and in the associated flag V_2 is S-invariant. We need to prove that $p'(x) \not\equiv 0$. By Taylor's theorem,

$$p(s + u(T)) = p(S) = u(T)p'(S) + \cdots = p(S) = 0$$

But then

$$p(T) = p(S + T - S) = P(S + T) - Sp'(S + T) + + \cdots = P(S + T) = 0$$

This is a contradiction as $m_T(x) = p(x)^2$.
For uniqueness, let $\sigma_1 : \mathbb{E} \to \mathcal{Z}_L(T)$, $\sigma_2 : \mathbb{E} \to \mathcal{Z}_L(T)$, be two canonical T-invariant $\mathbb{E}$-structures. Using a T-invariant subspace, we can reduce this to when (V, T) is dynamically equivalent to $(\mathbb{F}[x]/(p(x)), \mu_x)$. Then $\mathcal{Z}_L(T) = \mathbb{F}(T)$. Let α be a primitive element of $\mathbb{E}$ over $\mathbb{F}$, and $\sigma_i(\alpha) = S_i$, $i = 1, 2$. Let $S_i = f_i(T)$ where $f_i(x) \in \mathbb{F}[x]$. Using the induction hypothesis in the first part completes this proof. $\square$

Let $T \in L(V)$. A "$S + N$"-decomposition of T is a pair S, N such that

(i) $T = S + N$

(ii) S is dynamically semi-simple

(iii) N is nilpotent

(iv) $SN = NS$

Theorem 5.4.2. *Let $T \in L(V)$ and $m_T(x) = \prod_{i=1}^r p_i(x)^{d_i}$, where $p_i(x)$'s are monic irreducible polynomials in $\mathbb{F}[x]$. Then*

1. *T admits a "$S + N$"-decomposition iff for each i, either $d_i = 1$ or else $p_i'(x) \not\equiv 0$*

2. *If it exists, a "$S + N$"-decomposition is unique.*

3. *If $m_T(x) = p(x)^d$, $p(x) \in \mathbb{F}[x]$, $\mathbb{E} = \mathbb{F}[x]/(p(x))$ and a "$S + N$"- decomposition exists, then S defines the canonical T-invariant $\mathbb{E}$-structure on V. In particular, S strongly commutes with T, and so S, and hence N, are polynomials in T.*

Proof. By definition, $S, N \in \mathcal{Z}_L(T)$. This reduces the case to when $m_T(x) = p(x)^d$. Consider $d \geq 2$. Let $\mathbb{E} = \mathbb{F}[x]/(p(x))$ and $p'(x) \not\equiv 0$. The previous theorem tells us that under this condition, we have a polynomial $f(x) \in \mathbb{F}[x]$ such that $S = f(T)$ defines a canonical T-invariant $\mathbb{E}$-structure on V, particularly, $m_S(x) = p(x)$. S is dynamically semi-simple. Let $\bar{T}_i$, $\bar{S}_i$ be the operators induced by T, S on $V_i = \text{Ker } p(T)^i$, $i = 0, 1, \ldots d$. As S defines a canonical T-invariant $\mathbb{E}$-structure we have $\bar{T}_i = \bar{S}_i$. Then $N = T - S$ is nilpotent and $T = S + N$ is a "$S + N$"-decomposition of T.
Conversely, if $T = S + N$ is an "$S + N$"-decomposition of T. Then the induced operators on V_i are commuting dynamically semi-simple operators. Then, their nilpotent difference N_i must be 0. $\Rightarrow m_{\bar{T}_i}(x) = p(x) = m_{\bar{S}_i}(x)$. Then $m_S(x) = p(x) \Rightarrow \mathbb{F}[S] \cong \mathbb{E}$ and S defines a T-invariant $\mathbb{E}$-structure on V. So $p_i'(x) \not\equiv 0$. □

5.5 Affine Case

We will extend this theory to the affine case.
Let $\mathbb{F}$ be a field, V be the vector space over this field. Let $\mathbb{A}$ be the underlying affine case and $T = (A, v)$ be an affine map such that $x \mapsto Ax + v$. Let $S = (\alpha, a)$, $\alpha \in GL(V)$ be an element of $GA(V)$. Then,

$$S^{-1} = (\alpha^{-1}, -\alpha^{-1}a) \tag{5.1}$$

$$STS^{-1} = (\alpha A \alpha^{-1}, -\alpha A \alpha^{-1} a + \alpha v + a) \tag{5.2}$$

Let $\mathcal{C}_L(V)$ and $\mathcal{C}_A(V)$ denote the orbit-spaces $L(V)/GL(V)$ and $A(V)/GA(V)$. For $T \in L(V)$ and $A(V)$, let $[T]_L$ and $[T]_A$ denote its orbit in $\mathcal{C}_L(V)$ and $\mathcal{C}_A(V)$. The map $(A, v) \mapsto A$ is a homomorphism $l : A(V) \to L(V)$. The formula 5.2 shows that the map $[(A, v)]_A \mapsto [A]_L$ is a well-defined map $[l] : \mathcal{C}_A(V) \to \mathcal{C}_L(V)$.

Lemma 5.5.1. *Suppose $S = (A, v)$ and $T = (A, w)$ be in $A(V)$ such that $m_A(x) = (x - 1)^r$. Let s, t be the least non-negative integers $\leq r$ satisfying $(A - I)^s(v) = 0$ and $(A - I)^t(w) = 0$. Then S and T are in the same $GA(V)$-orbit iff $s = t$.*

Proof. Formula 5.2 shows us that (α, a) conjugates S into T iff $\alpha \in \mathcal{Z}_L^*(A)$ and $w = (I - A)a + \alpha v$. Since $m_A(X) = (x - 1)^r$, we have reduced it to the problem we solved in the previous section. Just set $N = I - A$ and consider the $\mathcal{Z}_L(A)$-invariant refined flag. Without loss of generality, let us assume $s \leq t$. Further, let $V_t = V_{t-1,k}$ and $v \in V_t - V_{t-1,k-1} \Rightarrow \alpha \in \mathcal{Z}_L^*(A)$ and $w = (I - A)a + \alpha v$ iff $s = t$. □

Theorem 5.5.2. *$[l]$ is a finite map, that is $[l]^{-1}([A])$ has only finitely many elements. For $A \in L(V)$, let $m_A(x) = (x-1)^r g(x)$ where $g(1) \neq 0$ be its minimal polynomial. Here $r \geq 0$ is an integer. Then $[l]^{-1}([A])$ has $r+1$ elements.*

Proof. Consider the case where $r = 0$. Then det $(I - A) \neq 0$. So, the equation $Ax + v = x$ has a unique solution in x. Let $\tau = (I, x_0)$. Then $\tau(A, v)\tau^{-1} = (A, 0) = A$. Any element in $l^{-1}(A)$ is conjugate to A. Then $[l]^{-1}$ has a unique element.
If $r > 0$, then $V = V_1 \oplus V_2$ where $V_1 =$ ker $(A-I)^r$ and $V_2 =$ Ker $g(A)$. Let $T = (A, v)$. Suppose, $v = v_1 + v_2$ where $v_i \in V_i$. Let x_0 be the solution in V_2 of the equation $Ax + v_2 = x$. Such a solution exists as det $(I-A) \restriction_{V_2} \neq 0$. Let $\tau = (I, x_o)$, then $\tau(A, v)\tau^{-1} = (A, v_1)$. We have shown that any element $(A, v) \in l^{-1}(A)$ is in the same $GA(V)$-orbit as (A, v_1) where $(A-I)^r(v_1) = 0$. Let s be the least non-negative integer such that $(A-I)^s(V_1)$. Then, as a consequence of Lemma 5.5.1, $[l]^{-1}([A])$ has $r+1$ elements. □

We can now determine the centralizer of an affine map.

Theorem 5.5.3. *Let $T = (A, v) \in A(V)$ and $V_i =$ Ker $(A-I)^i$*

1. *If T has a fixed point then $\mathcal{Z}_A(T)$ is conjugate to*

$$\{(B, w) \mid B \in \mathcal{Z}_L(A), \text{ and } w \in V_1\}$$

2. *Suppose T has no fixed point, and $m_A(x) = (x-1)^r g(x)$, $g(1) \neq 0$ is the minimal polynomial of A. Let $s \leq r$ be the least positive integer such that $(A-I)^s v = 0$. Then $\mathcal{Z}_A(T)$ is conjugate to*

$$\{(B, w) \mid B \in \mathcal{Z}_L(A), w \in V_{s+1}, (B-I)v = (A-I)w\}$$

An element (B, w) in $\mathcal{Z}_A^*(T)$ necessarily has eigenvalue 1 with multiplicity at least s.

Let $T = (A, v)$ in $A(V)$ and $S = (B, w) \in \mathcal{Z}_A(T)$. Then, $ST = TS$ is equivalent to

1. $BA = AB, B \in Z_L(T)$
2. $Bv + w = Aw + v \Rightarrow (B-I)v = (A-I)w$

Now, suppose T has a fixed point. Then by conjugation by an element in $GA(V)$ allows us to take $v = 0$. Then, we take the flag associated with A.

$$\mathcal{Z}_A(T) = \{(B, w) \mid B \in \mathcal{Z}_L(A), w \in V_1\}$$

where $V_1 = \text{Ker } (A - I)$.
Now, suppose T has no fixed point. Then, by Theorem 5.5.2, let us assume $m_A(x) = (x-1)^r g(x), g(1) \neq 0$ is the minimal polynomial of A and s is the least positive integer such that $(A-I)^s v = 0$. Then,

$$(A-I)^s(B-I)v = (B-I)(A-I)^s v = 0 = (A-I)^{s+1}w$$

$w \in V_{s+1}$ where $V_i = \text{Ker}(A-I)^i$.
Conversely, if $w \in V_{s+1}$, $V_s = V_{s-1,k}$ and $v \in V_s - V_{s-1,k-1}$ and $(A-I)w \in V_S$. Then, $\exists$ $C \in \mathcal{Z}_L(T)$ such that $Cv = (A-I)w$ and these C's are determined by the refined flag. For each C, let $B = C + I$. These B's are the (B, w)'s $\in \mathcal{Z}_A(T)$.
Further, $(B, w) \in \mathcal{Z}_A^*(T)$ iff $B \in \mathcal{Z}_L^*(A)$. Then $Bv \in V_s - V_{s-1,k-1}$. Since $(A-I)w \in V_{s-1,k-1} \Rightarrow Bv \equiv v \bmod V_{s-1,k-1}$. Then the linear map induced by B on $V_s/V_{s-1,k-1}$ has eigenvalue 1. Then B has eigenvalue 1 and the N-images of the corresponding eigenvector shows that the multiplicity of the eigenvaue 1 is at least s.

5.6 Parametrization Theorems

Theorem 5.6.1. *[Kul07]*

1. *A $GL(V)$-orbit in its action on $L(V)$ is parametrized by*
    - *(i) A primary partition $\pi : n = \sum_{i=1}^r n_i, n_i = m_i l_i$*
    - *(ii) The secondary partitions $l_i = \sum_{j=1}^r s_{i,j}\sigma_{i,j}$, where $s_{i,1} < s_{i,2} < \cdots < s_{i,r_i}$*
    - *(iii) An $\mathbb{F}$-isomorphism class of pairs $(\mathbb{E}_i, \alpha_i)$, where $\mathbb{E}_i$ is a simple field extension of $\mathbb{F}$ of degree m_i with α_i as its primitive element, for $i = 1, 2, \ldots, r$.*
2. *A $GA(V)$-orbit in its action on $A(V)$ is parametrized by the data $(i), (ii), (iii)$ given above and with $m(x) = (x-1)^u g(x), g(1) \neq 0$*
    - *(iv) A non-negative integer $s \leq u$.*

Theorem 5.6.2. *[Kul07]*

1. *An element of $L(V)$ is uniquely determined by: The data (i), (ii) , (iii) of Theorem 5.6.1, in particular the field extensions $\mathbb{E}_i = \mathbb{F}[x]/(p_i(x))$, and the primitive elements α_i.*

(iv) A decomposition $\mathcal{D}_\pi : V = \bigoplus_{i=1}^{r} V_i$ *of* V *patterned on the primary partition* π*.*

(v) Flags $\mathcal{F}((n_i, m_i; \{(s_{i,1}, \sigma_{i,1}), (s_{i,2}, \sigma_{i,2}), \ldots, (s_{i,ri}, \sigma_{i,ri})\})$ *of subspaces in* V_i*, patterned on the secondary partitions.*

(vi) Compatible $\mathbb{E}_i$*-structures on the sub-quotients in the flag in each* V_i*.*

2. *An element* T *of* $A(V)$ *is uniquely determined by the following data.*
   Case 1: (T *has a fixed point)*
   Choose a fixed point as the origin. So, T *may be identified with an element in* $L(V)$*. The data* $(i), \ldots, (vi)$ *is independent of the choice of the fixed point. These data and the affine subspace of fixed points determine* T*.*
   Case 2: (T *has no fixed point)*
   Express T *as* (B, v) *so that there exists* s*, a least positive integer such that* $(I-B)^s v = 0$*. Then the invariants* $(i), \ldots, (vi)$ *associated to* B *and* v *uniquely determine* T*.*

We will now provide the proof for the first part in Theorem 5.6.1 and the proof for the second part and Theorem 5.6.2 be along the same lines.
It is fairly obvious that we can independently read from a refined flag, the numerical data about exponents and multiplicities in the elementary divisors. An element $T \in L(V)$ is associated with

1. the minimal polynomial $m_T(x) = \prod_{i=1}^{r} p_i(x)^{d_i}$
2. the primary partition dim $V = \sum_{i=1}^{r}$ dim V_i where $V_i =$ Ker $p_i(T)^{d_i}$
3. the secondary partitions with $s_{i,j}$'s being the exponents in the elementary divisors $p_i(x)^{s_{i,j}}$'s and $\sigma'_{i,j}s$ being the multiplicities of $p_i(x)^{s_{i,j}}$'s.

Conversely, suppose we are given this data. Let $V = \oplus_{i=1}^{r} V_i$ is an arbitrary decomposition patterned over the primary partition. We can construct a flag in each V_i with type given by the pairs $(s_{i,j}, \sigma_{i,j})$'s. Let $\mathbb{E}_i = \mathbb{F}[x]/(p_i(x))$ and $\alpha = [x]$. Now, after equiping the sub-quotients in the flag in V_i with a compatible family of E_i-structures, we will take an arbitrary $\mathbb{E}_\beth$-basis $(e_1, \ldots, e_k)$ in the component $V_{0,1}$ of the flag. Then

$$(e_1, \alpha e_1, \alpha^2 e_1, \ldots, \alpha^{m_i - 1} e_1, e_2, \alpha e_2, \ldots, \alpha^{m_i - 1} e_k)$$

is an $\mathbb{F}$-basis of $V_{0,1}$. We can continue this process to all the components in the chain ending in V_1, and define the operator T on V_1 having the minimal polynomial $p(x)$. Now when we

consider the component $V_{1,1}$ in the flag, we notice that by construction dim ${}_{\mathbb{F}}V_{1,1}/V_1$ is $m_i k$ and it has an $\mathbb{E}_i$-structure. Choose $(e'_1, \dots, e'_k)$ in $V_{1,1}$ whose classes $[e'_i]$ modulo V_1 form an $\mathbb{E}_i$-basis. Then, if we define $T^j e'_u, 1 \leq j \leq m-1, 1 \leq u \leq k$ in $V_{1,1}$ so that their classes modulo V_1 are $[\alpha^j e_u]$. Now we define $p(T)e'_i = e_i$ in $V_{0,1}$ and $p(T)T^j e'_i = T^j e_i, 1 \leq j \leq m_i - 1$. If we continue this process we will obtain a basis of V_i and an operator $T \in L(V_i)$ which has the given secondary partition on V_i. Taking the direct sum we obtain an operator T on V having the minimal polynomial $m(x)$ and the given primary and secondary partitions.
Now, suppose T, T' are two elements in $L(V)$ having the same data.Then the dimension of a primary component V_i equals $m_i l_i$. Then by conjugating with an appropriate element of $GL(V)$, we can say that T, T' have the same primary components. This reduces the case to when $m_T(x) = m_{T'}(x) = p(x)^d$. Then by our hypothesis T, T' have the same secondary partitions. We can construct the flags and the bases e_j, e'_j's of V for these flags. Then, the an element $g \in GL(V)$, $ge_i \mapsto e'_i$ conjugates T into T'.

Theorem 5.6.3. *[Kul07]*

1. *A z-class in the $GL(V)$-action on $L(V)$ is parametrized by*

   (i) A primary partition π: $n = \prod_{i=1}^{r} n_i, n_i = m_i l_i$

   (ii) The secondary partition $l_i = \sum_{j=1}^{r_i} s_{i,j}\sigma_{i,j}$ where $s_{i,1} < s_{i,2} < \cdots < s_{i,r_i}$

   (iii) Simple field extensions $\mathbb{E}_i$, $1 \leq i \leq r$ of $\mathbb{F}$, $[\mathbb{E}_i : \mathbb{F}_i] = m_i$

2. *A z-class of (A, v) in the $GA(V)$-action on $A(V)$ is parametrized by the data $(i), (ii), (iii)$ in case $m_A(x)$ does not have 1 as an eigenvalue. In case $m_A(x) = (x-1)^u g(x)$, $g(1) \neq 0$, and $u > 0$, then the z-class of (A, v) is parametrized by the data $(i), (ii), (iii)$ and*

   (iv) A non-negative integer $s \leq u$.

Let S, T be in the same z-class in $L(V)$. Then $\mathcal{Z}^*_L(S)$ and $\mathcal{Z}^*_L(T)$ are conjugate in $GL(V)$.

Lemma 5.6.4. *Let T be in $L(V)$. Then $\mathcal{Z}_L(T)$ as an F-subalgebra of $L(V)$ and $\mathcal{Z}^*_L(T)$ as a subgroup of $GL(V)$ uniquely determine each other.*

Proof. $\mathcal{Z}_L(T)$ uniquely determines $\mathcal{Z}^*_L(T)$ as the multiplicative subgroup of units.
Conversely, let S be a non-invertible element in $\mathcal{Z}_L(T)$. Then $m_S(x) = x^k f(x)$, with $k > 0$, and $f(0) \neq 0$. Now if $V_0 = \text{Ker } S^k$ and $V_1 = \text{Ker } f(S)$, then $V = V_0 \oplus V_1$ is a T-invariant decomposition. Let J_{V_0,V_1} denote the operator which is identity on V_0, and zero on V_1. Then

$J_{V_0,V_1} \in Z_L(T)$ and $S_1 = S + J_{V_0,V_1} \in \mathcal{Z}_L^*(T)$, Then $\mathcal{Z}_L(T)$ is a linear span of $\mathcal{Z}_L^*(T)$ and the operators J_{V_0,V_1} correspond to all T-invariant decompositions $V = V_0 \oplus V_1$. $\square$

Now we may assume that $\mathcal{Z}_L(T)$ and $\mathcal{Z}_L(S)$ are conjugate in $GL(V)$ by an element u. We can replace S by uSu^{-1}, so we can assume $\mathcal{Z}_L(S) = \mathcal{Z}_L(T)$.
Let C be the center of $\mathcal{Z}_L(T)$. By Frobenius' bicommutant theorem, we have $C = \mathbb{F}[S] = \mathbb{F}[T]$. Although C does not determine T, every element of C leaves every T-invariant subspace invariant. Let $p_i(x)$ be the primes associated to T and $V = \oplus V_i$ the corresponding primary decomposition. Let W be a T-invariant subspace of V_i such that the pair $(W, T \restriction_W)$ is dynamically equivalent tooo $(\mathbb{F}[x]/(p_i(x)^d), \mu_x)$. Then $W_j = \text{Ker } p_i(x)^j, 0 \leq j \leq d$ are all the T-invariant subspaces of W. As a subsapce of V is T-invariant iff it is S-invariant, then W_j's are also all the S-invariant subspaces of W. The $m_{S\restriction_W}(x)$ will be of the form $q(x)^e$.
Now, $q(x)$ is such that the pair $(W, T \restriction_W)$ is dynamically equivalent to $(\mathbb{F}[x]/(p_i(x)^e), \mu_x)$ for some e. Then there is an operator $A \in \mathcal{Z}_L(T)$ which maps W onto U with the action of S. This implies that $V = \oplus V_i$ is a primary decomposition with respect to S. We now have a well-defined choice for the primary partition of n. When we restrict the action of $\mathcal{Z}_L(T) = \mathcal{Z}_L(S)$ to V_i, we see that the secondary partitions are well-defined invariants of $\mathcal{Z}_L(T) = \mathcal{Z}_L(S)$. Finally, a simple field extension of $\mathbb{F}[x]$ is a well-defined invariant of $\mathcal{Z}_L(S) = \mathcal{Z}_L(T)$. The converse is fairly straightforward to prove. Using the results in Section 5, we can prove the same for the affine case.

A consequence of Theorem 5.6.3 is the following theorem that describes the conditions in which there are finitely many z-classes.

Theorem 5.6.5. *[Kul07] Let V be an n-dimensional vector space over a field $\mathbb{F}$. Suppose $\mathbb{F}$ has the property that there are only finitely many extensions of $\mathbb{F}$ of degree at most n. Then there are finitely many z-classes of $GL(V)$ and $GA(V)$, actions on $L(V)$ and $A(V)$.*

Summary and further

In this thesis, we have computed the z-classes of symmetric groups, general linear groups and general affine groups. We observed that they have finitely-many z-classes and in case of general linear groups and general affine groups, this happens when the underlying firld has only finitely many extensions of each degree. We have also investigated the notion of "dynamical types" and described the Weyl group at a point x. Using this theory, we have also computed the z-classes of semisimple operators.

We could further use the theory of the generalised symmetric group to study the z-classes of alternating groups. [BKS17]. We can also look at various subgroups of general linear group, like, unitriangular matrices [Bhu19], semisimple matrices, symplectic matrices, unipotent matrices etc.

In fact, there is a lot of recent literature in the conjugacy classes of unipotent and unitriangular matrices. [Alp06] [VLA03], to name a few.

Bibliography

[Alp06] Jonathan L Alperin. Unipotent conjugacy in general linear groups. *Communications in Algebra*, 34(3):889–891, 2006.

[Bhu19] Sushil Bhunia. Conjugacy classes of centralizers in the group of upper triangular matrices. *Journal of Algebra and Its Applications*, 2019.

[BKS17] Sushil Bhunia, Dilpreet Kaur, and Anupam Singh. z-classes and rational conjugacy classes in alternating groups. *Journal of the Ramanujan Mathematical Society*, 2017.

[BR02] T.S. Blyth and E.F. Robertson. *Further Linear Algebra*. Springer Undergraduate Mathematics Series. Springer London, 2002.

[BS19] Sushil Bhunia and Anupam Singh. Conjugacy classes of centralizers in unitary groups. *Journal of Group Theory*, 22(2):231–251, 2019.

[Con] Keith Conrad. Conjugation in a group.

[DJ14] Rahul Dattatraya and Vikas Jadhav. z-classes in p-groups of order $\leq p^5$. *Bulletin of Allahabad Mathematical Society*, 10 2014.

[GK09a] Krishnendu Gongopadhyay and Ravi Kulkarni. z-classes of isometries of the hyperbolic space. *Conformal Geometry and Dynamics of the American Mathematical Society*, 13(4):91–109, 2009.

[GK09b] Krishnendu Gongopadhyay and Ravi S Kulkarni. The z-classes of isometries. *Journal of the Indian Mathematical Society*, 2009.

[Gon13] Krishnendu Gongopadhyay. The z-classes of quaternionic hyperbolic isometries. *Journal of Group Theory*, 16(6):941–964, 2013.

[Gou06] R. Gouraige. *z-classes in Central Simple Algebras*. PhD thesis, 2006.

[JK84] G.D. James and A. Kerber. *The Representation Theory of the Symmetric Group*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1984.

[KKJ16] Ravindra Kulkarni, Rahul Dattatraya Kitture, and Vikas S. Jadhav. z-classes in groups. *Journal of Algebra and Its Applications*, 15(07):1650131, 2016.

[Kul07] Ravi S Kulkarni. Dynamics of linear and affine maps. *Asian Journal of Mathematics*, 2007.

[Sin08] Anupam Singh. Conjugacy classes of centralizers in g_2. *Journal of the Ramanujan Mathematical Society*, 2008.

[SK07] Ravi. S. Kulkarni. Dynamical types and conjugacy classes of centralizers in groups. *Journal of the Ramanujan Mathematical Society*, 2007.

[Ste74] R. Steinberg. *Conjugacy classes in algebraic groups*. Lecture notes in mathematics. Springer-Verlag, 1974.

[VLA03] Antonio Vera-López and JM Arregi. Conjugacy classes in unitriangular matrices. *Linear Algebra and its Applications*, 370:85–124, 2003.