

## ABSTRACT

WANG, QIANG. Classification of  $K_{\mathbb{F}}$ -orbits of Unipotent Elements in Symmetric  $\mathbb{F}$ -varieties of  $\mathrm{SL}(n, \mathbb{F})$ . (Under the direction of Dr. Amassa Fauntleroy and Dr. Aloysius G. Helminck)

Richardson proved in 1982 that, given an algebraic group  $G$  and an involution  $\sigma$ , we could have only a finite number of  $K$ -orbits of unipotent elements in the symmetric variety  $P = G/K$  over an algebraically closed field, where  $K = G^{\sigma}$  is the fixed point group of some involution  $\sigma$ . A question arises naturally: what if the field is not algebraically closed? In this thesis we try to answer this question and go a little further by listing all  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P$  explicitly. We work on the symmetric  $\mathbb{F}$ -variety  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  for the special linear group over an arbitrary field  $\mathbb{F}$  of characteristic not 2. We classify all  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P$  for all inner involutions for the special linear group. For Cartan (outer) involution, we classify  $K$ -orbits for small  $n$  only and illustrate how to get the canonical form for general  $n$ . Further proofs are still needed. We also classify  $G_{\mathbb{F}}$ -orbits of unipotent elements in  $G_{\mathbb{F}}$ .

© Copyright 2010 by Qiang Wang

All Rights Reserved

Classification of  $K_{\mathbb{F}}$ -orbits of Unipotent Elements in Symmetric  $\mathbb{F}$ -varieties of  $\mathrm{SL}(n, \mathbb{F})$

by  
Qiang Wang

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2010

APPROVED BY:

---

Kailash Misra

---

Ernie Stitzinger

---

Amassa Fauntleroy  
Co-Chair of Advisory Committee

---

Aloysius G. Helminck  
Chair of Advisory Committee

## DEDICATION

To my parents and my advisors  
for their love and help

## BIOGRAPHY

**Qiang Wang** was born on Oct 26, 1978, in Qixia, Shangdong Province, the People's Republic of China. He has one sister and a two-and-a-half year old niece. He received his Bachelor of Science degree in Computational Mathematics from University of Petroleum, China (Huadong) in 2002, and his Master of Science degree in Computational Mathematics from Nanjing Normal University, China in 2005. He enrolled in the PhD program in Applied Mathematics at North Carolina State University in 2005.

## ACKNOWLEDGEMENTS

I would like to thank my advisors and mentors Dr. Fauntleroy and Dr. Helminck for all of their encouragement, support and friendship throughout my PhD study and thesis work. It's really a pleasing, exciting and learning experience each time I meet either or both of them for personal or academic reasons. They are there to listen and give valuable suggestions when I have problems in my research or personal life. I am grateful to them with all of my heart. I would like to thank my parents who brought me into this world, raised me up and taught me how to be a man, and whose love and support enable me to chase my mathematical dream. I would like to thank my sister for her support and for her caring for our parents in China while I am in the US. I also would like to thank many professors in the math department, from whom I learned knowledge and the right attitude toward math that I need for my academic career. I want to say thanks especially to Dr. Fulp, Dr. Griggs, Dr. Hong, Dr. Kheyfets, Dr. Li, Dr. Misra and Dr. Stitzinger, who helped me make my decision in the direction of my research, supported me in teaching and helped me with my personal life. I would like to thank Denise, Charlene, Frankie, Carolyn, Di, Nicole and Joel for all the help and nice words from them. I would also like to give special thanks to Ms. Brenda Currin, a former department secretary and nice lady with an ADORABLE heart, who left the math department in summer 2006 due to a family reason. I had been helped by her since I came to NC State, from language help to getting used to the way of living life in the US. I owe her a hug. I would like to thank math graduate students, graduated or not, and it's really a great pleasure to meet them and be friends with most of them. I cannot name them one by one (since it would be at least one third of all math graduate students), though, for all kinds of reasons, thanks have to go to: Ralph Abbey, John Absher, Michael Allocca, Julie Beier, Jennifer Bell, Kathryn Brennenman, Jonathan Brown, Catherine Buell, Ginny and James Cook, Kathleen Iwancio, Chirag Lakhani, Shirley Law, Katie Liszewski, Terrance Pendleton, Megan Sawyer, Rob Stapleton, Ryan Therkelsen, Kyle Thompson, Robert Watson, Thomas Wears and Qin Zhang. Qunlei (Sarah) Jiang deserves special thanks for many reasons. I would like to thank Yi He and Hong Yin, two of my former roommates, for a wonderful time we shared together and for suggestions and advice from both of them. I would like to thank Ron and Tish Allison, and Paul Chen: they are there for me with a smile when I need help.

Last but definitely not least, I would like to thank Dr. Weigo Li, my advisor at UPC (Huadong), and Dr. Yongzhong Song, my advisor at Nanjing Normal University. What I have learned and gotten from them benefits me in my whole life.

# TABLE OF CONTENTS

<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Background and motivation	1
1.1.1 Real symmetric spaces	1
1.1.2 Symmetric varieties	3
1.1.3 Symmetric $\mathbb{F}$ -variety	3
1.1.4 Fixed point group and symmetric $\mathbb{F}$ -varieties	4
1.1.5 Orbit decompositions of symmetric $\mathbb{F}$ -varieties	4
1.1.6 Problems we are to solve	5
1.2 Summary of results in this dissertation	6
1.2.1 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(2, \mathbb{F})$	6
1.2.2 Classification of $G_{\mathbb{F}}$ -orbits of unipotent elements in $\mathrm{SL}(n, \mathbb{F})$	7
1.2.3 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(n, \mathbb{F})$ for inner involutions	7
1.2.4 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(n, \mathbb{F})$ for Cartan/outer involution	11
<b>Chapter 2 Preliminaries</b>	<b>13</b>
2.1 Symmetric Spaces and symmetric $\mathbb{F}$ -varieties	14
2.2 Isomorphism classes of involutions	14
2.3 Jordan decomposition	15
<b>Chapter 3 Double Coset and Reduction</b>	<b>17</b>
3.1 Correspondence between $K$ -orbits ( $K_{\mathbb{F}}$ -orbits) and double cosets	17
3.2 Reduction	18
<b>Chapter 4 Classification Results for <math>\mathrm{SL}(2, \mathbb{F})</math></b>	<b>21</b>
4.1 Reducing to the key involution case	21
4.2 Classification of $K_{\mathbb{F}}$ -orbits for the involution with $q \in (\mathbb{F}^*)^2$	22
4.3 Classification of $K_{\mathbb{F}}$ -orbits for the involution with $q \neq 0 \notin (\mathbb{F}^*)^2$	25
<b>Chapter 5 Classification of <math>G_{\mathbb{F}}</math>-orbits of Unipotent Elements for <math>\mathrm{SL}(n, \mathbb{F})</math></b>	<b>27</b>
5.1 A definition and its properties	27
5.2 Classification of $G_{\mathbb{F}}$ -orbits of unipotent elements for $\mathrm{SL}(n, \mathbb{F})$	29
<b>Chapter 6 Classification Results for <math>\mathrm{SL}(n, \mathbb{F})</math> for Inner Involutions</b>	<b>37</b>
6.1 Key inner involution	37
6.1.1 To make $\tau(g)$ to be unipotent, $g$ needs to have a special form	38
6.1.2 Classification with $g$ in the special form	54
6.1.3 Examples	93
6.2 The first type of inner involution with different sizes of $I_1$ and $I_2$	96
6.3 Second type of inner involution	97
6.3.1 $\{\}$ power function	97
6.3.2 Translation to key inner involution	101
6.3.3 Classification for the second type of inner involution	104

<b>Chapter 7 Cartan/Outer Involution</b>	<b>117</b>
7.1 Redo the case $n = 2$	118
7.2 Reduction	121
7.3 Ideas	123
7.4 Characteristic polynomial $\det(\lambda I + J)$	134
7.5 Case $n = 3$	137
7.6 Case $n = 4$	139
7.7 Some remarks	140
<b>References</b>	<b>142</b>



# Chapter 1

## Introduction

Symmetric spaces have been studied for over 100 years. Initially they were only studied over the real numbers, but in the last 25 to 30 years symmilar spaces over other fields have become of importance in other areas of mathematics as well. Symmetric  $\mathbb{F}$ -varieties generalize both real reductive symmetric spaces and symmetric varieties. In the following we will give a brief introduction and a summary of our results.

Again in this chapter, we are assuming: field  $\mathbb{F}$  is arbitrary field of characteristic of not 2; linear algebraic group  $G$  defined over  $\mathbb{F}$  is also called  $\mathbb{F}$ -group; an involution  $\sigma$  is automorphism on some algebraic group  $G$  such that  $\sigma^2 = \text{Id}$ , the identity operator;  $K = G^\sigma$  is the fixed point group of  $G$  with respect to  $\sigma$ ;  $G_{\mathbb{F}}$  (resp.  $K_{\mathbb{F}}$ ) is the set of  $\mathbb{F}$ -rational points of  $G$  (resp.  $K$ ).

### 1.1 Background and motivation

#### 1.1.1 Real symmetric spaces

We define and discuss briefly real symmetric spaces in this subsection.

**Definition 1.1.1.** Given an involution  $\sigma$  of a real reductive group  $G$ , the real reductive symmetric spaces  $P_{\mathbb{R}}$  is defined to be the homogeneous space  $G_{\mathbb{R}}/K_{\mathbb{R}}$ , where  $\mathbb{R}$  is the field of the real numbers.

The homogeneous space  $P_{\mathbb{R}}$  can be identified with the following subvariety of  $G_{\mathbb{R}}$ :

$$P_{\mathbb{R}} \simeq \{g(\sigma(g))^{-1} | g \in G_{\mathbb{R}}\}.$$

We note that over the real numbers we can take  $G_{\mathbb{R}}$  to be a reductive Lie group of Harish-Chandra class.

*Remark 1.1.2.* The real reductive symmetric space  $P_{\mathbb{R}}$  is also called an affine symmetric space. If moreover  $K$  is compact, then  $P$  is also called a Riemannian symmetric space.

We can also define symmetric spaces in other ways. In differential geometry, representation theory and harmonic analysis, a symmetric space is a smooth manifold whose group of symmetries contains an "inversion symmetry" about every point. There are two ways to make this precise, via Riemannian geometry or via Lie theory; the Lie theoretic definition is more general and more algebraic. In Riemannian geometry, the inversions are geodesic symmetries, and these are required to be isometries, leading to the notion of a Riemannian symmetric space.

Here's a brief explanation, assuming the manifold is 2 dimensional. A symmetric space means it is a smooth surface such that every point on the surface can serve as a point for reflection through a point, such that any shortest distance from two points on the surface, is still the same, before and after the reflection. Think of a 2-dimensional Euclidean plane. There we have the concept of reflection through a point. After the reflection, all distances are preserved. So it is a symmetric space. The case gets a bit more complex when the surface is not a flat plane, for example, if it's a sphere instead. A sphere is also a symmetric space. Every point on the sphere can serve as the point for reflection through a point. After the operation, every geodesics is preserved. That is, any 2 points, P, and G, the shortest distance between them on the surface, is the same, before and after, the operation.

Some other basic examples of Riemannian symmetric spaces are Euclidean space, spheres, projective spaces, and hyperbolic spaces, each with their standard Riemannian metrics. More examples are provided by compact, semi-simple Lie groups equipped with a bi-invariant Riemannian metric.

For geometric properties of real symmetric spaces, please see for example [15].

The representations associated with reductive real symmetric spaces (i.e., a decomposition of  $L^2(G_{\mathbb{R}}/K_{\mathbb{R}})$  into irreducible components) had been studied intensively by many prominent mathematicians, starting with a study of compact groups and their representations by Cartan in [9], to a study of Riemannian symmetric spaces and real Lie groups by Harish-Chandra in [38], to a more recent study of the non-Riemannian symmetric spaces starting in the 1970's by work of Faraut in [13], Flensted-Jensen in [14] and Oshima and Sekiguchi in [30]. These are soon studied by many mathematicians, including Brylinski, Carmona, Delorme, Matsuki, Oshima, Schlichtkrull, van den Ban and many others (see for example [35], [36], [37], [7], [8],[12], [29]). In the mid 1980's a Plancherel formula for the general real reductive symmetric spaces was announced by Oshima, although a full proof was not published until 1996 by Delorme in [12]. See also van den Ban and Schlichtkrull for a different approach in [36] and [37].

In the late 1980's it seemed natural to generalise the concept of these real reductive symmetric spaces to similar spaces over the  $p$ -adic numbers and study representations associated with these spaces. At the same time generalizations of these real symmetric spaces to other base fields started to play a role in other areas, including mathematical physics, Lie theory, representation theory and differential geometry, though their best known role lies in represen-

tation theory and harmonic analysis. For example, for geometry see [10] and [11], for singularity theory see [26] and [24]. This prompted Helminck and Wang to commence a study of rational properties of these homogeneous spaces over other fields, see [16] for some results.

### 1.1.2 Symmetric varieties

Symmetric varieties were introduced in the setting of geometry and invariant theory as a class of spherical varieties.

**Definition 1.1.3.** For a reductive algebraic group  $G$  defined over an algebraically closed field of characteristic not 2 and an involution  $\sigma$  of  $G$  a symmetric variety  $P$  is defined as the subvariety

$$P = \tau(G) = \{g(\sigma(g))^{-1} | g \in G\}.$$

It was shown by Vust [39] that  $P$  is isomorphic to  $G/K$ .

We note that with this definition every linear algebraic group itself is a symmetric variety.

*Example 1.1.4 (Group case).* Consider  $G_1 = G \times G$  and  $\theta(x, y) = (y, x)$ , then  $K = \{(x, x) | x \in G\} \approx G$  is embedded diagonally and  $P = \{x, x^{-1} | x \in G\} \approx G$  embedded anti-diagonally.

Also as we can see from the definition, involutions play an essential role in the theory of symmetric spaces and symmetric varieties.

### 1.1.3 Symmetric $\mathbb{F}$ -variety

Symmetric  $\mathbb{F}$ -varieties generalize both real symmetric spaces and symmetric varieties. In the following we assume that  $\sigma$  is an involution of  $G$  defined over  $\mathbb{F}$ , i.e., keep the rational group  $G_{\mathbb{F}}$  invariant. We note that these involutions are also called  $\mathbb{F}$ -involutions.

**Definition 1.1.5.** Given a field  $\mathbb{F}$ , an algebraic group  $G$ , an involution  $\sigma$  defined over  $\mathbb{F}$  and a fixed point group  $K = G^{\sigma}$  of involution  $\sigma$ , we define the symmetric  $\mathbb{F}$ -variety  $P$  as  $\tau(G_{\mathbb{F}}) = \{g(\sigma(g))^{-1} | g \in G_{\mathbb{F}}\}$ . Here  $G_{\mathbb{F}}$  is the set of rational points of  $G$ .

*Remark 1.1.6.* Similarly we can define real symmetric spaces and symmetric varieties in this way and show that  $P \simeq G_{\mathbb{F}}/K_{\mathbb{F}}$ .

Given  $g, x \in G$ , the twisted action associated to  $\sigma$  is given by  $(g, x) \rightarrow g * x = gx(\sigma(g))^{-1}$ . This is also called  $\sigma$ -twisted conjugation. We can see that actually real symmetric spaces, symmetric varieties, symmetric  $\mathbb{F}$ -varieties are actually defined using twisted action. Let  $P' = \{g \in G | \sigma(g) = g^{-1}\}$ . We can see that  $P \subset P'$ . Both  $P$  and  $P'$  are invariant under the twisted action associated to  $\sigma$ . There are only a finite number of twisted  $G$ -orbits in  $P'$  and each such orbit is closed, see [32]. In particular  $P'$  is a connected closed  $\mathbb{F}$ -subvariety of  $G$ . For more details about  $\mathbb{F}$ -variety see [17].

### 1.1.4 Fixed point group and symmetric $\mathbb{F}$ -varieties

For  $\mathbb{F} = \mathbb{R}$ , the Plancherel formula was first determined in the case of the Riemannian symmetric spaces. The main reason for this is that in this case the structure of the corresponding reductive symmetric spaces is relatively simple. For example all elements of  $P$  are semisimple and the left regular representation decomposition is multiplicative-free.

For  $\mathbb{F} = \mathbb{Q}_p$ , one gets a generalization of the real Riemannian symmetric space. The fixed point group  $K_{\mathbb{Q}_p} = G_{\mathbb{Q}_p}^\sigma$  determines much of the structure of the corresponding symmetric  $\mathbb{F}$ -variety.

Moreover if  $K$  is compact, then from [16] it follows that  $P$  consists of semisimple elements.

**Proposition 1.1.7** ([16], Proposition 10.8). *Let  $G$  be a connected reductive algebraic  $\mathbb{F}$ -group with  $\text{char}(\mathbb{F}) = 0$  and  $P = \{g(\sigma(g))^{-1} | g \in G\}$ . Suppose that  $K \cap [G, G]$  is anisotropic over  $\mathbb{F}$ . Then  $P$  consists of semisimple elements.*

We note that for  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{Q}_p$  all  $\mathbb{F}$ -anisotropic subgroups are compact.

### 1.1.5 Orbit decompositions of symmetric $\mathbb{F}$ -varieties

The introduction here is from [19]. Orbit decompositions for symmetric  $\mathbb{F}$ -varieties play a fundamental role in the study of symmetric  $\mathbb{F}$ -varieties ( $\mathbb{F}$  is the field of characteristic not 2) and their applications to representation theory and many other areas of mathematics, such as geometry, the study of automorphic forms and character sheaves. In [19] Helminck studied orbit decomposition of symmetric  $\mathbb{F}$ -varieties in mainly following categories: orbits of parabolic  $\mathbb{F}$ -subgroups, the orbits of symmetric subgroups and Euclidean building.

The orbits of a parabolic  $\mathbb{F}$ -subgroup  $Q$  acting on the symmetric  $\mathbb{F}$ -variety  $G_{\mathbb{F}}/K_{\mathbb{F}}$  play a fundamental role in the study of representations associated with these symmetric spaces. These orbits were studied for many fields and can be characterized in several equivalent way. They can be characterized as the  $Q_{\mathbb{F}}$ -orbits acting on the symmetric  $\mathbb{F}$ -variety  $G_{\mathbb{F}}/K_{\mathbb{F}}$  by  $\sigma$ -twisted conjugation, or as the  $K_{\mathbb{F}}$ -orbits acting on the flag variety  $G_{\mathbb{F}}/K_{\mathbb{F}}$  by conjugation, or also as the set  $Q_{\mathbb{F}} \backslash G_{\mathbb{F}}/K_{\mathbb{F}}$  of  $(Q_{\mathbb{F}}, K_{\mathbb{F}})$ -double cosets in  $G_{\mathbb{F}}$ . The last one is same as the set of  $Q_{\mathbb{F}} \times K_{\mathbb{F}}$ -orbits on  $G_{\mathbb{F}}$ . For  $\mathbb{F}$  algebraically closed, and  $Q = B$  a Borel subgroup, these orbits were characterized by Springer [33] and a characterization of these orbits of general parabolic subgroups was given by Brion and Helminck in [5] and [18].

If the field  $\mathbb{F}$  is algebraically closed, the symmetric  $\mathbb{F}$ -varieties become symmetric varieties and we can consider the orbits of symmetric subgroups. The orbits of symmetric subgroups acting on symmetric varieties are of importance in representation theory. For  $K$  acting on the symmetric variety  $G/K$  those orbits were studied by Vust [39] and Richardson in [32], and for an arbitrary symmetric subgroup  $H$  acting on a symmetric variety these orbits were studied by

Helminck and Schwarz in [20]. In [19] Helminck also studied the case for two involutions of  $\sigma$  and  $\theta$ .

The analog of the Riemannian symmetric spaces for  $p$ -adic groups are the Euclidean buildings. Bruhat and Tits showed in [6] that most properties of the Riemannian symmetric spaces carry over to these Euclidean buildings. For example they showed that a compact group of isometries of a Euclidean building has a fixed point. They used this to show that in a simply connected  $p$ -adic group the maximal compact subgroups are parabolic subgroups. There are many other similarities between the Riemannian symmetric spaces and the Euclidean buildings. For example for both a geodesic joining two points is unique. There is also a  $p$ -adic curvature on the building. The Euclidean building also replaces the role of the Riemannian symmetric spaces in the cohomology of discrete subgroups.

### 1.1.6 Problems we are to solve

The problem we are mainly working on is part of  $K_{\mathbb{F}}$ -orbit decomposition of symmetric  $\mathbb{F}$ -varieties. We try to classify  $K_{\mathbb{F}}$ -orbits of unipotent elements in symmetric  $\mathbb{F}$ -variety  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  where  $K = G^{\sigma}$  is the fixed-point subgroup of  $G$ ,  $G_{\mathbb{F}}$  (resp.  $K_{\mathbb{F}}$ ) is the set of rational points of  $G$  (resp.  $K$ ), and  $\sigma$  an involution acting on  $G$ .

Using results from Lusztig [27] and Richardson [31] Richardson gave a finiteness result of  $K$ -orbits of unipotent elements in the symmetric space in [32] in 1982:

**Proposition 1.1.8** ([32], Proposition 7.4). *Let the field  $\mathbb{F}$  be algebraically closed, i.e.,  $\mathbb{F} = \overline{\mathbb{F}}$  and Characteristic of  $\mathbb{F}$  not be 2. Let  $\mathcal{U}(P)$  denote the set of unipotent elements in  $P = G/K$ . Then there are only a finite number of  $K$ -orbits of unipotent elements in  $\mathcal{U}(P)$ .*

A question arises naturally: what if the field  $\mathbb{F}$  is not algebraically closed? What's more, given an algebraic group and an involution, can we list those  $K$ -orbits (or  $K_{\mathbb{F}}$ -orbits if  $\mathbb{F} \neq \overline{\mathbb{F}}$ ) explicitly, i.e., do we have some canonical form for representatives of  $K_{\mathbb{F}}$ -orbits (resp.  $K_{\mathbb{F}}$ -orbits)?

These are open questions and we have one more conjecture: using results shown in this thesis, we can see that, for certain involutions and algebraic groups, if we work over fields  $\mathbb{R}$  and  $\mathbb{Q}_p$ , we still have a finite number of  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P$ , now, we want to ask: is this still true for all involutions and all algebraic groups?

For example, for rational field  $\mathbb{Q}$  we can definitely have infinitely  $K_{\mathbb{Q}}$ -orbits of unipotent elements in  $P$  (see Theorem 4.2.1): if we work on  $\mathrm{SL}(2, \mathbb{Q})$  and the involution is defined by  $\sigma(g) = p^{-1}gp$ , where  $p = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$ , then we will have  $2 \cdot |\mathbb{Q}^*/(\mathbb{Q}^*)^2| + 1$   $K$ -orbits of unipotent elements in  $P$  and we can particularly choose following elements

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

as representative elements in  $K_{\mathbb{Q}}$ -orbits, where  $\alpha$  is a representative of  $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ .

We answer the above questions for the special linear group  $\mathrm{SL}(n, \mathbb{F})$  for certain involutions. For inner involutions the finiteness of  $K$ -orbits of unipotent elements in  $P$  (or  $\mathcal{U}(P)$ ) depends on the cardinality of  $\mathbb{F}^*/(\mathbb{F}^*)^h$ ,  $\mathbb{F}_{\pm}^*/(\mathbb{F}_{\pm}^*)^h$  or  $\mathbb{F}^*(\sqrt{p})/(\mathbb{F}^*(\sqrt{p}))^{\{h\}}$  (see following chapters for definition of symbols here), where integer  $h$  is determined by sizes of Jordan blocks of a certain submatrix in  $g$ , and the "curly bracket" power function is defined in a way similar to the usual power function with conjugation involved. For outer involutions we have not done yet.

## 1.2 Summary of results in this dissertation

We first study what kind of elements could generate unipotent elements in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ , i.e., what kind of elements we need to study. Then we classify all  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P$  for  $G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$  for all inner involutions. According to classifications of involution in [22] and [23], we only need to study two types inner involutions, and we here list our results for inner involution accordingly. For outer involution, we study Cartan involution. We give out the  $K$ -orbits of small  $n$  ( $=2,3,4$ ) and illustrate how to get the canonical form for a general one. We have the correspondence between classification of  $K_{\mathbb{F}}$ -orbits and classification of  $(K_{\mathbb{F}}, K_{\mathbb{F}})$ -double cosets. We classify  $G_{\mathbb{F}}$ -orbits too.

### 1.2.1 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(2, \mathbb{F})$

To study the inner involution given by  $\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$ , we can study equivalently the inner involution defined by  $\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$  instead. From Theorem 4.2.1, we can see that we have  $2 \cdot |\mathbb{F}^*/(\mathbb{F}^*)^2| + 1$   $K_{\mathbb{F}}$ -orbits of unipotent elements in the symmetric space  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ . We can especially choose

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

as representative elements in  $K_{\mathbb{F}}$ -orbits, where  $\alpha$  are representatives of  $\mathbb{F}^*/(\mathbb{F}^*)^2$ .

For the inner involution given by  $\begin{bmatrix} & 1 \\ p & \end{bmatrix}$ , we have the identity orbit only (Theorem 4.3.1).

### 1.2.2 Classification of $G_{\mathbb{F}}$ -orbits of unipotent elements in $\mathrm{SL}(n, \mathbb{F})$

$G_{\mathbb{F}}$ -orbit of some element  $A \in G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$  is the set  $\{g^{-1}Ag | g \in G_{\mathbb{F}}\} = \{gAg^{-1} | g \in G_{\mathbb{F}}\}$ .

For classification of  $G_{\mathbb{F}}$ -orbits in  $\mathrm{SL}(n, \mathbb{F})$  we have from Theorem 5.2.1:

Let  $\mathbb{F}$  be an arbitrary field of characteristic of not 2. Suppose we have two unipotent elements  $A = XJ_AX^{-1}$  and  $B = YJ_BY^{-1}$  in  $G_{\mathbb{F}}$  (we can always have the decomposition since  $A$  and  $B$  are both unipotent) with same Jordan canonical form determined by partition (5.2.1):  $J_A = J_B$ . Then  $A$  and  $B$  is in the same  $G_{\mathbb{F}}$ -orbit, i.e., there exists  $P \in G_{\mathbb{F}}$  such that  $B = PAP^{-1}$ , if and only if  $\det(X)$  and  $\det(Y)$  are  $n_i\delta(\beta_i)$ -equivalent, or equivalently  $\det(X)/\det(Y) \in (\mathbb{F}^*)^h$ , where  $h$  is the greatest common factor of the sequence  $n_i\delta(\beta_i), i = 0, 1, \dots, r$ .

Furthermore, all  $G_{\mathbb{F}}$ -orbits can choose elements in the following form as representatives

$$XJX^{-1},$$

where  $J$  is the Jordan canonical form and

$$X = \begin{bmatrix} p & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$ .

### 1.2.3 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(n, \mathbb{F})$ for inner involutions

#### Key inner involution

We first study the key inner involution given by  $\begin{bmatrix} I & \\ & -I \end{bmatrix}$ . We show that, to classify  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  for  $\mathrm{SL}(n, \mathbb{F})$ , we only need to study following problem:

Let

$$K_{\mathbb{F}} = \{k = \begin{bmatrix} x & \\ & y \end{bmatrix} \in \mathrm{SL}(n, \mathbb{F}) | \det(k) = 1\}$$

$$J = \begin{bmatrix} & B \\ C & \end{bmatrix}, \bar{J} = \begin{bmatrix} & \bar{B} \\ \bar{C} & \end{bmatrix},$$

what kind of conditions should  $J$  and  $\bar{J}$  satisfy to have a  $k \in K_{\mathbb{F}}$  such that  $\bar{J} = kJk^{-1}$ ? Or

equivalently, if we use  $k$  to make  $J$  simple, what kind of simple/canonical form we could have for  $J$ ?

If  $C$  is nonsingular, ie  $\text{rank}(C) = \frac{n}{2}$ , we then have from Theorem 6.1.8:

Let

$$J = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}, \bar{J} = \begin{bmatrix} 0 & \bar{B} \\ \bar{C} & 0 \end{bmatrix},$$

be nilpotent matrices.

If  $\text{rank}(C) = \text{rank}(\bar{C}) = \frac{n}{2}$ , i.e.,  $\det(C) \cdot \det(\bar{C}) \neq 0$ , then

(I) If  $\mathbb{F} = \bar{\mathbb{F}}$ , then  $J \stackrel{K_{\mathbb{F}}}{\sim} \bar{J}$  if and only if  $\overline{BC}$  is similar to  $BC$  in the usual sense, or equivalently  $J^2 \stackrel{K_{\mathbb{F}}}{\sim} \bar{J}^2$ ;

(II) If  $\mathbb{F} \neq \bar{\mathbb{F}}$ , then  $J \stackrel{K_{\mathbb{F}}}{\sim} \bar{J}$  if and only if following two conditions are satisfied

1.  $\det(C)/\det(\bar{C}) \in (\mathbb{F}^*)^2$ ;

2. using condition 1 and approach I we can make  $\det(C) = \det(\bar{C})$ . Then  $BC$  and  $\overline{BC}$  are nilpotent and satisfy conditions in Theorem 5.2.4.

It's much different if  $C$  is singular. Suppose  $\text{rank}(C) = m < \frac{n}{2}$ . Using  $x$  and  $y$  we are able to make  $C = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ . In this case we have

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \bar{B} = \begin{bmatrix} \bar{B}_{11} & \bar{B}_{12} \\ \bar{B}_{21} & \bar{B}_{22} \end{bmatrix}$$

$$x = \begin{bmatrix} \alpha & 0 \\ x_{21} & x_{22} \end{bmatrix}$$

$$y = \begin{bmatrix} \alpha & y_{12} \\ 0 & y_{22} \end{bmatrix},$$

and  $B, x$  and  $y$  are partitioned accordingly. Using  $x$  and  $y$  we are able to make  $B_{12}, B_{21}$  and  $\bar{B}_{12}, \bar{B}_{21}$  into some special form. With  $B_{12}, B_{21}$  and  $\bar{B}_{12}, \bar{B}_{21}$  in certain special form we have from Theorem 6.1.21:

If  $\mathbb{F} = \bar{\mathbb{F}}$ , and if  $B$  and  $\bar{B}$  are in pre-described forms have solutions, we can find  $k \in K_{\mathbb{F}}$  such



that  $\bar{J} = kJk^{-1}$  if and only if following conditions hold for  $1 \leq i \leq r$ :

$$\begin{aligned}\text{rank}(B_{12}) &= \text{rank}(\bar{B}_{12}) \\ \text{rank}(B_{21}) &= \text{rank}(\bar{B}_{21}) \\ \text{rank}(b^{00}) &= \text{rank}(\bar{b}^{00}) \\ \text{rank}(b^i) &= \text{rank}(\bar{b}^i) \\ \text{rank}(c_j^i) &= \text{rank}(\bar{c}_j^i), j = 0, 1, 2, \\ \text{rank}(c^{00}) &= \text{rank}(\bar{c}^{00})\end{aligned}$$

and the above six conditions are numbered as (6.1.14)-(6.1.19).

Here  $b^{00}, \bar{b}^{00}, b^i, \bar{b}^i, c_j^i, \bar{c}_j^i, j = 0, 1, 2, c^{00}$  and  $\bar{c}^{00}$  come from the pre-described forms.

If the field is not algebraically closed, we have to add other conditions. In this case we have Theorem 6.1.25:

(I) If  $\mathbb{F} \neq \bar{\mathbb{F}}$ , all conditions (6.1.22), (6.1.23) and (6.1.25) are satisfied, if  $B_{12}$  and  $B_{21}$  are not square matrices, then all  $K_{\mathbb{F}}$ -orbits can choose following as representatives

$$c_0^i = \left[ \begin{array}{c|c} 0 & \\ \hline p & \\ & 1 \\ & \ddots \\ & 1 \\ \hline & 0 \end{array} \right] \text{ or } c_2^i = \left[ \begin{array}{c|c} 0 & 0 \\ \hline p & \\ & 1 \\ 0 & \ddots \\ & 1 \\ \hline 0 & 0 \end{array} \right], c_1^i = 0 \text{ or does not exist}$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence  $S1$ .

(II) If  $\mathbb{F} \neq \bar{\mathbb{F}}$ , all conditions (6.1.22), (6.1.23) and (6.1.26) are satisfied, if  $B_{12}$  and  $B_{21}$  are square matrices, then all  $K_{\mathbb{F}}$ -orbits can choose following as representatives

$$c_1^i = \left[ \begin{array}{c|c} 0 & 0 \\ \hline p & \\ & 1 \\ & \ddots \\ & 1 \\ \hline 0 & 0 \end{array} \right], c_2^i = 0$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence

$S2$ .

(III) If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , all conditions (6.1.22), (6.1.23), (6.1.25) and (6.1.26) are satisfied (there are some blocks in  $b_i$  satisfying (6.1.25) and some satisfying (6.1.26)), then all  $K_{\mathbb{F}}$ -orbits can choose as representatives from formula (6.1.27) and (6.1.28), where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence  $S3$ .

(IV) In all cases except the above three, all conditions in Theorem 6.1.21 have been good enough to serve as necessary and sufficient conditions.

Here sequences  $S1, S2, S3$  are defined in Subsection 6.1.2.

### The first type of inner involution with different sizes of identities

The case for the inner involution given by  $\begin{bmatrix} I_1 & \\ & I_2 \end{bmatrix}$  with different sizes of  $I_1$  and  $I_2$  can embed into the above case with a larger even integer.

### The second type of inner involutions

For classification of  $K_{\mathbb{F}}$ -orbits for the inner involution given by  $\begin{bmatrix} 1 & & \\ p & & \\ & \ddots & \\ & & 1 \\ & & & p \end{bmatrix}$ , we

show that we can study following problem equivalently:

Given  $J = \begin{bmatrix} \overline{D} \\ D \end{bmatrix}$ , using matrices from  $K_{\mathbb{F}} = \{k = \begin{bmatrix} x \\ \overline{x} \end{bmatrix}\}$ , what kind of simple/canonical form of  $J$  we can have? Theorem 6.3.8 answers this question:

We can choose elements  $g = I + J$  as representatives of  $K$ -orbits, where

$$J = \begin{bmatrix} \overline{D} \\ D \end{bmatrix} \text{ with } D = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix},$$

and  $J_i, i = 1, \dots, r$ , satisfy following conditions

1. The sizes of  $J_i$  are decreasing, especially if  $J_{i_0} = 0$ , then  $J_{i_0+1} = \dots = J_r = 0$ ;
2. If  $J_i \neq 0$ , then

$$J_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & q \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

where  $q = 1$  if  $1 \leq i \leq r - 1$ . If  $i = r$  ( $D$  has no zero Jordan block in this case), then  $q$  is representative element of  $\mathbb{F}^*(\sqrt{p})/(S^1 * (\mathbb{F}^*(\sqrt{p}))^{\{l\}})$  with  $\{\}$  being the "curly bracket" power function defined in Definition 6.3.1, and  $l$  is the greatest common factor of  $n_i$ ,  $n_i$  are from the partition of positive integer  $m = \frac{n}{2}$ :

$$m = \beta_1 n_1 + \cdots + \beta_s n_s.$$

which corresponds to the Jordan blocks  $J_i$  of  $D$ .

Given an integer  $h$  and a number  $x \in \mathbb{F}^*(\sqrt{p})$ , the  $\{\}$  power function in the above theorem is defined as  $x^{\{h\}} = \overbrace{x\bar{x}x\bar{x}\cdots}^h$ .

#### 1.2.4 Classification of $K_{\mathbb{F}}$ -orbits for $\mathrm{SL}(n, \mathbb{F})$ for Cartan/outer involution

We have not finished this part yet. We give out the canonical forms of  $K$ -orbits of unipotent elements in  $P = G/K$  over algebraically closed field  $\mathbb{F}$ , which consists of symmetric matrices, for small  $n$  only. We illustrate what kind of canonical form is expected for arbitrary  $n$ .

Representatives in  $K$ -orbits can be given as  $g = I + J$  with

- $n = 2$ :  $J = 0$  (the identity orbit)

$$J = \begin{bmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & -1 \end{bmatrix}, J = \begin{bmatrix} 1 & -\sqrt{-1} \\ -\sqrt{-1} & -1 \end{bmatrix}$$

- $n = 3$ :  $J = 0$  (the identity orbit)

$$J = \begin{bmatrix} 1 & \sqrt{-1} & 0 \\ \sqrt{-1} & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ (from the case } n = 2)$$

$$J = \begin{bmatrix} 0 & \sqrt{-1} & 0 \\ \sqrt{-1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

•  $n = 4$ :  $J = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 2\sqrt{-1} & 0 \\ 0 & 2\sqrt{-1} & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$   
and all orbits included in the case  $n = 3$ .

The canonical form for any  $n$  will be all  $K$ -orbits included in the case  $n - 1$  and one two  $K$ -orbits of rank  $n - 1$  in the following form:

$$J = \begin{bmatrix} \alpha & y_1 & & & \\ y_1 & -\alpha & y_2 & & \\ & y_2 & \alpha & y_3 & \\ & & y_3 & -\alpha & y_4 \\ & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

$\alpha = 0$  if  $n$  is odd and  $\alpha \neq 0$  if  $n$  is even. More specifically,  $\alpha$  should be able to take 1 if  $n$  is even.

If  $n = 2m + 1$  is odd, we should always have only one  $K$ -orbit of unipotent elements with rank  $n - 1$ , but we will have two  $K$ -orbits of unipotent elements since we have  $n - 1$  off-diagonals and are only able to introduce even numbers of  $-1$ 's. All  $K$ -orbits will be the ones of rank  $n - 1$  and the ones included in  $n - 1$  case. If  $n$  is odd, then only one  $K$ -orbit of rank  $n - 2$  from the  $K$ -orbits of  $n - 1$  case is included.

A good intuition is that all  $K$ -orbits here correspond to the Jordan canonical forms exactly.

## Chapter 2

# Preliminaries

Our basic references for reductive groups, Borel subgroups, and maximal tori will be the books of Borel [4], Humphreys [25], and Springer [34]. We will follow their notations and terminology.

Throughout this thesis,  $\mathbb{F}$  is a field of characteristic not 2, and  $G$  is a reductive algebraic group defined over  $\mathbb{F}$ . Let  $\sigma$  be an involution of  $G$  and  $K = G^\sigma$  be the fixed point group of  $G$ ,  $G_{\mathbb{F}}$  (resp.  $K_{\mathbb{F}}$ ) the set of rational points of  $G$  ( $K$ ).

Let  $V = \mathbb{F}^n$  be a finite dimensional vector space defined over  $\mathbb{F}$ ,  $M_n(\mathbb{F}) = M(n, \mathbb{F})$  the set of  $n \times n$  matrices with entries in  $\mathbb{F}$ , and  $\text{Id} \in \text{Aut}(G)$  the identity automorphism. For the identity matrix we often simply write  $I$  instead of  $\text{Id}$ .

Let

$$\text{GL}(n, \mathbb{F}) = \{A \in M_n(\mathbb{F}) \mid \det(A) \neq 0\}$$

and

$$\text{SL}(n, \mathbb{F}) = \{A \in \text{GL}(n, \mathbb{F}) \mid \det(A) = 1\}$$

Let  $\mathbb{F}^*$  denote the product group of all the nonzero elements. Given any positive integer  $h \in \mathbb{Z}^+$ , let

$$(\mathbb{F}^*)^h = \{a^h \mid 0 \neq a \in \mathbb{F}\}$$

The following is some notation we will use throughout the rest of this paper.

*Notation 2.0.1.* By  $m \in \mathbb{F}^*/(\mathbb{F}^*)^h$  we mean that  $m$  is the representative of the  $h$ -quotient group of  $\mathbb{F}^*$ , where  $h$  is any positive integer.

*Notation 2.0.2.* Given a matrix  $A$ ,  $A^T$  is the transpose of  $A$ .

## 2.1 Symmetric Spaces and symmetric $\mathbb{F}$ -varieties

Let  $\text{Aut}(G)$  denote the set of all automorphisms of  $G$ . We have following definitions.

**Definition 2.1.1.** An *involution* of  $G$  is an automorphism  $\sigma \in \text{Aut}(G)$  such that  $\sigma^2 = \text{Id}$  and  $\sigma \neq \text{Id}$ .

**Definition 2.1.2.** Given an involution  $\sigma$  of  $G$ ,  $K = G^\sigma$  the fixed point group of the involution  $\sigma$ , the *symmetric  $\mathbb{F}$ -variety*  $P_{\mathbb{F}}$  is defined as  $\tau(G_{\mathbb{F}}) = \{g\sigma(g)^{-1} | g \in G_{\mathbb{F}}\}$ , where  $G_{\mathbb{F}}$  is the set of  $\mathbb{F}$ -rational points of  $G$ . One can see that  $P \simeq G_{\mathbb{F}}/K_{\mathbb{F}}$ , where  $K_{\mathbb{F}}$  is the set of  $\mathbb{F}$ -rational points of  $K$ . For  $\mathbb{F} = \overline{\mathbb{F}}$  algebraically closed,  $P$  is also called a symmetric variety, and is the connected component containing the identity of  $Q = \{x \in G | \sigma(x) = x^{-1}\}$  if the connectedness is defined on  $G$ .

**Definition 2.1.3.** Given an involution  $\sigma$  of  $G$  and subgroup  $H$  of  $G$ , the  *$H$ -orbit* of some element  $g \in G$  is defined as the set  $\{h^{-1}gh | h \in H\} = \{hgh^{-1} | h \in H\}$ . Especially if  $H = K$  is the fixed point group of  $G$ , the  $K$ -orbit of  $g \in G$  is to be  $\{k^{-1}gk | k \in K\} = \{kgk^{-1} | k \in K\}$ . Similarly for arbitrary field  $\mathbb{F}$ , and  $H_{\mathbb{F}} = H \cap G_{\mathbb{F}}$  we define  *$H_{\mathbb{F}}$ -orbit* of some element  $g \in G_{\mathbb{F}}$  as the set  $\{h^{-1}gh | h \in H_{\mathbb{F}}\} = \{hgh^{-1} | h \in H_{\mathbb{F}}\}$ . Especially if  $H_{\mathbb{F}} = K_{\mathbb{F}}$ , the  $K_{\mathbb{F}}$ -orbit of  $g \in G_{\mathbb{F}}$  is to be  $\{k^{-1}gk | k \in K_{\mathbb{F}}\} = \{kgk^{-1} | k \in K_{\mathbb{F}}\}$ .

**Definition 2.1.4.** A matrix  $J$  is *nilpotent* if  $J^m = 0$  for some positive integer  $m$ . It's clear that  $J$  is nilpotent if and only if all eigenvalues of  $J$  are 0. A matrix  $g = I + J$  is *unipotent* if  $J$  is nilpotent, where  $I$  is the identity matrix.

## 2.2 Isomorphism classes of involutions

To classify  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  for all involutions on  $G_{\mathbb{F}}$ , we need first to classify involutions up to isomorphism. Before we define what we mean by isomorphism of involutions we need more notations.

For  $A \in \text{GL}(n, \mathbb{F})$ , let  $\text{Int}(A) = \mathcal{I}_A$  denote the automorphism defined by  $\mathcal{I}_A(X) = A^{-1}XA$ ,  $X \in \text{GL}(n, \mathbb{F})$ . An automorphism  $\Phi$  of  $G$  is called *of inner type* if  $\Phi = \mathcal{I}_A$  for some  $A \in \text{GL}(n, \mathbb{F})$ . Otherwise  $\Phi$  is called *of outer type*. Let  $\text{Int}_{\mathbb{F}}(G) = \{\text{Int}(x) | x \in G \text{ and } \text{Int}(G)(G_{\mathbb{F}}) \subseteq G_{\mathbb{F}}\}$  denote the set of inner automorphisms of  $G$  which keep  $G_{\mathbb{F}}$  invariant. Note that for  $G_{\mathbb{F}} = \text{SL}(n, \mathbb{F})$  one can consider conjugation by elements of  $\text{GL}(n, \overline{\mathbb{F}})$  instead of conjugation by elements of  $\text{SL}(n, \overline{\mathbb{F}})$ .

**Definition 2.2.1.**  $\sigma, \phi \in \text{Aut}(G)$  are said to be  $\mathbb{F}$ -conjugate or  $\mathbb{F}$ -isomorphic if and only if there is a  $\chi \in \text{Int}_{\mathbb{F}}(G)$  such that  $\chi^{-1}\sigma\chi = \phi$ . When the field  $\mathbb{F}$  is clear from the context, we simply say that they are conjugate or isomorphic.

We need following results on classification of involutions on  $G_{\mathbb{F}} = \text{SL}(n, \mathbb{F})$ :

**Theorem 2.2.2** ( $n = 2$ , [22]). *All the  $\mathbb{F}$ -isomorphism classes of involutions over  $G_{\mathbb{F}} = \text{SL}(2, \mathbb{F})$  can be represented by  $\text{Int}(X)$ , where  $X = \begin{bmatrix} 0 & 1 \\ p & 0 \end{bmatrix} \in \text{GL}(n, \mathbb{F})$ ,  $p \in \mathbb{F}^*/(\mathbb{F}^*)^2$ .*

**Theorem 2.2.3** ( $n \geq 3$ , inner, [23]). *Suppose the involution  $\sigma \in \text{Aut}(G)$  is of inner type. Then up to  $\mathbb{F}$ -isomorphism, we can write  $\sigma = \text{Int}(Y)$ , where  $Y$  is one of the following matrices:*

(1)  $\begin{bmatrix} I_1 & & \\ & -I_2 & \\ & & \end{bmatrix}$ ,  $I_1$  and  $I_2$  are identity matrices, and the size of  $I_1$  can be  $1, 2, \dots, \lfloor \frac{n}{2} \rfloor$ .  $\lfloor \frac{n}{2} \rfloor$  is the largest integer that is less than or equal to  $\frac{n}{2}$ ;

(2)  $\begin{bmatrix} 1 & & \\ p & & \\ & \ddots & \\ & & \ddots & \\ & & & 1 \\ & & & & p \end{bmatrix}$ ,  $p \in \mathbb{F}^*/(\mathbb{F}^*)^2$ .

**Theorem 2.2.4** ( $n \geq 3$ , outer, [23]). *Let  $\phi$  be the Cartan involution, i.e.,  $\phi(g) = (g^T)^{-1}$ . Then up to  $\mathbb{F}$ -isomorphism, any outer involution  $\sigma$  can be written as  $\sigma = \text{Int}(M) \circ \phi$ , where  $M$  is one of the following matrices:*

(1)  $M = \begin{bmatrix} m_1 & & \\ & \ddots & \\ & & m_n \end{bmatrix}$ ,  $m_i \in \mathbb{F}^*/(\mathbb{F}^*)^2$ ;

(2)  $M = \begin{bmatrix} & I \\ -I & \end{bmatrix}$ ,  $I$  is the identity matrix of size  $\frac{n}{2}$ . It's clear that this is possible only if  $n$  is an even integer.

## 2.3 Jordan decomposition

Jordan decomposition gives us nice canonical forms for arbitrary matrices, which is very useful in linear algebras theory. Here we sue following form of Jordan decomposition (see, for example, [1]).

If  $\mathbb{F}$  is an algebraically closed field, i.e.,  $\mathbb{F} = \overline{\mathbb{F}}$ , it's clear that, given a matrix  $A \in M_n(\mathbb{F})$ , we can find  $X \in \text{GL}(n, \mathbb{F})$  such that

$$X^{-1}AX = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix},$$

where

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix} = \lambda_i I + \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Since  $\mathbb{F} = \overline{\mathbb{F}}$ , we can actually take  $X \in \mathrm{SL}(n, \mathbb{F})$  in above multiplications. This is not true any more if  $\mathbb{F} \neq \overline{\mathbb{F}}$ . We give corresponding result in Theorem 5.2.1. In Chapter 6, We actually reduce a matrix pair (or matrix pencil)  $(B, C)$  simultaneously, i.e.,  $B \rightarrow x^{-1}By, C \rightarrow y^{-1}Cx$ , and we also consider the case of  $D \rightarrow (\overline{x})^{-1}Dx$ .

*Remark 2.3.1.* Even if  $\mathbb{F} \neq \overline{\mathbb{F}}$ , if  $A$  is unipotent or nilpotent, we still have the above Jordan decomposition, though one can only assume  $X \in \mathrm{GL}(n, \mathbb{F})$ .



## Chapter 3

# Double Coset and Reduction

There is a correspondence between  $K$ -orbits ( $K_{\mathbb{F}}$ -orbits) and double cosets, and one way to classify  $K$ -orbits ( $K_{\mathbb{F}}$ -orbits) is to classify corresponding double cosets. In this chapter we also verify the corresponding classification results for two isomorphic involutions, which allows us to use results from [22] and [23] to work on some specific involutions only.

### 3.1 Correspondence between $K$ -orbits ( $K_{\mathbb{F}}$ -orbits) and double cosets

We first need to define formally what we mean by double coset:

**Definition 3.1.1.** Given two sets  $K_1, K_2 \in G$  (resp.  $G_{\mathbb{F}}$ ) and  $g \in G$  (resp.  $G_{\mathbb{F}}$ ), we define the  $(K_1, K_2)$ -coset of  $g$  (double coset) as

$$K_1 g K_2 = \{k_1 g k_2 | k_1 \in K_1, k_2 \in K_2\}.$$

Especially  $(K_{\mathbb{F}}, K_{\mathbb{F}})$ -coset is

$$K_{\mathbb{F}} g K_{\mathbb{F}} = \{k_1 g k_2 | k_1, k_2 \in K_{\mathbb{F}}\}.$$

The following result from [4] gives us an intuition of types of involutions we should study:

**Lemma 3.1.2.** 1. If the field  $\mathbb{F}$  is algebraically closed, then  $|\text{Aut}(G)/\text{Int}(G)| = 2$  for  $n \geq 3$ .

2. Any outer automorphism can be written as  $\text{Int}(M) \circ \phi$ , where  $\phi$  is a fixed outer automorphism.

*Remark 3.1.3.* Cartan involution  $\phi$  defined by  $\phi(g) = (g^T)^{-1}$  is an outer involution and actually it's the most important outer one.

Following calculations are trivial:

$$\begin{aligned}
\tau(g) &= \tau(\bar{g}) \\
g\sigma(g)^{-1} &= \bar{g}\sigma(\bar{g})^{-1} \\
I &= g^{-1}\bar{g}\sigma(\bar{g})^{-1}\sigma(g) \\
I &= (g^{-1}\bar{g})\sigma(g^{-1}\bar{g})^{-1} \\
&\Leftrightarrow g^{-1}\bar{g} \in K_{\mathbb{F}} \\
&\Leftrightarrow \bar{g} = gk \text{ for some } k \in K_{\mathbb{F}}.
\end{aligned}$$

This actually gives us a quick look of why  $P \simeq G_{\mathbb{F}}/K_{\mathbb{F}}$ . See [16] for a strict proof. For all possible  $k_1, k_2 \in K_{\mathbb{F}}$ ,  $k_1 g k_2$  are always in the same  $K_{\mathbb{F}}$ -orbit. It's true vice versa. This is actually following theorem:

**Theorem 3.1.4.** *Given  $g, \bar{g} \in G_{\mathbb{F}}$ ,  $\tau(g)$  and  $\tau(\bar{g})$  are in the same  $K_{\mathbb{F}}$ -orbit if and only*

$$K_{\mathbb{F}}gK_{\mathbb{F}} = K_{\mathbb{F}}\bar{g}K_{\mathbb{F}}.$$

*Remark 3.1.5.* For any  $g, \bar{g} \in G_{\mathbb{F}}$  we have either  $K_{\mathbb{F}}gK_{\mathbb{F}} = K_{\mathbb{F}}\bar{g}K_{\mathbb{F}}$  or  $K_{\mathbb{F}}gK_{\mathbb{F}} \cap K_{\mathbb{F}}\bar{g}K_{\mathbb{F}} = \phi$ , the empty set.

*Remark 3.1.6.* Properties and decompositions/classifications of double cosets have been studied over years. Bala and Carter have studied double cosets of unipotent elements in algebraic groups in [2], [3], which gave us Bala-Carter theorem. Matsuki studied double coset decompositions of reductive Lie groups arising from two involutions in [28]. Richardson studied  $(K_1, K_2)$  double cosets on  $G$  in [32], where  $K_1$  and  $K_2$  are some subgroups of  $G$  containing the identity component of  $K$ . Helminck and Schwarz considered real double cosets of semisimple elements in [21].

## 3.2 Reduction

Let  $\theta, \sigma$  be two involutions with  $\theta = \theta_p^{-1} \circ \sigma \circ \theta_p$ ,  $\theta_p$  is an inner involution defined by  $\theta_p(g) = p^{-1}gp$ , where  $p \in \text{GL}(n, \mathbb{F})$ ,  $g \in G$ , and  $G$  is some linear algebraic group. What we will show is that, to study properties of pair  $(G, \theta)$ , we can equivalently study  $(\theta_p(G), \sigma)$ .

It's clear that  $\theta_p(G)$  is still a linear algebraic group and  $\sigma$  is an involution on  $\theta_p(G)$ . Let

$\tau_\theta(g) = g \cdot \theta(g^{-1})$ . We have

$$\begin{aligned} g &\rightarrow \theta(g) = \theta_p^{-1} \circ \sigma \circ \theta_p(g) \\ \Rightarrow \theta_p(\theta(g)) &= \sigma(\theta_p(g)) \\ \Rightarrow \theta_p(g) &\rightarrow \sigma(\theta_p(g)) \end{aligned}$$

and

$$\begin{aligned} \theta_p(\tau_\theta(g)) &= \theta_p(g(\theta(g^{-1}))) \\ &= \theta_p(g) \cdot \theta_p \circ \theta(g^{-1}) \\ &= \theta_p(g) \cdot \sigma(\theta_p(g^{-1})) \\ &= \theta_p(g) \cdot \sigma((\theta_p(g))^{-1}) \\ &= \tau_\sigma(\theta_p(g)) \end{aligned}$$

and

$$\begin{aligned} \Rightarrow g &\rightarrow \tau_\theta(g) \Leftrightarrow \theta_p(g) \rightarrow \tau_\sigma(\theta_p(g)) \\ \Rightarrow \theta_p(P_\theta) &= P_\sigma, \end{aligned}$$

where  $P_\theta = G/G^\theta$  and  $P_\sigma = \theta_p(G)/(\theta_p(G))^\sigma$ ,  $G^\theta$  and  $(\theta_p(G))^\sigma$  are corresponding fixed-point subgroup of  $G$  and  $\theta_p(G)$ .

Also

$$\begin{aligned} g &= \theta(g) \Leftrightarrow \theta_p(g) = \theta_p(\theta(g)) = \sigma(\theta_p(g)) \\ \Rightarrow g &\in G^\theta \Leftrightarrow \theta_p(g) \in (\theta_p(G))^\sigma \\ \Rightarrow (\theta_p(G))^\sigma &= \theta_p(G^\theta). \end{aligned}$$

Similarly

$$g^{-1} = \theta_p(g) \Leftrightarrow (\theta_p(g))^{-1} = \theta_p(g^{-1}) = \theta_p(\theta_p(g)) = \sigma(\theta_p(g)).$$

For  $K$ -orbits. Let  $x \in P_\theta, k \in G^\theta, x_p = \theta_p(x) \in P_\sigma$  and  $k_p = \theta_p(k) \in (\theta_p(G))^\sigma$ , then

$$\theta_p(k \cdot x \cdot k^{-1}) = \theta_p(k) \cdot \theta_p(x) \cdot \theta_p(k)^{-1} = k_p \cdot x_p \cdot k_p^{-1}.$$

This means that we can study the  $K$ -orbits of pair  $(\theta_p(G), \sigma)$  instead of  $(G, \theta)$  if  $\theta$  and  $\sigma$  are conjugate/isomorphic/equivalent and if it's easier to study. Once we get result for  $(\theta_p(G), \sigma)$ ,

we get corresponding results for  $(G, \theta)$  immediately.

Following are true too:

$g$  is semisimple in  $G \Leftrightarrow g_p = \theta_p(g)$  is semisimple in  $\theta_p(G)$ .

$g$  is unipotent in  $G \Leftrightarrow g_p = \theta_p(g)$  is unipotent in  $\theta_p(G)$ .

$T$  is a maximal torus in  $G \Leftrightarrow T_p = \theta_p(T)$  is a maximal torus in  $\theta_p(G)$ .

$A$  is a maximal  $\theta$ -split torus in  $G \Leftrightarrow A_p = \theta_p(A)$  is a maximal  $\sigma$ -split torus in  $\theta_p(G)$ .

We also have similar correspondence in Lie Algebras, Borel subgroups, and root system of  $G$  and  $\theta_p(G)$ .

*Remark 3.2.1.* We have same equivalence results for rational group pairs  $(G_{\mathbb{F}}, \theta)$  and  $(\theta_p(G_{\mathbb{F}}), \sigma)$ .

*Remark 3.2.2.* In conclusion, we only need to study involutions specified in Theorem 2.2.2, Theorem 2.2.3 and Theorem 2.2.4 to get the classification of  $K_{\mathbb{F}}$ -orbits of unipotent elements for all  $\mathbb{F}$ -involutions on  $\mathrm{SL}(n, \mathbb{F})$ .

## Chapter 4

# Classification Results for $\mathrm{SL}(2, \mathbb{F})$

In this chapter we classify  $K_{\mathbb{F}}$ -orbits of unipotent elements for  $\mathrm{SL}(2, \mathbb{F})$ . We have inner involutions only in this case, which makes things easier. We translate the classification problem to the "key" involution case first, then study classification problem for the key involution, and go back to the original problem to get results in those cases. the idea of reduction and the results we get here give big hints on how to work on the general case.

### 4.1 Reducing to the key involution case

Let's go to the problem we work on: classification of  $K$ -orbits of unipotent elements. Let  $G_{\mathbb{F}} = \mathrm{SL}(2, \mathbb{F})$ ,  $K_{\mathbb{F}} = K_{\sigma, \mathbb{F}} = G_{\mathbb{F}}^{\sigma}$ ,  $K_{\theta, \mathbb{F}} = G_{\mathbb{F}}^{\theta}$ , and define  $P_{\sigma}, P_{\theta}$  and  $P_{\tilde{\theta}}$  in a similar way.

Theorem 2.2.2 tells us that we only need to study the inner involution  $\theta_q$  with  $q = \begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix}$  (we abuse our notations here), where  $q \in \mathbb{F}^*/(\mathbb{F}^*)^2$ .

Let

$$p = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, p_1 = \begin{bmatrix} 1 & \\ & \sqrt{q} \end{bmatrix}, \text{ and } g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G_{\mathbb{F}} = \mathrm{SL}(2, \mathbb{F})$$

and

$$\begin{aligned}\sigma(g) &= \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \\ \theta(g) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \tilde{\theta}(g) &= \begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix} \\ \theta_{p_1}(g) &= p_1^{-1} g p_1\end{aligned}$$

From

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = p \cdot \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \cdot p^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

and

$$\begin{aligned}\begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix} &= \begin{bmatrix} 1 & \\ & \sqrt{q} \end{bmatrix} \begin{bmatrix} 0 & \sqrt{q} \\ \sqrt{q} & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & \sqrt{q} \end{bmatrix}^{-1} \\ &= \sqrt{q} \left( \begin{bmatrix} 1 & \\ & \sqrt{q} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & \sqrt{q} \end{bmatrix}^{-1} \right)\end{aligned}$$

we can see that

$$\begin{aligned}\theta &= \theta_p^{-1} \circ \sigma \circ \theta_p, \\ \tilde{\theta} &= \theta_{p_1}^{-1} \circ \theta \circ \theta_{p_1}.\end{aligned}$$

## 4.2 Classification of $K_{\mathbb{F}}$ -orbits for the involution with $q \in (\mathbb{F}^*)^2$

In this case we have  $q \in (\mathbb{F}^*)^2$ . We can take  $q = 1$  here.

Here is the main result for the case  $n = 2$ :

**Theorem 4.2.1.** *Let  $G_{\mathbb{F}} = \mathrm{SL}(2, \mathbb{F})$  and we use the above notations. If  $\mathrm{char}(\mathbb{F}) \neq 2$ , for pair  $(G_{\mathbb{F}}, \theta)$ , we can study  $(G_{\mathbb{F}}, \sigma)$  equivalently. For  $(G_{\mathbb{F}}, \sigma)$  we have  $2 \cdot |\mathbb{F}^*|/(\mathbb{F}^*)^2 + 1$   $K_{\mathbb{F}}$ -orbits of unipotent elements in the symmetric  $\mathbb{F}$ -variety  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ . We can especially choose*

$$\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

as representative elements in  $K_{\mathbb{F}}$ -orbits, where  $\alpha$  is a representative of  $\mathbb{F}^*/(\mathbb{F}^*)^2$ .

*Proof.* Since  $\theta = \theta_p^{-1} \circ \sigma \circ \theta_p$ , we study  $(\theta_p(G_{\mathbb{F}}), \sigma)$  instead of  $(G_{\mathbb{F}}, \theta)$ . Also, from

$$\begin{aligned}
\theta_p(g) &= \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} a+c & b+d \\ -a+c & -b+d \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} a+b+c+d & -a+b-c+d \\ -a-b+c+d & a-b-c+d \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} (a+d) + (b+c) & (-a+d) + (b-c) \\ (-a+d) - (b-c) & (a+d) - (b+c) \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2}[(a+d) + (b+c)] & \frac{1}{2}[(-a+d) + (b-c)] \\ \frac{1}{2}[(-a+d) - (b-c)] & \frac{1}{2}[(a+d) - (b+c)] \end{bmatrix} \\
&= \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix} \stackrel{\text{define}}{=} \bar{g}
\end{aligned}$$

we can see that  $\theta_p(G_{\mathbb{F}}) = G_{\mathbb{F}}$  and

$$K_{\mathbb{F}} = K_{\sigma, \mathbb{F}} = G_{\mathbb{F}}^{\sigma} = \left\{ k = \begin{bmatrix} \gamma & \\ & \gamma^{-1} \end{bmatrix} \in \text{SL}(2, \mathbb{F}) \mid \gamma \in \mathbb{F}^* \right\}.$$

Now let's focus on  $(G_{\mathbb{F}}, \sigma)$ . Let's look at elements in  $P_{\sigma}$ :  $(g^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix})$

$$\begin{aligned}
\tau_{\sigma}(g) &= g \cdot \sigma(g^{-1}) \\
&= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \\
&= \begin{bmatrix} a & -b \\ c & -d \end{bmatrix} \begin{bmatrix} d & b \\ -c & -a \end{bmatrix} \\
&= \begin{bmatrix} ad + bc & 2ab \\ 2cd & ad + bc \end{bmatrix} \\
&= (ad + bc)I + 2 \begin{bmatrix} 0 & ab \\ cd & 0 \end{bmatrix} = A + B
\end{aligned}$$

If  $abcd \neq 0$ , then  $B$  has two different eigenvalues in  $\overline{\mathbb{F}}$ , which implies that  $\tau_{\sigma}(g) = A + B$  is semisimple in  $P_{\sigma}$  and not unipotent.

If  $a = 0$ , then

$$\begin{aligned}
\det(\tau_{\sigma}(g)) &= 1 = ad - bc \rightarrow bc = -1 \\
\Rightarrow \tau_{\sigma}(g) &= -I + 2 \begin{bmatrix} 0 & 0 \\ cd & 0 \end{bmatrix},
\end{aligned}$$

which is not a unipotent either since the only eigenvalue is  $\lambda = -1$ . This also gives us the clue of proving corresponding part in the case  $n \geq 3$ .

We have same conclusion for the case  $d = 0$ .

If  $c = 0$  but  $b \neq 0$ ,  $\tau_{\sigma}(g) = \begin{bmatrix} 1 & 2ab \\ 0 & 1 \end{bmatrix}$  is definitely unipotent, and  $(k = \begin{bmatrix} \gamma & \\ & \gamma^{-1} \end{bmatrix})$

$$\begin{aligned}
k * \tau_{\sigma}(g) &= k \cdot \tau_{\sigma}(g) \cdot k^{-1} \\
&= \begin{bmatrix} \gamma & \\ & \gamma^{-1} \end{bmatrix} \begin{bmatrix} 1 & 2ab \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma & \\ & \gamma^{-1} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & 2ab\gamma^2 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

It's clear that we have  $|\mathbb{F}^*/(\mathbb{F}^*)^2|$   $K_{\sigma, \mathbb{F}}$ -orbits of unipotent elements and their representatives could be  $\begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}$  with  $\alpha \in \mathbb{F}^*/(\mathbb{F}^*)^2$ .

We have similar conclusion for the case  $b = 0$  but  $c \neq 0$ .



If  $b = c = 0$ , then  $\tau_\sigma(g) = I$  and  $k * \tau_\sigma(g) = I$ : we have only one  $K_{\sigma, \mathbb{F}}$ -orbit with only one element  $I$ .  $\square$

### 4.3 Classification of $K_{\mathbb{F}}$ -orbits for the involution with $q \neq 0 \notin (\mathbb{F}^*)^2$

Using the above result and calculations in the proof of the above theorem, we have

**Theorem 4.3.1.** *Let  $G_{\mathbb{F}} = \mathrm{SL}(2, \mathbb{F})$ . If  $|\mathbb{F}^*/(\mathbb{F}^*)^2| > 1$ ,  $\tilde{\theta}(g) = \begin{bmatrix} 0 & 1 \\ p & 0 \end{bmatrix}^{-1} g \begin{bmatrix} 0 & 1 \\ p & 0 \end{bmatrix}$ , where  $p \neq 0 \notin (\mathbb{F}^*)^2$ , then we have only one  $K_{\tilde{\theta}, \mathbb{F}}$ -orbit of unipotent element in the symmetric  $\mathbb{F}$ -variety  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ , the identity element  $I$ .*

*Proof.* Similar to Theorem 4.2.1, we will study the pair  $(\theta_{p_1}(G_{\mathbb{F}}), \theta)$  instead of  $(G_{\mathbb{F}}, \tilde{\theta})$  since

$$\tilde{\theta} = \theta_{p_1}^{-1} \circ \theta \circ \theta_{p_1}.$$

Let  $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  and we look at what  $\theta_{p_1}(G_{\mathbb{F}})$  looks like:

$$\begin{aligned} \theta_{p_1}(g) &= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{q} \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{q} \end{bmatrix} \\ &= \begin{bmatrix} a & \sqrt{q}b \\ \frac{c}{\sqrt{q}} & d \end{bmatrix} = \begin{bmatrix} a & \sqrt{q}b \\ \sqrt{q}\frac{c}{q} & d \end{bmatrix}. \end{aligned}$$

Using calculations in Theorem 4.2.1, we can see that

To get unipotent elements in the symmetric  $\mathbb{F}$ -variety, for pair  $(G_{\mathbb{F}}, \theta)$  we have either  $\bar{b} = \frac{1}{2}[(-a + d) + (b - c)] = 0$  or  $\bar{c} = \frac{1}{2}[(-a + d) - (b - c)] = 0$ . Also, to get  $\theta_{p_1}(G_{\mathbb{F}})$  from  $G_{\mathbb{F}}$  we only need to replace  $b$  and  $c$  by  $\sqrt{q}b$  and  $\sqrt{q}\frac{c}{q}$ . Therefore to get unipotent elements in  $P_{\theta_{p_1}}$ , we will have

$$\begin{aligned} 0 = \bar{b} &= \frac{1}{2}[(-a + d) + (\sqrt{q}b - \sqrt{q}\frac{c}{q})] = \frac{1}{2}[(-a + d) + \sqrt{q}(b - \frac{c}{q})] \\ \Rightarrow \left\{ \begin{bmatrix} a - d = 0 \\ b - \frac{c}{q} = 0 \end{bmatrix} \right. & \quad (\sqrt{q} \notin \mathbb{F}) \\ \Rightarrow \bar{c} = \frac{1}{2}[(-a + d) - (\sqrt{q}b - \sqrt{q}\frac{c}{q})] &= \frac{1}{2}[(-a + d) - \sqrt{q}(b - \frac{c}{q})] = 0. \end{aligned}$$

From  $\bar{c} = 0$  we have same conclusion. Therefore we can have only one unipotent in  $P_\theta$  or  $P_{\tilde{\theta}}$ , the identity element  $I$ .  $\square$

*Remark 4.3.2.* For case  $n = 2$ , we first reduce to the classification for our key involution and therefore finish classification for all involutions. We have two parts in our proof of classification for  $q \in (\mathbb{F})^2$ : 1, we first find out only certain types of unipotent elements can generate unipotent elements in symmetric  $\mathbb{F}$ -variety  $P$ , then 2, classify those particular types of unipotent elements only. For the case  $q \neq 0 \notin (\mathbb{F})^2$ , we translate it into the one for our key involution, and then use relation between two problems and special properties of translated problem to get the result. The results in this section are simple but important, since what we will do for the case  $n \geq 3$  is exactly what we have done for the case  $n = 2$ .

## Chapter 5

# Classification of $G_{\mathbb{F}}$ -orbits of Unipotent Elements for $\mathrm{SL}(n, \mathbb{F})$

We here study  $G_{\mathbb{F}}$ -orbits of unipotent elements in  $G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$  and then apply this result to our classification problem for inner involutions. We first introduce one definition we need and then prove the result in the following section.

### 5.1 A definition and its properties

We first introduce some notations, one definition and its properties that we need in this chapter and in Chapter 6.

We define notation  $\delta$  for some integer  $i$  in following way:

$$\delta(i) = \begin{cases} 0, & \text{if } i = 0, \\ 1, & \text{if } i \neq 0. \end{cases}$$

Given a sequence of integers  $n_0, n_1, \dots, n_r$ , we are able to define following equation:

$$\prod_{0 \leq i \leq r} x_i^{n_i} = p \tag{5.1.1}$$

where  $p \in \mathbb{F}^*$  is given and  $x_i, i = 0, \dots, r$  are unknowns.

*Remark 5.1.1.* It doesn't matter whether  $n_{i_0} = 0$  for some  $i_0$ , and we actually only need to consider all nonzero integers in the sequence if it is the case.

Using above equation we are able to define an equivalent class in the field  $\mathbb{F}$ :

**Definition 5.1.2.** Given  $f_1, f_2 \in \mathbb{F}^*$  and a sequence of integers  $n_i, i = 0, 1, \dots, r$ , we say that  $f_1$  and  $f_2$  are  $n_i$ -equivalent, denoted by  $f_1 \stackrel{n_i}{\sim} f_2$ , if and only if the equation (5.1.1) has at least one set of solution in  $\mathbb{F}^*$  with  $p = f_1/f_2$ .

It's easy to check that the above definition indeed defines an equivalence relation in  $\mathbb{F}^*$ .

We have following lemma for the equivalent classes in  $\mathbb{F}^*$ :

**Lemma 5.1.3.** *Suppose the greatest common factor of nonzero integers  $n_i, i = 0, 1, \dots, r$ , is  $h$ . We have then*

$$f_1 \stackrel{n_i}{\sim} f_2 \Leftrightarrow f_1/f_2 \in (\mathbb{F}^*)^h.$$

*Proof.* It's obvious that we only need to prove the case that all  $n_i$  are all positive integers.

Necessity. Suppose  $f_1 \stackrel{n_i}{\sim} f_2$  and  $n_i = h l_i$  where  $h$  is the great common factor. We then have for  $x_i$ 's in equation (5.1.1) for  $p = f_1/f_2$

$$\begin{aligned} f_1/f_2 &= \prod_{0 \leq i \leq r} x_i^{n_i} \\ &= \prod_{0 \leq i \leq r} x_i^{h l_i} \\ &= \left( \prod_{0 \leq i \leq r} x_i^{l_i} \right)^h \in (\mathbb{F}^*)^h. \end{aligned}$$

Sufficiency. Suppose  $f_1/f_2 = p^h \in (\mathbb{F}^*)^h$ . Since  $h$  is the greatest common factor of  $n_i$ , we are able to find integers  $u_i$  such that

$$h = \sum_{0 \leq i \leq r} u_i n_i.$$

Let  $x_i = p^{u_i}$ , then

$$\begin{aligned} \prod_{0 \leq i \leq r} x_i^{n_i} &= \prod_{0 \leq i \leq r} (p^{u_i})^{n_i} \\ &= p^{\sum_{0 \leq i \leq r} u_i n_i} \\ &= p^h = f_1/f_2. \end{aligned}$$

So we have  $f_1 \stackrel{n_i}{\sim} f_2$ .

□

## 5.2 Classification of $G_{\mathbb{F}}$ -orbits of unipotent elements for $\mathrm{SL}(n, \mathbb{F})$

Let  $G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$ , and we define  $G_{\mathbb{F}}$ -orbit of  $A \in G_{\mathbb{F}}$  as the set

$$\{gAg^{-1} | g \in G_{\mathbb{F}}\} = \{g^{-1}Ag | g \in G_{\mathbb{F}}\}.$$

Let  $n \in \mathbb{N}$  be partitioned in the following way:

$$n = \beta_0 n_0 + \beta_1 n_1 + \cdots + \beta_r n_r, \quad (5.2.1)$$

where  $1 = n_0 < n_1 < n_2 < \cdots < n_{r-1} < n_r \leq n$ , and  $\beta_i, i = 0, \dots, r$ , are nonnegative integers.

In this section, we discuss the classification of  $G_{\mathbb{F}}$ -orbits of unipotent elements in  $G_{\mathbb{F}}$ . The above partition of  $n$  corresponds to sizes of Jordan blocks in the Jordan canonical form of a unipotent element in  $G$ , i.e., we have  $\beta_i$  Jordan blocks of size  $n_i, i = 0, 1, \dots, r$ .

We also note that, whether  $\mathbb{F} = \overline{\mathbb{F}}$  or not, given a nilpotent matrix  $A$ , we can always find matrix  $X \in \mathrm{GL}(n, \mathbb{F})$  such that  $A = X^{-1}JX$ , where  $J$  is the usual Jordan canonical form.

Here is the main result in this chapter though not used in our thesis directly:

**Theorem 5.2.1.** *Let  $\mathbb{F}$  be an arbitrary field. Suppose we have two unipotent elements  $A = XJ_AX^{-1}$  and  $B = YJ_BY^{-1}$  in  $G$  with same Jordan canonical form determined by partition (5.2.1):  $J_A = J_B$ . Then  $A$  and  $B$  is in the same  $G_{\mathbb{F}}$ -orbit, i.e., there exists  $P \in G_{\mathbb{F}}$  such that  $B = PAP^{-1}$ , if and only if  $\det(X)$  and  $\det(Y)$  are  $n_i\delta(\beta_i)$ -equivalent, or equivalently  $\det(X)/\det(Y) \in (\mathbb{F}^*)^h$ , where  $h$  is the greatest common factor of the sequence  $n_i\delta(\beta_i), i = 0, 1, \dots, r$ .*

*Furthermore, all  $G$ -orbits can choose elements in the following form as representatives*

$$XJX^{-1}, \quad (5.2.2)$$

*where  $J$  is the Jordan canonical form and*

$$X = \begin{bmatrix} p & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$ .

*Proof.* If  $\mathbb{F} = \overline{\mathbb{F}}$ , we can see that two elements in  $G_{\mathbb{F}}$  are  $G_{\mathbb{F}}$ -similar if and only if their Jordan canonical forms are same. When  $\mathbb{F} \neq \overline{\mathbb{F}}$ , we need more conditions.

If  $A \in G_{\mathbb{F}} = \text{SL}(n, \mathbb{F})$  is unipotent, then its only eigenvalue is  $\lambda = 1$ . Let  $V = \mathbb{F}^n$  and define an operator  $p$  on any  $v \in V$

$$pv \triangleq (A - \lambda I)v.$$

We can then define subspaces (it's clear that  $p^{n_r}V = \{0\}$ )

$$V^0 = V$$

$$V^j = \{v \in V | p^{j-1}v \neq 0 \text{ but } p^jv = 0\}, j = 1, \dots, n_r.$$

It's clear that  $p^{j-i}V^j \subseteq V^i$  for all  $0 \leq i \leq j \leq n_r$ .

Now we define  $\tilde{V}^j, j = n_r, \dots, 1$  as in following expression

$$\begin{aligned} V^{n_r} &= \tilde{V}^{n_r}; \\ V^{n_r-1} &= p(V^{n_r}) \oplus \tilde{V}^{n_r-1}; \\ &\vdots \\ V^j &= p(V^{j+1}) \oplus \tilde{V}^j; \\ &\vdots \\ V^1 &= p(V^2) \oplus \tilde{V}^1. \end{aligned}$$

We assume  $m_i = \dim(V^i)$  and  $\tilde{m}_i = \dim(\tilde{V}^i)$ . It's clear that

$$\begin{aligned}
\tilde{m}_{n_r} &= m_{n_r}; \\
\tilde{m}_i &= m_{i+1} - m_i, i = n_r - 1, \dots, 1; \\
m_i &\geq m_j \text{ if } i \leq j.
\end{aligned}$$

We can always reduce  $A$  to its Jordan canonical form in field  $\mathbb{F}$ :  $A = XJ_AX^{-1}$  because  $A$  is unipotent. We here want to reduce  $A$  into a special canonical form, still denoted by  $J_A$ , to get the bi-condition. Here is what I mean. It's not hard to see that, to get the Jordan canonical form, we can choose a basis from  $\tilde{V}^1, pV^2, \tilde{V}^2, pV^3, \dots, \tilde{V}^i, pV^{i+1}, \dots, pV^{n_r}, \tilde{V}^{n_r}$ , and we still denote the basis by  $\tilde{V}^1, pV^2, \tilde{V}^2, pV^3, \dots, \tilde{V}^i, pV^{i+1}, \dots, pV^{n_r}, \tilde{V}^{n_r}$ , and therefore we have

$$A[\tilde{V}^1, pV^2, \dots, \tilde{V}^{n_r}] = [\tilde{V}^1, pV^2, \dots, \tilde{V}^{n_r}]J_A$$

or

$$AX = XJ_A$$

where  $J_A$  is determined by arranging (generalized) eigenvectors in the above way, and then also determined by  $m_i$  (and therefore  $\tilde{m}_i$ ),  $i = 1, \dots, n_r$  only and is unique. Now let's see what kind of  $X$  we can take. We start from  $\tilde{V}^{n_r}$  down to  $\tilde{V}^1$ . We assume  $\tilde{m}_i = \dim(\tilde{V}^i) \neq 0, i = 1, \dots, n_r$  first. This means that  $n_0 = 1, n_1 = 2, n_2 = 3, \dots, n_r = r + 1$ .

First we fix a basis  $X_0 = [\tilde{V}_0^1, pV_0^2, \tilde{V}_0^2, pV_0^{n_r}, \tilde{V}_0^{n_r}]$  which gives us the canonical form  $J_A$ . It's clear that all possible  $\tilde{V}^{n_r}$  has the form  $\tilde{V}_0^{n_r}$ , where  $D_{n_r} \in \text{GL}(\tilde{m}_{n_r}, \mathbb{F})$  could any nonsingular matrix.  $pV^{n_r} = p\tilde{V}^{n_r} = p\tilde{V}_0^{n_r} D_{n_r}$ . Similarly all possible  $\tilde{V}^{n_r-1}$  has the form

$$\begin{aligned}
\tilde{V}_0^{n_r-1} D_{n_r-1} + pV_0^{n_r} \cdot * &= \tilde{V}_0^{n_r-1} D_{n_r-1} + p\tilde{V}_0^{n_r} \cdot * \\
&= \begin{bmatrix} \tilde{V}_0^{n_r-1} & p\tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_{n_r-1} \\ * \end{bmatrix}.
\end{aligned}$$

where  $D_{n_r-1} \in \text{GL}(\tilde{m}_{n_r-1}, \mathbb{F})$  and  $*$  is any  $m_{n_r}$  by  $\tilde{m}_{n_r-1}$  matrix. So all possible  $V^{n_r-1} = [\tilde{V}^{n_r-1}, pV^{n_r}]$ , has the form

$$V^{n_r-1} = \begin{bmatrix} \tilde{V}_0^{n_r-1} & p \tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_{n_r-1} & 0 \\ * & D_{n_r} \end{bmatrix} = V_0^{n_r-1} \begin{bmatrix} D_{n_r-1} & 0 \\ * & D_{n_r} \end{bmatrix}.$$

Again, all possible  $\tilde{V}^{n_r-2}$  has the form

$$\begin{aligned} \tilde{V}_0^{n_r-2} D_{n_r-2} + p V^{n_r-1} \cdot * &= \tilde{V}_0^{n_r-2} D_{n_r-2} + p \begin{bmatrix} \tilde{V}_0^{n_r-1} & p \tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_{n_r-1} & 0 \\ * & D_{n_r} \end{bmatrix} \cdot * \\ &= \begin{bmatrix} \tilde{V}_0^{n_r-2} & p \tilde{V}_0^{n_r-1} & p^2 \tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_{n_r-2} \\ * \\ * \end{bmatrix}. \end{aligned}$$

This gives us

$$\begin{aligned} V^{n_r-2} &= \begin{bmatrix} \tilde{V}^{n_r-2} & p V_0^{n_r-1} \end{bmatrix} \\ &= \begin{bmatrix} \tilde{V}_0^{n_r-2} & p \tilde{V}_0^{n_r-1} & p^2 \tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_{n_r-2} \\ * & D_{n_r-1} \\ * & * & D_{n_r} \end{bmatrix} \\ &= V_0^{n_r-2} \begin{bmatrix} D_{n_r-2} \\ * & D_{n_r-1} \\ * & * & D_{n_r} \end{bmatrix}. \end{aligned}$$

Repeating it we can get a general formula for  $V^i$ :

$$V^i = \begin{bmatrix} \tilde{V}_0^i & p \tilde{V}_0^{i+1} & \dots & p^{n_r-i} \tilde{V}_0^{n_r} \end{bmatrix} \begin{bmatrix} D_i \\ * & D_{i+1} \\ * & * & \ddots \\ \vdots & \vdots & \ddots & D_{n_r-1} \\ * & * & \dots & * & D_{n_r} \end{bmatrix}$$



$$\begin{aligned}
&= V_0^i \begin{bmatrix} D_i & & & & \\ * & D_{i+1} & & & \\ * & * & \ddots & & \\ \vdots & \vdots & \ddots & D_{n_r-1} & \\ * & * & \cdots & * & D_{n_r} \end{bmatrix} \\
&= V_0^i \tilde{D}_i.
\end{aligned}$$

where  $D_j \in \text{GL}(\tilde{m}_j, \mathbb{F}), i \leq j \leq n_r$ .

We then can get the formula for the general  $X$ :

$$\begin{aligned}
X &= \begin{bmatrix} V^1 & V^2 & \cdots & V^{n_r} \end{bmatrix} \\
&= \begin{bmatrix} \tilde{V}^1 & pV^2 & \tilde{V}^2 & pV^3 & \cdots & pV^{n_r} & \tilde{V}^{n_r} \end{bmatrix}
\end{aligned}$$

where  $V^i = \begin{bmatrix} \tilde{V}^i & pV^{i+1} \end{bmatrix}, 1 \leq i \leq n_r - 1$  and  $V^{n_r} = \tilde{V}^{n_r}$ .

Now it's

$$\begin{aligned}
X &= X_0 \begin{bmatrix} \tilde{D}_1 & & & \\ & \tilde{D}_2 & & \\ & & \ddots & \\ & & & \tilde{D}_{n_r} \end{bmatrix} \\
&= X_0 D.
\end{aligned}$$

Now we consider the general case, i.e., there could be some  $i_0$  such that  $\dim(\tilde{V}^{i_0}) = 0$ . In this case we have

$$\begin{aligned}
V^{i_0} &= pV^{i_0+1} \\
&= pV_0^{i_0+1} \tilde{D}_{i_0+1} \\
&= V_0^{i_0} \tilde{D}_{i_0+1},
\end{aligned}$$

This implies  $\tilde{D}_{i_0} = \tilde{D}_{i_0+1}$ . Therefore, generally  $\tilde{D}_i$  (and therefore  $D_i$ ) appear in  $D$  if and only if  $\tilde{m}_i = \dim(\tilde{V}^i) > 0$ . Using correspondence between sizes of Jordan blocks and partition of  $n$  in (5.2.1), we can see that if  $\tilde{m}_i \neq 0$ , we then actually have  $\beta_i = \tilde{m}_{n_i}, 0 \leq i \leq r$ .

Let  $\omega_i = \det(D_{n_i}) \in \mathbb{F}^*, 0 \leq i \leq r$ . We then have

$$\begin{aligned}\det(D) &= \prod_{1 \leq i \leq n_r} \det(\tilde{D}_i) \\ &= \prod_{0 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)}.\end{aligned}$$

We only need to consider  $n_i$  instead of  $i$  here since, if for some index  $i_0$ ,  $\beta_{i_0} = \tilde{m}_{n_{i_0}} = 0$ , then  $D_{n_{i_0}}$  would not show up and therefore  $\omega_{i_0}^{n_{i_0} \delta(\beta_{i_0})} = \omega_{i_0}^0 = 1$  in the above expression.

If  $\delta(\beta_i) = 1$ , then  $D_{n_i}$  will appear  $n_i$  times in  $D$ , and therefore  $\omega_i$  appears  $n_i$  times in the above.

This tells us that all possible  $X$  in  $A = XJ_AX^{-1}$  has determinant:

$$\det(X) = \det(X_0) \cdot \prod_{0 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)},$$

i.e., an  $n_i \delta(\beta_i)$ -equivalent class in  $\mathbb{F}^*$ .

Now we assume

$$A = XJ_AX^{-1}, B = YJ_BY^{-1} \tag{5.2.3}$$

are  $G_{\mathbb{F}}$ -similar, i.e., there exists  $P \in G_{\mathbb{F}}$  such that  $B = PAP^{-1}$ . Here  $J_A$  and  $J_B$  are in the special canonical form described above. We then have  $J_A = J_B$  immediately. From

$$B = YJ_BY^{-1} = YJ_AY^{-1} = PAP^{-1} = PXJ_AX^{-1}P^{-1}$$

We know that there exists a nonsingular diagonal block matrix  $D$  as described above such that

$$PX = YD \Leftrightarrow P = X^{-1}YD.$$

$$\begin{aligned}
P \in G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F}) &\Leftrightarrow 1 = \det(X^{-1}YD) = \det(X)^{-1} \det(Y) \det(D) \\
&\Leftrightarrow \prod_{1 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)} = \frac{\det(X)}{\det(Y)}.
\end{aligned}$$

All possible  $\det(X)$  and  $\det(Y)$  are  $\det(X_0) \det(D_X)$  and  $\det(Y_0) \det(D_Y)$ , where  $X_0$  and  $Y_0$  are fixed. We now have

$$\prod_{1 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)} = \frac{\det(X)}{\det(Y)} = \frac{\det(X_0) \det(D_X)}{\det(Y_0) \det(D_Y)}$$

so

$$\begin{aligned}
\frac{\det(X_0)}{\det(Y_0)} &= \frac{\det(D_Y)}{\det(D_X)} \prod_{1 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)} \\
&= \frac{\prod_{1 \leq i \leq r} (\omega_{i,Y})^{n_i \delta(\beta_i)}}{\prod_{1 \leq i \leq r} (\omega_{i,X})^{n_i \delta(\beta_i)}} \cdot \prod_{1 \leq i \leq r} (\omega_i)^{n_i \delta(\beta_i)} \\
&= \prod_{1 \leq i \leq r} \left( \frac{\omega_{i,Y} \cdot \omega_i}{\omega_{i,X}} \right)^{n_i \delta(\beta_i)}
\end{aligned}$$

So we can see that  $\det(X_0)$  and  $\det(Y_0)$  must be in the same  $n_i \delta(\beta_i)$ -equivalent class, or no matter what  $X$  and  $Y$  we choose in (5.2.3),  $\det(X)$  and  $\det(Y)$  must be in the same  $n_i \delta(\beta_i)$ -equivalent class, and therefore using Lemma 5.1.3 we can see that  $\det(X)/\det(Y) \in (\mathbb{F}^*)^h$ , where  $h$  is the greatest common factor of  $i \delta(\beta_i)$ .

On the other hand, in (5.2.3), assume  $\det(X)$  and  $\det(Y)$  are in the same  $n_i \delta(\beta_i)$ -equivalent class. If  $\det(X) = \det(Y)$  we can take  $P = YX^{-1} \in G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$  directly. If  $\det(X) \neq \det(Y)$ , according to the above analysis, we can take another matrix  $\tilde{Y}$  such that for some block diagonal matrix  $D$

$$B = YJ_B Y^{-1} = \tilde{Y} J_B \tilde{Y}^{-1} \quad \text{with} \quad \tilde{Y} = YD$$

and  $\det(X) = \det(\tilde{Y})$  because  $\det(X)$ ,  $\det(Y)$  and  $\det(\tilde{Y})$  are in the same  $n_i \delta(\beta_i)$ -equivalent class. Now we can take  $P = \tilde{Y} X^{-1} \in G_{\mathbb{F}}$ .

We can also see that, even if  $J_A$  and  $J_B$  are in the traditional canonical form, we still have

same conclusion/bi-condition here.

For each  $G_{\mathbb{F}}$ -orbit of unipotent elements in  $G_{\mathbb{F}}$ , we can always take elements in the form (5.2.2) as our representatives in the  $G_{\mathbb{F}}$ -orbits of unipotent elements. □

*Remark 5.2.2.* If in the partition (5.2.1)  $\beta_0 > 0$ , i.e., we always have zero Jordan block, we can see that the greatest common factor  $h$  is always 1 and if this is the case, the usual Jordan canonical form is good enough to distinguish different  $G_{\mathbb{F}}$ -orbits.

We actually need a little more for Theorem 6.1.8. First we introduce

*Notation 5.2.3.* By  $p \in \mathbb{F}_{\pm}^*/(\mathbb{F}_{\pm}^*)^h$ , we mean that given  $p_1, p_2 \in \mathbb{F}^*/(\mathbb{F}^*)^h$ ,  $p_1$  and  $p_2$  represent same equivalence class in  $\mathbb{F}_{\pm}^*/(\mathbb{F}_{\pm}^*)^h$  if and only if  $p_2 = \pm p_1$ . It is true that this indeed defines an equivalence relation in  $\mathbb{F}^*$  (or  $\mathbb{F}^*/(\mathbb{F}^*)^h$ ).

Now we have

**Theorem 5.2.4.** *Given  $A, B$  and  $P$  as in Theorem 5.2.1, if we allow the determinant of  $P$  to be  $\pm 1$  instead of 1, then*

1. *if  $-1 \in (\mathbb{F}^*)^h$ , we will require same condition and have same representatives of " $\pm G_{\mathbb{F}}$ -orbits" as in Theorem 5.2.1;*
2. *if  $-1 \notin (\mathbb{F}^*)^h$ , we then require instead  $\pm \frac{\det(X)}{\det(Y)} \in (\mathbb{F}^*)^h$  as condition of being in the same " $\pm G_{\mathbb{F}}$ -orbit"; for representatives, we choose instead  $p \in \mathbb{F}_{\pm}^*/(\mathbb{F}_{\pm}^*)^h$ .*

*Proof.* We only use the determinant condition in the last part of the proof of Theorem 5.2.1, and the rest is from the definition of the equivalence class  $\mathbb{F}_{\pm}^*/(\mathbb{F}_{\pm}^*)^h$ . □

## Chapter 6

# Classification Results for $\mathrm{SL}(n, \mathbb{F})$ for Inner Involutions

Now we go to the case  $n \geq 3$ . Let  $G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F})$ . We have two types of involutions to study, inner involutions and outer involutions. We study inner involutions case in this chapter. From Theorem 2.2.3 we know that we only need to study two types of inner involutions. We will study a key inner involution first, and then apply conclusions obtained for the key inner involution to the other cases.

### 6.1 Key inner involution

The key involution  $\sigma$  we will study is defined as

$$\sigma(g) = p^{-1}gp, \quad p = \begin{bmatrix} I & \\ & -I \end{bmatrix} = p^{-1}.$$

Here  $n \in 2\mathbb{Z}^+$ .

In this case,

$$K_{\mathbb{F}} = K_{\sigma, \mathbb{F}} = G_{\mathbb{F}}^{\sigma} = \left\{ k = \begin{bmatrix} P & \\ & Q \end{bmatrix} \in \mathrm{SL}(n, \mathbb{F}) \mid P, Q \in \mathrm{GL}\left(\frac{n}{2}, \mathbb{F}\right), \det(P)\det(Q) = 1 \right\}.$$

Again we will do same thing as we did in the case  $n = 2$ :

1. Show that  $g$  (or  $\tau(g)$ ) needs to have a special form to yield a unipotent element in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ ;
2. Classify all  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P$ .

### 6.1.1 To make $\tau(g)$ to be unipotent, $g$ needs to have a special form

Our main results in this subsection are Theorem 6.1.1 and Theorem 6.1.3.

**Theorem 6.1.1.** *Let  $g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in G_{\mathbb{F}}$ . If  $\det(A)\det(D) \neq 0$ , then all unipotent elements in  $P$  can be generated by a special type of  $g$ :  $g = \begin{bmatrix} I & B \\ C & I \end{bmatrix} = I + J$  with  $J = \begin{bmatrix} & B \\ C & \end{bmatrix}$  nilpotent.*

To prove it, we need following lemma first.

**Lemma 6.1.2.** *Let  $J = \begin{bmatrix} & B \\ C & \end{bmatrix}$ , then  $\lambda$  is an eigenvalue of  $J$  if and only if  $-\lambda$  is an eigenvalue of  $J$ . So we have  $\det(I + J) \neq 0 \Leftrightarrow \det(I - J) \neq 0$ .*

*Proof.* Let  $\lambda \neq 0$ , then from

$$\begin{bmatrix} I & 0 \\ -\frac{1}{\lambda}C & I \end{bmatrix} \begin{bmatrix} \lambda I & B \\ C & \lambda I \end{bmatrix} = \begin{bmatrix} \lambda I & B \\ 0 & \lambda I - \frac{1}{\lambda}CB \end{bmatrix}$$

we have

$$\begin{aligned} \det(\lambda I - J) &= \det(\lambda I) \det(\lambda I - \frac{1}{\lambda}CB) \\ &= \det(\lambda I \cdot (\lambda I - \frac{1}{\lambda}CB)) \\ &= \det(\lambda^2 I - CB) \end{aligned}$$

So the characteristic polynomial of  $J$  is a polynomial of  $\lambda^2$ , which gives us the conclusion.  $\square$

*Proof of Theorem 6.1.1.* First we prove that  $g = \begin{bmatrix} I & B \\ C & I \end{bmatrix} = I + J$  with  $J$  nilpotent generates unipotent elements in  $P$ . We then prove that all unipotent elements in  $P$  can be generated by this type of  $g$ . Also, the order of  $I$  should be clear from its context and it should not be confusing if it stands for identities with different sizes.

It's clear that  $J^2 = \begin{bmatrix} BC & \\ & CB \end{bmatrix}$  is diagonal, and

$$\begin{aligned} g &= I + J \\ &= (I + J)(I - J)(I - J)^{-1} \text{ (Lemma 6.1.2)} \\ &= (I - J^2)(I - J)^{-1}, \end{aligned}$$

so  $g^{-1} = (I - J)(I - J^2)^{-1}$  and ( $p$  and  $I - J^2$  commute)

$$\begin{aligned}
\tau(g) &= g\sigma(g^{-1}) \\
&= gp^{-1}g^{-1}p \\
&= (I + J)p(I - J)(I - J^2)^{-1}p \\
&= (I + J)p(I - J)p(I - J^2)^{-1} \\
&= (I + J)[p(I - J)p](I - J^2)^{-1} \\
&= (I + J)(I - pJp)(I - J^2)^{-1} \\
&= (I + J)(I + J)(I - J^2)^{-1} \quad (pJp = -J) \\
&= (I + 2J + J^2)(I - J^2)^{-1} \\
&= [(I - J^2) + (2J + 2J^2)](I - J^2)^{-1} \\
&= I + 2J(I + J)(I - J^2)^{-1} \\
&= I + 2J(I - J)^{-1}
\end{aligned}$$

then  $\tau(g)$  is unipotent if and only if  $2J(I - J)^{-1}$  or  $J(I - J)^{-1}$  is nilpotent. The eigenvalues of  $J(I - J)^{-1}$  have the form  $\lambda = \frac{\lambda_J}{1 - \lambda_J}$ , where  $\lambda_J$  is an eigenvalue of  $J$ . This is because  $J(I - J)^{-1}$  is a rational function of  $J$  and the eigenvalues  $\lambda$  of  $J(I - J)^{-1}$  are just same rational function of eigenvalues  $\lambda_J$  of  $J$ . So  $\lambda = 0 \Leftrightarrow \lambda_J = 0$ , i.e.,  $J$  is nilpotent.

Let's see a general case. Let

$$\begin{aligned}
g &= \begin{bmatrix} A & B \\ C & D \end{bmatrix} \\
Q &= \begin{bmatrix} A^{-1} & \\ & D^{-1} \end{bmatrix} \\
\bar{g} = gQ &= \begin{bmatrix} I & BD^{-1} \\ CA^{-1} & I \end{bmatrix} = \begin{bmatrix} I & \bar{B} \\ \bar{C} & I \end{bmatrix}
\end{aligned}$$

then

$$\begin{aligned}
\tau(g) &= gp^{-1}g^{-1}p \\
&= gQ \cdot Q^{-1}p^{-1}g^{-1}p \\
&= \bar{g} \cdot p^{-1}Q^{-1}g^{-1}p \quad (Qp = pQ) \\
&= \bar{g} \cdot p^{-1}(gQ)^{-1}p \\
&= \bar{g} \cdot p^{-1}\bar{g}^{-1}p \\
&= \tau(\bar{g})
\end{aligned}$$

This means that, if  $\det(A)\det(D) \neq 0$ , we can always assume  $A = D = I$ . To show that  $\bar{g} \in \mathrm{SL}(n, \mathbb{F})$ , i.e.,  $\det(\bar{g}) = 1$  we only need to notice that, if we assume  $\bar{g} = I + \bar{J}$

$$\tau(g) = \tau(\bar{g}) \text{ unipotent} \Rightarrow \bar{J} \text{ nilpotent} \Rightarrow \det(\bar{g}) = 1.$$

□

**Theorem 6.1.3.** *Let  $g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in G_{\mathbb{F}}$ . If  $\det(A)\det(D) = 0$ , then  $\tau(g)$  has one eigenvalue  $\lambda = -1$  which implies that  $\tau(g)$  is not unipotent.*

We assume the field  $\mathbb{F} = \overline{\mathbb{F}}$  here. If we prove the theorem for the case  $\mathbb{F} = \overline{\mathbb{F}}$ , the theorem will also be true for the case  $\mathbb{F} \neq \overline{\mathbb{F}}$ :  $\tau(g)$  has eigenvalue  $\lambda = -1$  in  $\overline{\mathbb{F}}$  and it'd have eigenvalue  $\lambda = -1$  in  $\mathbb{F}$  too.

We put the proof of Theorem 6.1.3 into two cases:

Case 1. One of  $A$  and  $D$  is nonsingular;

Case 2. Both  $A$  and  $D$  are singular.

Proof of Case 1. Without loss of generality, we assume  $A$  is nonsingular and  $D$  is singular. Same as in Theorem 6.1.1, we can assume  $A = I$  and  $D = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ , where the identity matrices in  $A$  and  $D$  should be taken properly. Now we have

$$g = \begin{bmatrix} I & B \\ C & D \end{bmatrix} \text{ with } D = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Thanks to the nonsingularity of  $A = I$ , we can write out  $g^{-1}$  explicitly. Here is a quick look: ( $g$  nonsingular  $\Rightarrow D - CB$  nonsingular)

$$\begin{bmatrix} I & -B \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & (D - CB)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -C & I \end{bmatrix} \begin{bmatrix} I & B \\ C & D \end{bmatrix} = I$$

which implies



$$\begin{aligned}
g^{-1} &= \begin{bmatrix} I & -B \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & (D - CB)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -C & I \end{bmatrix} \\
&= \begin{bmatrix} I & -B(D - CB)^{-1} \\ 0 & (D - CB)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -C & I \end{bmatrix} \\
&= \left[ \begin{array}{c|c} I + B(D - CB)^{-1}C & -B(D - CB)^{-1} \\ \hline -(D - CB)^{-1}C & (D - CB)^{-1} \end{array} \right].
\end{aligned}$$

From  $g \cdot g^{-1} = I$  we have  $(g \cdot g^{-1})_{21} = 0$  if we partition  $g \cdot g^{-1}$  into two by two block matrix, and therefore

$$\begin{aligned}
0 &= (g \cdot g^{-1})_{21} \\
&= \begin{bmatrix} C & D \end{bmatrix} \cdot \begin{bmatrix} I + B(D - CB)^{-1}C \\ -(D - CB)^{-1}C \end{bmatrix} \\
&= C \cdot [I + B(D - CB)^{-1}C] - D(D - CB)^{-1}C
\end{aligned}$$

and we then get

$$C(I + B(D - CB)^{-1}C) = D(D - CB)^{-1}C. \quad (6.1.1)$$

Then

$$\begin{aligned}
\tau(g) &= gp^{-1}g^{-1}p \\
&= \begin{bmatrix} I & B \\ C & D \end{bmatrix} \begin{bmatrix} I & \\ & -I \end{bmatrix} \left[ \begin{array}{c|c} I + B(D - CB)^{-1}C & -B(D - CB)^{-1} \\ \hline -(D - CB)^{-1}C & (D - CB)^{-1} \end{array} \right] \begin{bmatrix} I & \\ & -I \end{bmatrix} \\
&= \begin{bmatrix} I & -B \\ C & -D \end{bmatrix} \left[ \begin{array}{c|c} I + B(D - CB)^{-1}C & B(D - CB)^{-1} \\ \hline -(D - CB)^{-1}C & -(D - CB)^{-1} \end{array} \right] \\
&= \begin{bmatrix} I + B(D - CB)^{-1}C + B(D - CB)^{-1}C & B(D - CB)^{-1} + B(D - CB)^{-1} \\ C(I + B(D - CB)^{-1}C) + D(D - CB)^{-1}C & CB(D - CB)^{-1} + D(D - CB)^{-1} \end{bmatrix} \\
&\stackrel{6.1.1}{=} \begin{bmatrix} I + 2B(D - CB)^{-1}C & 2B(D - CB)^{-1} \\ 2D(D - CB)^{-1}C & (D + CB)(D - CB)^{-1} \end{bmatrix} \\
&= \begin{bmatrix} I + 2B(D - CB)^{-1}C & 2B(D - CB)^{-1} \\ 2D(D - CB)^{-1}C & -I + 2D(D - CB)^{-1} \end{bmatrix}. \tag{6.1.2}
\end{aligned}$$

Now we look at the second block row of  $\tau(g)$

$$\begin{bmatrix} 2D(D - CB)^{-1}C & -I + 2D(D - CB)^{-1} \end{bmatrix}.$$

Since  $D = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ , suppose the last  $t$  rows of  $D$  are zero, then this means that, in the last  $t$  rows of  $\tau(g)$ , the only nonzero elements are -1 which are from  $-I$ . They are all on the diagonal positions of  $\tau(g)$ , which tells us that  $\tau(g)$  has an eigenvalue  $\lambda = -1$  whose multiplicity is at least  $t$ .

Proof of Case 2:  $A$  and  $D$  are both singular. Without loss of generality, we assume that  $\text{rank}(A) \leq \text{rank}(D)$ . Similarly we can take  $A$  and  $D$  in following special form:

$$A = \begin{bmatrix} I_1 & 0 \\ 0 & 0 \end{bmatrix}, D = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}.$$

So  $A$  has more zero rows than  $D$  does.

In this case we will again find the eigenvalue  $\lambda = -1$  of  $\tau(g)$  explicitly. Actually we will find a submatrix of  $\tau(g)$  in  $\pm \text{SL}(m, \mathbb{F})$  for some integer  $m$ , which gives us the eigenvalue  $\lambda = -1$  of  $\tau(g)$ .

We first partition  $B$  and  $C$  according to ranks of  $A$  and  $D$

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

Now we make full use of freedom of choosing elements in  $K_{\mathbb{F}}gK_{\mathbb{F}}$ , i.e., make  $g$  to be a really special one. Notice that all elements in the same  $K_{\mathbb{F}}$ -orbit of any element in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  have same eigenvalues.

We will discuss three subcases:

1.  $B_{22} \neq 0$  and  $C_{22} \neq 0$ .
2.  $B_{22} = 0$  but  $C_{22} \neq 0$ .
3.  $C_{22} = 0$ .

First subcase. If  $B_{22} \neq 0$  and  $C_{22} \neq 0$ , then there exists nonsingular matrices  $f$  and  $h$  of proper size such that  $fC_{22}h = \begin{bmatrix} 0 & 0 \\ 0 & I_3 \end{bmatrix}$  (if  $C_{22}$  is nonsingular, we simply remove certain zero

blocks in the above form). Similarly, if  $B_{22} \neq 0$ , we can find  $s$  and  $t$  such that  $sB_{22}t = \begin{bmatrix} 0 & 0 \\ 0 & I_4 \end{bmatrix}$ .

The identity matrices in  $fC_{22}h$  and  $sB_{22}t$  need not be same. Let

$$k_1 = \left[ \begin{array}{cc|cc} \alpha I & 0 & & 0 \\ 0 & s & & \\ \hline & & I & 0 \\ 0 & & 0 & f \end{array} \right], k_2 = \left[ \begin{array}{cc|cc} \beta I & 0 & & 0 \\ 0 & h & & \\ \hline & & I & 0 \\ 0 & & 0 & t \end{array} \right],$$

where  $\alpha, \beta \in \overline{\mathbb{F}}$  make  $\det(k_1) = \det(k_2) = 1$ .

Then

$$\begin{aligned}
k_1 g k_2 &= \left[ \begin{array}{cc|c} \alpha I & 0 & 0 \\ 0 & s & \\ \hline & I & 0 \\ & 0 & f \end{array} \right] \left[ \begin{array}{cc|cc} * & 0 & B_{11} & B_{12} \\ 0 & 0 & B_{21} & B_{22} \\ \hline C_{11} & C_{12} & I & 0 \\ C_{21} & C_{22} & 0 & 0 \end{array} \right] \left[ \begin{array}{cc|c} \beta I & 0 & 0 \\ 0 & h & \\ \hline & I & 0 \\ & 0 & t \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & \alpha B_{11} & \alpha B_{12} t \\ 0 & 0 & s B_{21} & s B_{22} t \\ \hline \beta C_{11} & C_{12} h & I & 0 \\ \beta f C_{21} & f C_{22} h & 0 & 0 \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & \alpha B_{11} & \alpha B_{12} t \\ 0 & 0 & s B_{21} & 0 \\ \hline \beta C_{11} & C_{12} h & I & 0 \\ \beta f C_{21} & 0 & 0 & 0 \\ \hline & & I_4 & \\ & & & I_3 \end{array} \right].
\end{aligned}$$

Now let the last column of  $C_{12}h$  be  $u$ , and the last row of  $sB_{21}$  be  $v$ . Let

$$k_3 = \left[ \begin{array}{cc|cc} I_1 & 0 & & \\ 0 & I & & \\ \hline & & I_2 & -u \\ & & 0 & I \end{array} \right], k_4 = \left[ \begin{array}{cc|cc} I_1 & 0 & & \\ 0 & I & & \\ \hline & & I_2 & 0 \\ & & -v & I \end{array} \right].$$

Now

$$\begin{aligned}
& k_3(k_1 g k_2) k_4 \\
&= \left[ \begin{array}{cc|c} I_1 & 0 & 0 \\ 0 & I & \\ \hline 0 & I_2 & -u \\ & 0 & I \end{array} \right] \left[ \begin{array}{cc|cc} * & 0 & \alpha B_{11} & \alpha B_{12} t \\ 0 & 0 & s B_{21} & 0 \\ \hline \beta C_{11} & C_{12} h & & \\ \beta f C_{21} & 0 & I & 0 \\ & I_3 & 0 & 0 \end{array} \right] \left[ \begin{array}{cc|c} I_1 & 0 & 0 \\ 0 & I & \\ \hline 0 & I_2 & 0 \\ & -v & I \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & * & \alpha B_{12} t \\ 0 & 0 & * & 0 \\ & & 0 & I_4 \\ \hline * & * & 0 & \\ \beta f C_{21} & 0 & & \\ & I_3 & I & 0 \\ & & 0 & 0 \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & * & (\alpha B_{12} t)_1 & (\alpha B_{12} t)_2 \\ 0 & 0 & * & 0 & \\ & & 0 & I_4 & \\ & & & & 1 \\ \hline * & * & 0 & & \\ \hline (\beta f C_{21})_1 & 0 & & I & 0 \\ (\beta f C_{21})_2 & I_3 & & 0 & 0 \\ & & 1 & & \end{array} \right] \\
&= \begin{bmatrix} (k_3(k_1 g k_2) k_4)_{11} & (k_3(k_1 g k_2) k_4)_{12} \\ (k_3(k_1 g k_2) k_4)_{21} & (k_3(k_1 g k_2) k_4)_{22} \end{bmatrix}
\end{aligned}$$

By adding the last column of  $\begin{bmatrix} (k_3(k_1 g k_2) k_4)_{11} \\ (k_3(k_1 g k_2) k_4)_{21} \end{bmatrix}$  to all other columns we are able to make  $(\beta f C_{21})_2 = 0$ . This is equivalent to multiply  $k_3(k_1 g k_2) k_4$  by some element  $k_6$  in  $K_{\mathbb{F}}$  from the right. In the same way we are able to make  $(\alpha B_{12} t)_2 = 0$  by multiplying  $(k_3(k_1 g k_2) k_4) k_6$  by some  $k_5 \in K_{\mathbb{F}}$  from the left, i.e., we get following form

$$k_5(k_3(k_1 g k_2)k_4)k_6 = \left[ \begin{array}{cc|cc|c} & & * & 0 & \\ & & 0 & 0 & \\ \hline & & * & 0 & \\ & & 0 & 0 & \\ \hline & & 0 & & 1 \\ \hline * & * & 0 & & \\ \hline (\beta f C_{21})_1 & 0 & & & \\ & I_5 & & & \\ \hline 0 & & 1 & & \end{array} \right] \begin{array}{c} * \\ (\alpha B_{12} t)_1 \\ 0 \\ 0 \\ 1 \\ I \\ 0 \\ 0 \end{array} \right] \triangleq \bar{g}.$$

Now on the  $(n/2 - 1)$ -th and  $n$ -th rows and columns of  $\bar{g}$  there only two nonzero elements, 1, and we actually form following submatrix of  $\bar{g}$

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

We claim that the  $(n/2 - 1)$ -th and  $n$ -th rows and columns of  $(\bar{g})^{-1}$  is same as  $\bar{g}$ . Let  $P$  be a permutation matrix such that

$$P\bar{g}P^T = \begin{bmatrix} * & 0 \\ 0 & Q \end{bmatrix},$$

then

$$\begin{aligned} (P\bar{g}P^T)^{-1} &= \begin{bmatrix} * & 0 \\ 0 & Q \end{bmatrix}^{-1} \\ &= \begin{bmatrix} * & 0 \\ 0 & Q \end{bmatrix} (Q^{-1} = Q) \\ &= P(\bar{g})^{-1}P^T \end{aligned}$$

This implies that  $(\bar{g})^{-1} = P^T(P(\bar{g})^{-1}P^T)P$  has same "shape" as  $\bar{g}$  and elements on the  $(n/2 - 1)$ -th and  $n$ -th rows and columns of  $(\bar{g})^{-1}$  are same as those of  $\bar{g}$ .

Now let's look at  $\tau(\bar{g}) = \bar{g}p^{-1}(\bar{g})^{-1}p$ . The operation on  $(n/2 - 1)$ -th and  $n$ -th rows and

columns of  $\tau(\bar{g})$  is actually same as

$$\begin{aligned}
Qp^{-1}Q^{-1}p &= QpQp \\
&= (Qp)^2 \\
&= \left( \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right)^2 \\
&= \left( \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right)^2 \\
&= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}$$

All nonzero elements on  $(n/2 - 1)$ -th and  $n$ -th rows and columns of  $\tau(\bar{g})$  lie in  $Qp^{-1}Q^{-1}p$ , and it's obvious that  $\tau(\bar{g})$  has an eigenvalue  $\lambda = -1$ .

The second subcase. If  $B_{22} = 0$  but  $C_{22} \neq 0$ , following similar procedure we can assume  $g$  to have following special form:

$$g = \left[ \begin{array}{cc|cc} * & 0 & * & * \\ 0 & 0 & B_{21} & 0 \\ \hline * & * & 0 & \\ \hline * & 0 & I_2 & 0 \\ & I_5 & 0 & 0 \\ \hline 0 & & & 1 \end{array} \right].$$

$B_{21}$  must be full row-rank, and therefore there exist nonsingular  $f$  and  $h$  such that  $fB_{21}h = \begin{bmatrix} 0 & I \end{bmatrix}$ .

Let

$$k_1 = \left[ \begin{array}{cc|cc} \alpha I & 0 & 0 & \\ 0 & f & h^{-1} & 0 \\ \hline 0 & & 0 & I \end{array} \right], k_2 = \left[ \begin{array}{cc|cc} \beta I & 0 & 0 & \\ 0 & I & h & 0 \\ \hline 0 & & 0 & I \end{array} \right],$$

then

$$\begin{aligned}
k_1 g k_2 &= \left[ \begin{array}{cc|c} \alpha I & 0 & 0 \\ 0 & f & \\ \hline 0 & h^{-1} & 0 \\ & 0 & I \end{array} \right] \left[ \begin{array}{cc|cc} * & 0 & * & * \\ 0 & 0 & B_{21} & 0 \\ \hline * & * & 0 & \\ * & 0 & I_5 & \\ \hline 0 & & 1 & \end{array} \right] \left[ \begin{array}{cc|c} \beta I & 0 & 0 \\ 0 & I & \\ \hline 0 & h & 0 \\ & 0 & I \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & * & * \\ 0 & 0 & f B_{21} h & 0 \\ \hline * & * & 0 & \\ * & 0 & I_2 & 0 \\ * & I_5 & 0 & 0 \\ \hline 0 & & 1 & \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} * & 0 & * & * \\ 0 & 0 & 0 & I \\ \hline * & * & 0 & \\ * & 0 & I_2 & 0 \\ * & I_5 & 0 & 0 \\ \hline 0 & & 1 & \end{array} \right]
\end{aligned}$$

Suppose  $\text{rank}(D) = n_1$  and let  $\tilde{P}$  be the permutation matrix that permutes the  $(n/2 + n_1)$ -th and  $n$ -th row of  $k_1 g k_2$ . Let

$$k_3 = \left[ \begin{array}{c|cc} I & & \\ \hline & \alpha I & 0 \\ & 0 & \tilde{P} \end{array} \right],$$

then



$$\begin{aligned}
k_3(k_1 g k_2) &= \left[ \begin{array}{ccc|ccc} * & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \\
&= \overline{g}.
\end{aligned}$$

Considering the  $n/2$ -th and  $(n/2 + n_1)$ -th rows and columns of  $\overline{g}$  we can get submatrix  $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  again and similar to the first subcase, we can find permutation matrix  $P$  such that

$$\begin{aligned}
P\overline{g}P^T &= \left[ \begin{array}{c|c} * & * \\ \hline 0 & Q \end{array} \right] \\
P(\overline{g}p)P^T &= \left[ \begin{array}{c|cc} * & & * \\ \hline & 0 & -1 \\ 0 & 1 & 0 \end{array} \right] \\
P(p\overline{g})P^T &= \left[ \begin{array}{c|cc} * & & * \\ \hline & 0 & 1 \\ 0 & -1 & 0 \end{array} \right]
\end{aligned}$$

and

$$\begin{aligned}
P(\overline{g})^{-1}p^{-1}P^T &= P(p\overline{g})^{-1}P^T \\
&= (P(p\overline{g})P^T)^{-1} \\
&= \left[ \begin{array}{c|cc} * & & * \\ \hline & 0 & -1 \\ 0 & 1 & 0 \end{array} \right].
\end{aligned}$$

Then we have

$$\begin{aligned}
P\tau(\bar{g})P^T &= P(k_3(k_1gk_2))p^{-1}(\bar{g})^{-1}pP^T \\
&= P(\bar{g})pP^TP(\bar{g})^{-1}pP^T \\
&= (P(\bar{g}p)P^T)(P(\bar{g})^{-1}p^{-1}P^T) \quad (p = p^{-1}) \\
&= \left[ \begin{array}{c|cc} * & & * \\ \hline 0 & 0 & -1 \\ & 1 & 0 \end{array} \right] \left[ \begin{array}{c|cc} * & & * \\ \hline 0 & 0 & -1 \\ & 1 & 0 \end{array} \right] \\
&= \left[ \begin{array}{c|cc} * & & * \\ \hline 0 & -1 & 0 \\ & 0 & -1 \end{array} \right].
\end{aligned}$$

which tells us that  $P\tau(\bar{g})P^T$  or  $\tau(\bar{g})$  has an eigenvalue  $\lambda = -1$ .

Third subcase. We assume  $C_{21} = 0$  and we have immediately that  $C_{12}$  has full column-rank, so there exist nonsingular  $f$  and  $h$  such that  $fC_{12}h = \begin{bmatrix} I_3 \\ 0 \end{bmatrix}$ . Let

$$k_1 = \left[ \begin{array}{c|cc} I_1 & & \\ \hline & \alpha I & \\ & f & \\ & & I \end{array} \right], k_2 = \left[ \begin{array}{c|cc} I_1 & & \\ \hline & h & \\ & f^{-1} & \\ & & \beta I \end{array} \right],$$

then

$$\begin{aligned}
k_1 g k_2 &= \left[ \begin{array}{c|c} I_1 & \\ \hline \alpha I & \\ \hline & f \\ & I \end{array} \right] \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline * & C_{12} & I_2 & 0 \\ * & 0 & 0 & 0 \end{array} \right] \left[ \begin{array}{c|c} I_1 & \\ \hline h & \\ \hline & f^{-1} \\ & \beta I \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline * & f C_{12} h & I_2 & 0 \\ * & 0 & 0 & 0 \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline \overline{C}_{21} & I_3 & I_2 & 0 \\ & 0 & 0 & 0 \\ \hline * & 0 & & \end{array} \right] = \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline \overline{C}_{21}^1 & I_3 & I_2 & 0 \\ \overline{C}_{21}^2 & 0 & 0 & 0 \\ \hline * & 0 & & \end{array} \right].
\end{aligned}$$

Let

$$k_3 = \left[ \begin{array}{cc|c} I_1 & 0 & \\ -\overline{C}_{21}^1 & I_3 & \\ \hline & I_2 & \\ & & I_6 \end{array} \right],$$

then

$$\begin{aligned}
(k_1 g k_2) k_3 &= \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline \overline{C}_{21}^1 & I_3 & I_2 & 0 \\ \overline{C}_{21}^2 & 0 & 0 & 0 \\ * & 0 & & \end{array} \right] \left[ \begin{array}{cc|c} I_1 & 0 & \\ -\overline{C}_{21}^1 & I_3 & \\ \hline & & I_2 \\ & & I_6 \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline 0 & I_3 & I_2 & 0 \\ \overline{C}_{21}^2 & 0 & 0 & 0 \\ * & 0 & & \end{array} \right] \\
&= \tilde{g}.
\end{aligned}$$

Let

$$L = \left[ \begin{array}{cc|cc} I_1 & & & \\ & I_3 & I_3 & 0 & 0 \\ \hline & & I_3 & 0 \\ & & I_4 & 0 \\ & & 0 & 0 & I_6 \end{array} \right] \Rightarrow L^{-1} = \left[ \begin{array}{cc|cc} I_1 & & & \\ & I_3 & -I_3 & 0 & 0 \\ \hline & & I_3 & 0 \\ & & I_4 & 0 \\ & & 0 & 0 & I_6 \end{array} \right],$$

then

$$\begin{aligned}
L \tilde{g} &= \left[ \begin{array}{cc|cc} I_1 & & & \\ & I_3 & I_3 & 0 & 0 \\ \hline & & I_3 & 0 \\ & & I_4 & 0 \\ & & 0 & 0 & I_6 \end{array} \right] \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & 0 & * & * \\ \hline 0 & I_3 & I_2 & 0 \\ \overline{C}_{21}^2 & 0 & 0 & 0 \\ * & 0 & & \end{array} \right] \\
&= \left[ \begin{array}{cc|cc} I_1 & 0 & * & * \\ 0 & I_3 & * & * \\ \hline 0 & I_3 & I_2 & 0 \\ \overline{C}_{21}^2 & 0 & 0 & 0 \\ * & 0 & & \end{array} \right] = \begin{bmatrix} I & B \\ C & D \end{bmatrix} = \bar{g},
\end{aligned}$$

where  $D = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$ . So

$$\begin{aligned}\tilde{g} &= L^{-1}\bar{g}, \\ \tilde{g}^{-1} &= (\bar{g})^{-1}L\end{aligned}$$

$$\begin{aligned}p^{-1}Lp &= \begin{bmatrix} I & \\ & -I \end{bmatrix} \left[ \begin{array}{c|cc} I_1 & & \\ \hline & I_3 & \\ \hline & & I_3 & 0 & 0 \\ \hline & & I_3 & & 0 \\ & & I_4 & & 0 \\ \hline & & 0 & 0 & I_6 \end{array} \right] \begin{bmatrix} I & \\ & -I \end{bmatrix} \\ &= \begin{bmatrix} I_1 & & \\ \hline & I_3 & \\ \hline & & -I_3 & 0 & 0 \\ \hline & & I_3 & & 0 \\ & & I_4 & & 0 \\ \hline & & 0 & 0 & I_6 \end{bmatrix} \\ &= L^{-1}\end{aligned}$$

and

$$\begin{aligned}\tau(\tilde{g}) &= \tilde{g} p^{-1} \tilde{g}^{-1} p \\ &= L^{-1}\bar{g}p^{-1}(\bar{g})^{-1}Lp \\ &= L^{-1}(\bar{g}p^{-1}(\bar{g})^{-1}p)(p^{-1}Lp) \\ &= L^{-1}(\bar{g}p^{-1}(\bar{g})^{-1}p)L^{-1} \\ &= \begin{bmatrix} I_1 & & \\ \hline & I_3 & \\ \hline & & -I_3 & 0 & 0 \\ \hline & & I_3 & & 0 \\ & & I_4 & & 0 \\ \hline & & 0 & 0 & I_6 \end{bmatrix} \cdot \begin{bmatrix} I + 2B(D - CB)^{-1}C & 2B(D - CB)^{-1} \\ 2D(D - CB)^{-1}C & -I + 2D(D - CB)^{-1} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} I_1 & & \\ \hline & I_3 & \\ \hline & & -I_3 & 0 & 0 \\ \hline & & I_3 & & 0 \\ & & I_4 & & 0 \\ \hline & & 0 & 0 & I_6 \end{bmatrix}.\end{aligned}$$

The expression for  $\bar{g}p^{-1}(\bar{g})^{-1}p$  is from (6.1.2). It's clear that the last  $\text{rank}(I_6)$  rows and columns of  $\tau(\tilde{g})$  are same as the last  $\text{rank}(I_6)$  rows and columns of  $\bar{g}p^{-1}(\bar{g})^{-1}p = \tau(\bar{g})$ . From the analysis in the case that  $A = I$ ,  $D$  is singular, we can see that  $\tau(\tilde{g})$  has one eigenvalue  $\lambda = -1$ . This finishes the proof of Theorem 6.1.3.

Following lemma is needed in Section 6.1.2.

**Lemma 6.1.4.** *Let  $g \in \text{GL}(n, \mathbb{F})$  be a nonsingular matrix. Given  $s \leq n$  and indices  $\{i_1, \dots, i_s\}$  and  $\{j_1, \dots, j_s\}$ , if all nonzero entries of  $g$  lie only in rows  $\{i_1, \dots, i_s\}$  lie in columns  $\{j_1, \dots, j_s\}$ , the submatrix  $T$  of  $g$  formed by elements lying in the above rows and columns is nonsingular.*

*Proof.* By choosing special permutation matrices  $P$  and  $Q$  we are able to make  $T$  be at the right bottom corner of  $PgQ$ :

$$PgQ = \begin{bmatrix} \bar{g}_{11} & \bar{g}_{12} \\ \bar{g}_{12} & \bar{g}_{22} \end{bmatrix} = \begin{bmatrix} \bar{g}_{11} & \bar{g}_{12} \\ 0 & T \end{bmatrix}$$

$\bar{g}_{21} = 0$  because the only nonzero elements of  $PgQ$  lie in  $T$ .  $T$  is a square submatrix and it's obvious that  $T$  is nonsingular.  $\square$

### 6.1.2 Classification with $g$ in the special form

We assume here  $n$  is an even positive integer, i.e.,  $n \in 2\mathbb{Z}^+$  and  $n \geq 4$ .

Let  $G_{\mathbb{F}} = \text{SL}(n, \mathbb{F})$ ,  $K_{\mathbb{F}} = G_{\mathbb{F}}^{\sigma} = \left\{ \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \right\}$ , where  $\mathbb{F}$  is a field of characteristic not 2, not

necessarily algebraically closed, the involution  $\sigma$  is defined as  $\sigma(g) = p^{-1}gp$ ,  $p = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$ , and  $K_{\mathbb{F}} = G_{\mathbb{F}}^{\sigma}$  is the fix-point subgroup of  $\sigma$ .  $\tau(g) = g(\sigma(g))^{-1}$ .

We've seen that all unipotent elements in  $P$  can be generated by  $g = I + J$ , where  $J = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}$ , and using correspondence with double cosets in Chapter 3, we know that  $\tau(g) \stackrel{K_{\mathbb{F}}}{\sim} \tau(\bar{g}) \Leftrightarrow K_{\mathbb{F}}gK_{\mathbb{F}} = K_{\mathbb{F}}\bar{g}K_{\mathbb{F}}$ , or  $\bar{g} \in K_{\mathbb{F}}gK_{\mathbb{F}}$ , or  $\exists k, \bar{k} \in K_{\mathbb{F}}$  s.t.  $\bar{g} = kg\bar{k}$ . With the special form of  $g$  and  $\bar{g}$  we have immediately  $\bar{k} = k^{-1} \in K_{\mathbb{F}}$ .

Now

$$\begin{aligned}
\bar{g} &= kgk^{-1} \\
\begin{bmatrix} I & \bar{B} \\ \bar{C} & I \end{bmatrix} &= \begin{bmatrix} x & 0 \\ 0 & y \end{bmatrix} \begin{bmatrix} I & B \\ C & I \end{bmatrix} \begin{bmatrix} x^{-1} & 0 \\ 0 & y^{-1} \end{bmatrix} \\
\text{(I)} \quad &\Leftrightarrow \begin{cases} \bar{B} = xBy^{-1} \\ \bar{C} = yCx^{-1} \end{cases} \\
\text{(II)} \quad &\Leftrightarrow \begin{cases} xB = \bar{B}y \\ yC = \bar{C}x \end{cases}
\end{aligned}$$

*Remark 6.1.5.* It's obvious that  $\text{rank}(B) = \text{rank}(\bar{B})$  and  $\text{rank}(C) = \text{rank}(\bar{C})$ . We hereafter assume this condition to be true in the rest of this chapter.

*Remark 6.1.6.* There are two ways of achieving our goals:

1. Get bi-conditions which two  $K_{\mathbb{F}}$ -similar elements should satisfy directly.
2. Reduce  $g(\bar{g})$  to the canonical form directly and show that the canonical form can be chosen as a representative element in the  $K_{\mathbb{F}}$ -orbits.

We will combine both here. It's true that  $\exists k \in K_{\mathbb{F}}$  s.t.  $\bar{g} = kgk^{-1} \Leftrightarrow \exists k_1, k_2, k_3 \in K_{\mathbb{F}}$  s.t.  $k_1 \bar{g} k_1^{-1} = k_3 (k_2 g k_2^{-1}) k_3^{-1}$ . This means that we are able to take some  $k_1$  and  $k_2$  to make  $k_1 \bar{g} k_1^{-1}$  and  $k_2 g k_2^{-1}$  to be special first to make our discussion easier. This observation is more related to the second approach, and we will reduce  $g$  and  $\bar{g}$  to their " $K_{\mathbb{F}}$ -canonical form" directly and then show that those canonical forms can be chosen as representative elements in their  $K_{\mathbb{F}}$ -orbits, i.e., we "mainly" follow the second approach while we use the first approach to help us see where we should go and what kind of conditions they should satisfy.

If we use  $x, y$  acts on  $B, C$  like in **(I)**, this means that, more likely, we are "simplifying"  $g$  or  $J$ .

If we use equations in **(II)**, this means that, while it's still possible for us to simplify  $g$  or  $J$ , we are more likely either to reduce the PROBLEM, or to consider conditions that  $g/J$  (or equivalently,  $B, C$ ) should satisfy.

*Remark 6.1.7.* If  $\mathbb{F} = \bar{\mathbb{F}}$ , then any  $k = \begin{bmatrix} x & \\ & y \end{bmatrix} \in \text{GL}(n, \mathbb{F})$  can be taken in our proof, i.e., we are able to forget the determinant condition.

We will put our classification into two subcases: 1,  $\det C \neq 0$ , and 2,  $\det C = 0$ . The first case is relatively easy and the second one requires much more work.

**Case**  $\det C \neq 0$

If  $\text{rank}(C) = \frac{n}{2}$ , i.e.,  $\det C \neq 0$ , and field  $\mathbb{F} = \overline{\mathbb{F}}$  we are able to assume  $C = \overline{C} = I$ , and in this case  $\bar{g} = kgk^{-1} \Leftrightarrow \overline{B} = xBx^{-1}$ , i.e.,  $\overline{B}$  and  $B$  are similar  $\Rightarrow$  the Jordan canonical forms give us all possible  $K_{\mathbb{F}}$ -orbits. This observation is included in the following Theorem:

**Theorem 6.1.8.** *Let*

$$J = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}, \overline{J} = \begin{bmatrix} 0 & \overline{B} \\ \overline{C} & 0 \end{bmatrix},$$

*be nilpotent matrices.*

*If  $\text{rank}(C) = \text{rank}(\overline{C}) = \frac{n}{2}$ , i.e.,  $\det(C) \cdot \det(\overline{C}) \neq 0$ , then*

*(I) If  $\mathbb{F} = \overline{\mathbb{F}}$ , then  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}$  if and only if  $\overline{BC}$  is similar to  $BC$  in the usual sense, or equivalently  $J^2 \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}^2$ ;*

*(II) If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , then  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}$  if and only if following two conditions are satisfied*

*1.  $\det(C)/\det(\overline{C}) \in (\mathbb{F}^*)^2$ ;*

*2. using condition 1 and approach I we can make  $\det(C) = \det(\overline{C})$ . Then  $BC$  and  $\overline{BC}$  are nilpotent and satisfy conditions in Theorem 5.2.4.*

*Proof.* From

$$\begin{aligned} \overline{C} &= yCx^{-1} \\ 1 &= \det(x) \cdot \det(y) \end{aligned}$$

we have

$$\begin{aligned} \det(C)/\det(\overline{C}) &\in (k^*)^2 \\ \det(x) &= \pm \sqrt{\det(C)/\det(\overline{C})} \end{aligned} \tag{6.1.3}$$

From the the first one we have

$$y = \overline{C}x^{-1} \text{ or } y^{-1} = Cx^{-1}\overline{C}^{-1} \tag{6.1.4}$$

plug it into

$$\begin{aligned} \overline{B} &= xBy^{-1} = xBCx^{-1}\overline{C}^{-1} \\ \Rightarrow \overline{BC} &= xBCx^{-1} \end{aligned} \tag{6.1.5}$$

Since  $y$  is determined by  $x$  completely in (6.1.4), we only need to choose proper  $x$  to make all conditions satisfied. We have two conditions for  $x$ : (6.1.3) and (6.1.5). If  $\mathbb{F} = \overline{\mathbb{F}}$  we are done.



If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , first, using approach I we are able to take proper  $x_1$  and  $y_1$  to make  $\det(C) = \det(\overline{C})$ , and this gives us  $\det(x) = \pm 1$ . The rest is from Theorem 5.2.4.  $\square$

*Remark 6.1.9.* In the above Theorem 6.1.8, we are able to choose as representatives  $C$  and  $B$  to have following special form:

$$C = \begin{bmatrix} p & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix} \text{ or } B_2 = \begin{bmatrix} J_0 & & \\ & J_1 & \\ & & \ddots \\ & & & J_r \end{bmatrix}$$

where

$$J_i = \begin{bmatrix} 0 & q_i & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}, 1 \leq i \leq r,$$

If  $B = B_1$  then  $q_1 = q/p, q_i = 1, i \geq 2, p \in \mathbb{F}^*/(\mathbb{F}^*)^2$  and  $q \in \mathbb{F}_\pm^*/(\mathbb{F}_\pm^*)^h$ , where  $h$  is the greatest common factor of the sequence  $\text{size}(J_1), \dots, \text{size}(J_r)$ . If  $B = B_2$ , then  $q_i = 1, i = 1, \dots, r$ . Note that we have  $J_0 = 0$  in  $B_2$ , and in this case, from Remark 5.2.2 we can see that  $C$  and the Jordan canonical form of  $BC$  ( $\overline{BC}$ ) have been good enough to distinguish the  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$ .

**Case  $\det C = 0$**

Now we assume  $m = \text{rank}(C) = \text{rank}(\overline{C}) < \frac{n}{2}$ , and we then are able to choose  $x, y$  to make  $C = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ . This is true whether the field  $\mathbb{F} = \overline{\mathbb{F}}$  or not. It's like "transferring" determinant

conditions to  $B$  in the following procedure. Do same thing to  $\overline{C}$ , i.e., make  $\overline{C} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} = C$ .

Here we are concerned with the case  $\tau(g)$  unipotent  $\Leftrightarrow J$  nilpotent  $\Leftrightarrow BC$  nilpotent  $\Leftrightarrow$

$(B_{11})^N = 0$ , where  $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$  is partitioned according to the partition of  $C$ , ie  $B_{11} \in \mathbb{F}^{m \times m}$ . Throughout this paper we assume  $B$  (and therefore  $\overline{B}$ ) to be partitioned in this way.

To keep  $C = \overline{C} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ ,  $x, y$  need to satisfy

$$\begin{aligned} & \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ \Rightarrow & \begin{cases} x_{11} = y_{11} \stackrel{\text{def}}{=} \alpha \\ x_{12} = 0 \\ y_{21} = 0 \end{cases} \Rightarrow (\det \alpha)^2 \cdot \det x_{22} \cdot \det y_{22} = \det x \cdot \det y = 1 \neq 0. \end{aligned}$$

It's clear that  $y^{-1}$  has same shape as  $y$  (upper-triangular):

$$y^{-1} = \begin{bmatrix} \alpha^{-1} & * \\ 0 & y_{22}^{-1} \end{bmatrix}.$$

From  $\overline{B} = xBy^{-1}$  we have

$$\begin{aligned} & xB = \overline{B}y \\ & \begin{bmatrix} \alpha & 0 \\ x_{21} & x_{22} \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} \overline{B}_{11} & \overline{B}_{12} \\ \overline{B}_{21} & \overline{B}_{22} \end{bmatrix} \cdot \begin{bmatrix} \alpha & y_{12} \\ 0 & y_{22} \end{bmatrix} \end{aligned}$$

we get four equations equivalently:

$$\alpha B_{11} = \overline{B}_{11} \alpha \tag{6.1.6}$$

$$\alpha B_{12} = \overline{B}_{11} y_{12} + \overline{B}_{12} y_{22} \tag{6.1.7}$$

$$x_{21} B_{11} + x_{22} B_{21} = \overline{B}_{21} \alpha \tag{6.1.8}$$

$$x_{21} B_{12} + x_{22} B_{22} = \overline{B}_{21} y_{12} + \overline{B}_{22} y_{22}. \tag{6.1.9}$$

We have immediately  $B_{11} \sim \overline{B}_{11}$ , i.e.,  $B_{11}$  is similar to  $\overline{B}_{11}$ , and therefore, if we use approach (I) to make  $B_{11}$  and  $\overline{B}_{11}$  into their Jordan canonical form, we are able to assume, and therefore assume from now on

$$B_{11} = \overline{B}_{11} = \begin{bmatrix} J_0 & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_r \end{bmatrix},$$

where  $J_i, i = 0, 1, 2, \dots, r$  are the usual nilpotent Jordan blocks,  $J_0$  is the only zero Jordan block, sizes of  $J_i, i = 1, 2, \dots, r$  are nondecreasing. We are able to do this because we can take special  $x_{22}$  or  $y_{22}$  to make the determinant to be one.

We list all "manipulations" we will use to reduce  $B$  and  $\bar{B}$  in the following lemma:

**Lemma 6.1.10.** *Without changing  $B_{11}$  and  $C$ , we can apply following "manipulations" on  $B_{ij}, i, j = 1, 2$  (using approach (I)):*

- (a) *add columns of  $B_{11}$  to columns of  $B_{12}$  ( $B_{21}$  and  $B_{22}$  might change too);*
- (b) *add rows of  $B_{11}$  to rows of  $B_{21}$  ( $B_{12}$  and  $B_{22}$  might change too);*
- (c) *multiply  $\begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix}$  from right by some nonsingular matrix;*
- (d) *multiply  $[B_{21}, B_{22}]$  from left by some nonsingular matrix;*
- (e) *add the last row of any sub-block of  $[B_{11}, B_{12}]$  to the first row of any sub-block of  $[B_{11}, B_{12}]$  or to any row "belonging to"  $J_0$  ( $B_{12}$  changes correspondingly);*
- (f) *add the first column of any sub-block of  $\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$  to the last column of any sub-block of  $\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$  or to any column "belonging to"  $J_0$  ( $B_{21}$  changes correspondingly);*
- (g) *multiply the first  $\text{size}(J_0)$  rows of  $B_{12}$  from left by any nonsingular matrix;*
- (h) *multiply the first  $\text{size}(J_0)$  columns of  $B_{21}$  from right by any nonsingular matrix;*
- (i) *add rows in  $B_{12}$  corresponding to the last rows of each sub-block of  $B_{11}$  to  $B_{22}$  (this does not change  $B_{21}$ );*
- (j) *add columns in  $B_{21}$  corresponding to the first columns of each sub-block of  $B_{11}$  to  $B_{22}$  (this does not change  $B_{12}$ ).*

*Proof.* We only need to take proper  $x$  and  $y$  in (I). □

*Remark 6.1.11.* Using Lemma 6.1.10 we are able to assume and will assume in the rest of this chapter:

- (1) If  $(B_{11})_{ij} \neq 0$ , then the  $i$ -th rows of  $B_{12}$  and the  $j$ -th column of  $B_{21}$  are zeros..

To see this, we simply add proper multiple of the  $j$ -th column of  $B_{11}$  to corresponding columns of  $B_{12}$  to "clear up" the  $i$ -th rows of  $B_{12}$ . This will NOT affect any other rows of  $B_{12}$

since  $(B_{11})_{ij}$  is the only nonzero element in the  $j$ -th column of  $B_{11}$ . We are able to "clear up"  $j$ -th column of  $B_{21}$  too.

(2)  $B_{12} = [B_{12}^1, 0]$  and  $B_{21} = \begin{bmatrix} B_{21}^1 \\ 0 \end{bmatrix}$ , where  $B_{12}^1$  ( $B_{21}^1$ ) consists of linearly independent columns (rows, respectively) only.

To see this, we only need to choose proper  $y_{22}$  in (I) to multiply  $B_{12}$  from right by  $y_{22}^{-1}$ . We have similar observation for  $B_{21}$ .

(3) Part (1) and Part (2) are true for all fields.

(4) In the case that the field  $\mathbb{F} \neq \overline{\mathbb{F}}$ , the determinants of  $k$  generated in (a), (b), (e), (f), (i) and (j) are always 1, so we need to be careful when we apply (c), (d), (g) and (h).

Following Lemma 6.1.12 plays a key role in the problem reduction.

**Lemma 6.1.12.** *Let  $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$  and  $\overline{B} = \begin{bmatrix} \overline{B}_{11} & \overline{B}_{12} \\ \overline{B}_{21} & \overline{B}_{22} \end{bmatrix}$ . If  $B_{12} = \overline{B}_{12} = B_{12}^1 \in \mathbb{F}^{s \times t}$ ,  $B_{21} = \overline{B}_{21}$ ,  $\overline{B}_{22} = 0$ , and  $B_{22}$  is arbitrary, then  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J} (\Leftrightarrow g \stackrel{K_{\mathbb{F}}}{\sim} \overline{g} \Leftrightarrow \tau(g) \stackrel{K_{\mathbb{F}}}{\sim} \tau(\overline{g}))$ .*

*Proof.* We only need to show that equations (6.1.6)-(6.1.9) have a solution in  $K_{\mathbb{F}}$ .

Let  $\alpha = I$ ,  $x_{22} = I$ ,  $y_{22} = I$ ,  $y_{12} = 0$ , and  $x_{21}$  satisfy  $x_{21}B_{11} = 0$ . Then  $\det(x) \cdot \det(y) = 1$  and such  $x_{21}$  always exists. Actually we are able to choose  $x_{21}$  in the following way:

If  $(B_{11})_{i_0 j_0} = 1 \neq 0$ , we make the  $i_0$ -th column of  $x_{21}$  to be zero. At same time the  $i_0$ -th row of  $B_{12}$  is zero too (see Remark 6.1.11), and this implies that, if some row of  $B_{12}$  is not zero, the corresponding column in  $x_{21}$  need not be zero either. Since  $B_{12} = B_{12}^1$  has full column rank, we are able to choose  $s = \text{rank}(B_{12}^1)$  rows in  $B_{12}$  to form a nonsingular matrix in  $B_{12}$ . We assume the  $i_1, \dots, i_s$  rows are linearly independent. Then we let all columns of  $x_{21}$  be zeros except the  $i_1, \dots, i_s$  columns, which are to be determined. Let

$$\tilde{B}_{12} = \begin{bmatrix} (B_{12})_{i_1, :} \\ (B_{12})_{i_s, :} \end{bmatrix}, \tilde{x}_{21} = [(x_{21})_{:, i_1}, \dots, (x_{21})_{:, i_s}],$$

then in equation (6.1.9)  $x_{21} \cdot B_{12}$  can be replaced by  $\tilde{x}_{21} \cdot \tilde{B}_{12}$  since  $x_{21} \cdot B_{12} = \tilde{x}_{21} \cdot \tilde{B}_{12}$ .  $\det(\tilde{B}_{12}) \neq 0$  and  $\tilde{x}_{21}$  can be anything  $\Rightarrow$  equation (6.1.9) always has solution  $\Rightarrow$  equations (6.1.6)-(6.1.9) are all satisfied.  $\square$

*Remark 6.1.13.* (1) This means that, if  $B_{12}$  has full column rank, we are always able to assume  $B_{22} = 0$  directly.

(2) If  $B_{21}$  has full row rank we are able to do same thing.

**Corollary 6.1.14.** *We are able to almost always assume*

$$B = \left[ \begin{array}{c|c|c} B_{11} & B_{12}^1 & 0 \\ \hline B_{21}^1 & 0 & 0 \\ \hline 0 & 0 & I \end{array} \right],$$

where either  $B_{12}^1$  has full column rank or  $B_{21}^1$  has full row rank, and  $I$  is the identity matrix of proper size.

The only exception is the case that,  $\mathbb{F} \neq \overline{\mathbb{F}}$ ,  $B_{12} = 0$ ,  $B_{21} = 0$ ,  $\text{rank}(B_{22}) = n/2 - \text{rank}(C)$ , and  $\det(B_{22}) \neq 0 \notin (\mathbb{F}^*)^2$ .

*Proof.* We assume  $\text{rank}(B_{12}^1) \geq \text{rank}(B_{21}^1)$  and  $\text{rank}(B_{12}^1) \geq 1$  here. The proof for the case  $\text{rank}(B_{12}^1) \leq \text{rank}(B_{21}^1)$  and  $\text{rank}(B_{21}^1) \geq 1$  is similar.

By "manipulations" in Lemma 6.1.10, we are able to make  $B$  have the form

$$B = \left[ \begin{array}{c|c|c} B_{11} & B_{12}^1 & 0 \\ \hline B_{21}^1 & B_{22}^{11} & B_{22}^{12} \\ \hline 0 & B_{22}^{21} & B_{22}^{22} \end{array} \right]$$

Here  $B_{22}^{22}$  need not be square since  $\text{rank}(B_{12}^1)$  need not equal  $\text{rank}(B_{21}^1)$ .

By Lemma 6.1.12 we are able to assume  $B_{22}^{11} = 0$ . Using part (i) in Lemma 6.1.10, we are able to make  $B_{22}^{21} = 0$  for following two reasons:

1. all rows in  $B_{11}$  corresponding to nonzero rows in  $B_{12}^1$  are zeros; and
2.  $B_{12}^1$  has full column rank.

Similarly we are able to make  $B_{22}^{12} = 0$  (and redo the proof of Lemma 6.1.12 in a same way if we like). Using part (c) and (d) in Lemma 6.1.10, we are able to multiply  $B_{22}^{22}$  by two nonsingular matrices from left and right to make  $B_{22}^{22} = \left[ \begin{array}{c|c} 0 & 0 \\ \hline 0 & I \end{array} \right]$ , without affecting  $B_{21}^1$  and  $B_{12}^1$  (just need to make sure that these two matrices have proper sizes). For the determinant condition here, we actually "transfer" it to columns of  $B_{12}^1$  since  $B_{12}^1 \neq 0$  here.

If  $\text{rank}(B_{12}^1) = \text{rank}(B_{21}^1) = 0$  and  $\det B_{22} = 0$ , we can transfer the determinant condition to the zero diagonal part of  $B_{22}$ , so we are fine in this case.

If  $\text{rank}(B_{12}^1) = \text{rank}(B_{21}^1) = 0$  and  $\det B_{22} \neq 0$  and  $\mathbb{F} \neq \overline{\mathbb{F}}$ , we cannot "transfer" determinant condition to  $B_{12}$  or  $B_{21}$ , or "some zero" in  $B_{22}$ . We have

$$\det(\alpha)^2 \cdot \det(x_{22}) \cdot \det(y_{22}) = 1$$

and use  $x_{22}$  and  $y_{22}$  to reduce  $B_{22}$  to a canonical form

$$x_{22}B_{22}y_{22}^{-1} = \overline{B}_{22} \Rightarrow \det(\alpha)^2 \cdot \det(x_{22})^2 = \det(\overline{B}_{22})/\det(B_{22}).$$

In this case we are able to choose following  $g \in G_{\mathbb{F}}$  as representative

$$g = I + J, J = \begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix},$$

$$B_{11} = \begin{bmatrix} J_0 & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_r \end{bmatrix}, B_{22} = \begin{bmatrix} p & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix},$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^2$ . □

*Remark 6.1.15.* We assume from now on in this chapter  $\max\{\text{rank}(B_{12}^1), \text{rank}(B_{21}^1)\} \geq 1$  and  $B$  and  $\overline{B}$  always have the form in Corollary 6.1.14.

Now we are able to introduce following theorem, which reduces the original problem to a special case only.

**Theorem 6.1.16.** *From  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}$  we have*

- (a)  $\text{rank}([B_{11}, B_{12}]) = \text{rank}([B_{11}, \overline{B}_{12}]) \Rightarrow \text{rank}(B_{12}^1) = \text{rank}(\overline{B}_{12}^1);$   
 $\text{rank}\left(\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} B_{11} \\ \overline{B}_{21} \end{bmatrix}\right) \Rightarrow \text{rank}(B_{21}^1) = \text{rank}(\overline{B}_{21}^1);$
- (b) *Part (a) implies  $\text{rank}(B_{22}^{22}) = \text{rank}(\overline{B}_{22}^{22})$  (or equivalently  $B_{22} = \overline{B}_{22}$ );*
- (c) *we have nonsingular solution for equations (6.1.6)-(6.1.9) for new  $\tilde{J}$  and  $\tilde{\overline{J}}$  with  $\tilde{B} =$*   

$$\left[ \begin{array}{c|c} B_{11} & B_{12}^1 \\ \hline \overline{B}_{21}^1 & 0 \\ \hline 0 & \end{array} \right] \text{ and } \tilde{\overline{B}} = \left[ \begin{array}{c|c} B_{11} & \overline{B}_{12}^1 \\ \hline \overline{B}_{21}^1 & 0 \\ \hline 0 & \end{array} \right], \text{ if } \text{rank}(B_{12}^1) \geq \text{rank}(B_{21}^1). \text{ We have same conclu-}$$
  
*sion if  $\text{rank}(B_{12}^1) \leq \text{rank}(B_{21}^1)$ .*

*Conversely, if part (a)-(c) are true for  $J$  and  $\overline{J}$ , we have  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}$ .*

*Remark 6.1.17.* Theorem 6.1.16 allows us to study a special included in part (c) only, for which corresponding equations (6.1.6)-(6.1.9) are much simpler. In this case, we able to find the

" $K_{\mathbb{F}}$ -canonical form of  $B$  (or  $J$ ) directly, for which equations (6.1.6)-(6.1.9) are automatically satisfied.

*Proof.* (a)  $J \stackrel{K_{\mathbb{F}}}{\sim} \bar{J} \Rightarrow xB = \bar{B}y \Rightarrow \alpha \cdot [B_{11}, B_{12}] = [B_{11}, \bar{B}_{12}] \cdot y \stackrel{\det \alpha \cdot \det y \neq 0}{\Rightarrow} \text{rank}([B_{11}, B_{12}]) = \text{rank}(B_{11}, \bar{B}_{12})$ .

From Remark 6.1.11 we have

$$\begin{aligned} \text{rank}([B_{11}, B_{12}]) &= \text{rank}(B_{11}) + \text{rank}(B_{12}) \\ \text{rank}([B_{11}, \bar{B}_{12}]) &= \text{rank}(B_{11}) + \text{rank}(\bar{B}_{12}) \end{aligned}$$

which implies

$$\text{rank}(B_{12}) = \text{rank}(\bar{B}_{12}).$$

The proof for the second part of part (a) is similar.

(b) Since  $B$  and  $\bar{B}$  are in the form of Corollary 6.1.14, we can see that

$$\begin{aligned} \text{rank}(B) &= \text{rank}(B_{11}) + \text{rank}(B_{12}^1) + \text{rank}(B_{21}^1) + \text{rank}(B_{22}^{22}) \\ \text{rank}(\bar{B}) &= \text{rank}(B_{11}) + \text{rank}(\bar{B}_{12}^1) + \text{rank}(\bar{B}_{21}^1) + \text{rank}(\bar{B}_{22}^{22}) \\ \text{rank}(B) &= \text{rank}(\bar{B}) \end{aligned}$$

which implies

$$\text{rank}(B_{22}^{22}) = \text{rank}(\bar{B}_{22}^{22}) \text{ or } \text{rank}(B_{22}) = \text{rank}(\bar{B}_{22}).$$

(c) We will show that we actually are able to pick up corresponding rows and columns in  $x$  and  $y$  to form the new  $\tilde{x}$  and  $\tilde{y}$  to satisfy all four equations for the new  $\tilde{B}$  and  $\tilde{\bar{B}}$  with  $\det(\tilde{x}) \det(\tilde{y}) = 1$ .

We here only prove the case that  $\text{rank}(B_{12}^1) \geq \text{rank}(B_{21}^1)$ . Let  $s = \text{rank}(B_{12}^1)$ . we take same  $\alpha$ , the first  $s$  columns of  $y_{12}$ , the first  $s$  rows of  $x_{21}$  and  $s \times s$  principle submatrix of  $x_{22}$  and  $y_{22}$  to form the new  $\tilde{x}$  and  $\tilde{y}$ . Equations (6.1.6)-(6.1.8) are satisfied immediately. For equation (6.1.9), since we are considering  $s \times s$  principle submatrix,  $x_{22} \cdot B_{22}$  and  $\bar{B}_{22} \cdot y_{22}$  disappear in equation (6.1.9), which actually corresponds to the condition  $\tilde{B}_{22} = \tilde{\bar{B}}_{22}$ .

The corresponding part of  $x_{21} \cdot B_{12}$  and  $\bar{B}_{21} \cdot y_{12}$  in original equation (6.1.9) are just new equation (6.1.9) in the new problem, and this implies that for new problem equation (6.1.6)-(6.1.9) have solution(s). We still need to show that we are able to take nonsingular  $\tilde{x}_{22}$  and  $\tilde{y}_{22}$  from  $x_{22}$  and  $y_{22}$ , and  $\det(\tilde{x}) \cdot \det(\tilde{y}) = 1$  which is requirement of the original problem.

Let  $B_{11} \in \mathbb{F}^{t \times t}$ . The only possibly nonzero rows of  $B_{12}$  and  $\bar{B}_{12}$  are those corresponding to

zero rows in  $B_{11}$ , so in the addition "+" of equation (6.1.7)

$$\alpha B_{12} = \overline{B}_{11}y_{12} + \overline{B}_{12}y_{22}$$

$B_{11} \cdot y_{12}$  and  $\overline{B}_{12} \cdot y_{22}$  have no "intersection". The last  $t - s$  columns of  $B_{12}$  are zeros  $\Rightarrow$  the last  $t - s$  columns of  $\alpha B_{12} = 0 \Rightarrow$  the last  $t - s$  columns of  $\overline{B}_{12}y_{22} = 0$ . Let

$$y_{22} = \begin{bmatrix} y_{22}^1 \\ y_{22}^2 \end{bmatrix} = \left[ \begin{array}{c|c} y_{22}^{11} & y_{22}^{12} \\ \hline y_{22}^{21} & y_{22}^{22} \end{array} \right], x_{22} = [x_{22}^1, x_{22}^2] = \left[ \begin{array}{c|c} x_{22}^{11} & x_{22}^{12} \\ \hline x_{22}^{21} & x_{22}^{22} \end{array} \right],$$

where  $y_{22}^{11} = \tilde{y}_{22} \in \mathbb{F}^{s \times s}$  and  $x_{22}^{11} = \tilde{x}_{22} \in \mathbb{F}^{s \times s}$ . Therefore

$$\begin{aligned} \overline{B}_{12} \cdot y_{22} &= [\overline{B}_{12}^1, 0] \cdot \begin{bmatrix} y_{22}^1 \\ y_{22}^2 \end{bmatrix} = \overline{B}_{12}^1 \cdot y_{22}^1 = [\overline{B}_{12}^1 \cdot y_{22}^{11}, \overline{B}_{12}^1 \cdot y_{22}^{12}] = [\overline{B}_{12}^1 \cdot y_{22}^{11}, 0] \\ \Rightarrow \overline{B}_{12}^1 \cdot y_{22}^{12} &= 0 \Rightarrow y_{22}^{12} = 0. \end{aligned}$$

and in equation (6.1.7) we have no condition on  $y_{22}^{22}$  since it multiplies the zero block.

If  $\text{rank}(B_{12}^1) = \text{rank}(B_{21}^1)$ , from equation (6.1.8) we have similar conclusion:  $x_{22}^{21} = 0$  and no condition on  $x_{22}^{22}$ .

For equation (6.1.9)

$$x_{21}B_{12} + x_{22}B_{22} = \overline{B}_{21}y_{12} + \overline{B}_{22}y_{22}$$

we let

$$\begin{aligned} M &= x_{21}B_{12} + x_{22}B_{22} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \\ N &= \overline{B}_{21}y_{12} + \overline{B}_{22}y_{22} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \end{aligned}$$

where  $M_{11}, N_{11} \in \mathbb{F}^{s \times s}$ .

Thanks to the special structure of  $B_{22} = \overline{B}_{22}$ ,  $x_{22}^{22}$  and  $y_{22}^{22}$  appear in  $M_{22}$  and  $N_{22}$  only. The corresponding part of the equation (6.1.9) is

$$M_{22} = N_{22}$$

or

$$x_{22}^{22} \cdot \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \cdot y_{22}^{22}$$

and this is the ONLY condition on  $x_{22}^{22}$  and  $y_{22}^{22}$  from equations (6.1.6)-(6.1.9). We can definitely



take

$$x_{22}^{22} = y_{22}^{22} = I$$

to satisfy all four equations, and we get at same time that

$$\begin{aligned} 1 &= \det(x) \det(y) = \det(\tilde{x}) \det(x_{22}^{22}) \cdot \det(\tilde{y}) \det(y_{22}^{22}) \\ \Rightarrow \det(\tilde{x}) \det(\tilde{y}) &= 1 \end{aligned}$$

Now we assume  $\text{rank}(B_{12}^1) > \text{rank}(B_{21}^1) = l$ . We still assume  $x_{22}$  and  $y_{22}$  have above form but  $x_{22}^{11} \in \mathbb{F}^{l \times l}$ .

We similarly have  $x_{22}^{21} \cdot B_{21}^1 = 0 \Rightarrow x_{22}^{21} = 0$  ( $B_{21}^1$  has full row rank). This implies  $x_{22}^{11}$  is nonsingular, but its size is less than "required":  $m < s$ .

Let ( $Q = x_{22}^{22}$ )

$$x_{22} = \left[ \begin{array}{c|c|c} x_{22}^{11} & P & * \\ \hline 0 & Q^{11} & Q^{12} \\ \hline 0 & Q^{21} & Q^{22} \end{array} \right],$$

where

$$\tilde{x}_{22} = \left[ \begin{array}{c|c} x_{22}^{11} & P \\ \hline 0 & Q^{11} \end{array} \right] \in \mathbb{F}^{s \times s}.$$

Similar to the case  $\text{rank}(B_{12}^1) = \text{rank}(B_{21}^1)$ , the condition on  $Q^{21}$  and  $Q^{22}$  from equation (6.1.6)-(6.1.9) is still  $M_{22} = N_{22}$  and actually there is no condition on  $Q^{21}$ . We again are able to take  $Q^{22} = y_{22}^{22} = I$ , and let  $Q^{21} = 0$ . We then still have  $\det(\tilde{x}) \det(\tilde{y}) = 1$ , which proves part (c).

(II) Actually part (c) and condition  $\text{rank}(B) = \text{rank}(\overline{B})$  (which guarantees  $\text{rank}(B_{22}^{22}) = \text{rank}(\overline{B}_{22}^{22})$ ) have been enough to make  $J \stackrel{K_{\mathbb{F}}}{\sim} \overline{J}$ . Similar to the "converse process" in the proof of part (c), we simply take  $x_{22}^{22} = I, y_{22}^{22} = I$ , and fill all other spots with zeros.  $\square$

From now on we focus on the case in Lemma 6.1.12 only:

$s = \text{rank}(B_{12}) = \text{rank}(B_{12}^1) = \text{rank}(\overline{B}_{12}) = \text{rank}(\overline{B}_{12}^1)$ , i.e.,  $B_{12}$  and  $\overline{B}_{12}$  have same full column rank. The "canonical" form for the general  $B$  can be obtained by "patching" some identity matrix of proper size at the right bottom corner. We assume here that  $\text{rank}(B_{12}^1) \geq \text{rank}(B_{21}^1)$ . For the case that  $\text{rank}(B_{21}) = \text{rank}(B_{21}^1) = \text{rank}(\overline{B}_{21}) = \text{rank}(\overline{B}_{21}^1)$  we can simply do the same thing.

Let's look at what  $\alpha$  looks like first. We partition  $\alpha$  according to Jordan blocks of  $B_{11}$ :  $\alpha = (\alpha_{ij})$  with  $\alpha_{ij} = (x_{ij}^{kl})$ .

$\alpha \cdot B_{11} = B_{11} \alpha \Leftrightarrow \alpha_{ij} \cdot J_j = J_i \cdot \alpha_{ij}$  which implies:

If  $i = j = 0$ ,  $\alpha_{00}$  can be anything;

If  $i \cdot j \neq 0$ , we have:  $\alpha_{ij}^{k,l} = \alpha_{ij}^{k+1,l+1}$ , which means that all elements on the same diagonal are all same, and all elements under the upper main diagonal are zeros, and especially,

If  $i = 0, j \neq 0$ , all possibly nonzero elements lie in the last column of  $\alpha_{0j}$ ;

If  $i \neq 0, j = 0$ , all possibly nonzero elements lie in the first row of  $\alpha_{i0}$ .

To be specific,  $\alpha$  looks like, for example,

$$\alpha = \begin{bmatrix} & & & & & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & & & & & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & & & & * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & & & & & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & & & & & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & * & * & * & * & * & * & * & * & 0 & * & * & * & * & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & 0 & 0 & * & * & * & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ * & * & * & * & * & * & * & * & * & 0 & * & * & * & * & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & * & * & * & * & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & * & * & * & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & * & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \quad (6.1.10)$$

$$= \begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} & \alpha_{03} \\ \alpha_{10} & \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{20} & \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{30} & \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix},$$

where, for each  $i, j \geq 1$ , all elements on the same diagonal are same.

We partition  $B_{12}$  (or  $B_{12}^1$ ) and  $B_{21}$  correspondingly. Since we only need to consider rows in  $B_{12}$  corresponding to the zero rows in  $B_{11}$ , sometimes we only write down these possibly nonzero rows in  $B_{12}$ , i.e., we "squeeze" all must zero rows out of  $B_{12}$ , and still use  $B_{12}$  to represent the squeezed  $B_{12}$ . Here is what I mean by "squeezing". Corresponding to the above

$\alpha$  the original  $B_{12}$  the "squeezed"  $B_{12}$  (defined after  $\triangleq$ ) are

$$\begin{aligned}
 B_{12} = & \left[ \begin{array}{cccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \end{array} \right] = \left[ \begin{array}{c} (b^0)_1 \\ (b^0)_2 \\ (b^0)_3 \\ (b^0)_4 \\ (b^0)_5 \\ \hline 0 \\ 0 \\ 0 \\ (b^1)_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ (b^1)_2 \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ (b^1)_3 \end{array} \right] \triangleq \left[ \begin{array}{c} (b^0)_1 \\ (b^0)_2 \\ (b^0)_3 \\ (b^0)_4 \\ (b^0)_5 \\ \hline (b^1)_1 \\ (b^1)_2 \\ (b^1)_3 \end{array} \right] = \left[ \begin{array}{c} b^0 \\ \hline b^1 \end{array} \right]
 \end{aligned}$$

Similarly  $B_{21}$  and the squeezed  $B_{21}$  are

$$\begin{aligned}
 B_{21} = & \left[ \begin{array}{cc|cc|cc|cc} * & * & * & 0 & * & 0 & 0 & * & 0 & 0 & 0 \\ * & * & * & 0 & * & 0 & 0 & * & 0 & 0 & 0 \\ * & * & * & 0 & * & 0 & 0 & * & 0 & 0 & 0 \\ * & * & * & 0 & * & 0 & 0 & * & 0 & 0 & 0 \end{array} \right] \\
 = & \left[ \begin{array}{cc|cc|cc|cc} (c^0)_1 & (c^0)_2 & (c^1)_1 & 0 & (c^1)_2 & 0 & 0 & (c^1)_3 & 0 & 0 & 0 \end{array} \right] \\
 \triangleq & \left[ \begin{array}{cc|cc|cc} (c^0)_1 & (c^0)_2 & (c^1)_1 & (c^1)_2 & (c^1)_3 \end{array} \right] \text{ (squeezed)} \\
 = & \left[ \begin{array}{c|c} c^0 & c^1 \end{array} \right],
 \end{aligned}$$

The meaning/definition/notations of  $B_{12}$  and  $B_{21}$  should be clear from its context. We then

are able to assume

$$B_{12} = \left[ \begin{array}{c|c} b^0 & \\ \hline & b^1 \end{array} \right] = \left[ \begin{array}{c|c} b^{00} & b^{01} \\ \hline b^{10} & b^{11} \end{array} \right], B_{21} = \left[ \begin{array}{c} c^0 \\ \parallel \\ c^1 \end{array} \right] = \left[ \begin{array}{c|c} c^{00} & c^{01} \\ \hline c^{10} & c^{11} \end{array} \right].$$

Same for  $\overline{B}_{12}$  and  $\overline{B}_{21}$ . While we still use  $B_{12}, B_{21}, \overline{B}_{12}$  and  $\overline{B}_{21}$  to represent the squeezed form, in many cases we also use lower-case letters  $b$  and  $c$  to represent the squeezed form of  $B_{12}$  and  $B_{21}$  in order to make our notations not too complicated.

Using part (c) in Lemma 6.1.10 we are able to assume  $b^1 = [b^{10}, b^{11}] = [0, b^{11}]$ , i.e., the last  $\text{rank}(b^1)$  columns of  $b^1$  are the only nonzero columns and they are linearly independent, and we can further make  $b^{11}$  to be in an upper-triangular form. Now using part (e) in Lemma 6.1.10 we can make  $b^{01} = 0$  ( $b^{11}$  has full column rank) and therefore

$$B_{12} = \left[ \begin{array}{c|c} b^{00} & 0 \\ \hline 0 & b^{11} \end{array} \right].$$

Using part (g) in Lemma 6.1.10 we are able to make  $b^{00} = \begin{bmatrix} 0 \\ I \end{bmatrix}$  and make  $c^{00}$  to be diagonal-like submatrix. This is because  $B_{12}$  has full column rank. The determinant condition now is "transferred" to  $B_{21}$ . If  $\mathbb{F} = \overline{\mathbb{F}}$ , we can make those "diagonals" in  $B_{21}$  to be one since we don't need to consider the determinant condition in this case (Remark 6.1.7).

**Theorem 6.1.18.** *Let  $B_{21} = [(B_{21})^1, (B_{21})^2], \overline{B}_{21} = [(\overline{B}_{21})^1, (\overline{B}_{21})^2]$ , where  $(B_{21})^1, (\overline{B}_{21})^1 \in \mathbb{F}^{s \times s}$ . If  $\alpha = \alpha_{00}$ , i.e.,  $B_{11} = 0$ , we have*

(I) (a) *If  $s = \text{rank}(B_{12}) = \text{rank}(B_{21}) = \text{rank}((B_{21})^2)$ ,  $(B_{21})^1$  exists but  $(B_{21})^1 = 0$ ,  $\mathbb{F} \neq \overline{\mathbb{F}}$ , then equations (6.1.6)-(6.1.9) have solution(s) if and only if  $\det((B_{21})^2)/\det((\overline{B}_{21})^2) \in (\mathbb{F}^*)^2$ ;*

(I) (b) *If  $s = \text{rank}(B_{12}) = \text{rank}(B_{21}) = \text{rank}((B_{21})^2)$ ,  $B_{21} = (B_{21})^2$ , i.e.,  $B_{12}$  and  $B_{21}$  are square matrices, and  $(B_{21})^2$  does not exist (this means that  $B_{12} = b^{00} = I$ ),  $\mathbb{F} \neq \overline{\mathbb{F}}$ , then equations (6.1.6)-(6.1.9) have solution(s) if and only if  $\det((B_{21})^2)/\det((\overline{B}_{21})^2) \in (\mathbb{F}^*)^4$ ;*

(II) *If  $s = \text{rank}(B_{12}) = \text{rank}(B_{21}) = \text{rank}((B_{21})^1)$ ,  $(B_{21})^2 = 0$ ,  $\mathbb{F} \neq \overline{\mathbb{F}}$ , then equations (6.1.6)-(6.1.9) have solution(s) if and only if  $(B_{21})^1 \in \mathbb{F}^{s \times s}$  and  $\det((B_{21})^1)/\det((\overline{B}_{21})^1) \in (\mathbb{F}^*)^3$ ;*

(III) *In all other cases, i.e., except the above two cases, equations (6.1.6)-(6.1.9) have solution(s) if and only if  $\begin{cases} \text{rank}(B_{21}) = \text{rank}(\overline{B}_{21}) \\ \text{rank}((B_{21})^1) = \text{rank}((\overline{B}_{21})^1) \end{cases}$ ,*

*Proof.* We proof (III) first.

Let  $B_{12} = B_{12}^1 = \begin{bmatrix} 0 \\ I \end{bmatrix}$  in this case ( $B_{12}$  has full column rank).

Necessity. Let  $\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$  with  $\alpha_{22} \in k^{s \times s}$ .

From equation (6.1.7):

$$\alpha B_{12} = B_{11}y_{12} + \bar{B}_{12}y_{22} = \bar{B}_{12}y_{22} = B_{12}y_{22},$$

we have  $\alpha_{12} = 0$ , which implies that  $\alpha_{11}$  and  $\alpha_{22}$  are both nonsingular.

From equation (6.1.8):

$$x_{21}B_{11} + x_{22}B_{21} = \bar{B}_{21}\alpha \text{ or } x_{22}B_{21} = \bar{B}_{21}\alpha,$$

we have

$$x_{22}(B_{21})^2 = (\bar{B}_{21})^2\alpha_{22} \Rightarrow \text{rank}((\bar{B}_{21})^2) = \text{rank}((B_{21})^2) \quad (\det(x_{22}) \cdot \det(\alpha_{22}) \neq 0).$$

Sufficiency. Let  $x_{21} = 0, y_{12} = 0 \Rightarrow$  equation (6.1.9) is satisfied.

Equation (6.1.6) is always true.

Let  $\alpha = \begin{bmatrix} \alpha_{11} & 0 \\ \alpha_{21} & \alpha_{22} \end{bmatrix}, y_{22} = \alpha_{11} \Rightarrow$  equation (6.1.7) is satisfied,  $\det(\alpha_{11}) \cdot \det(\alpha_{22}) \neq 0$ ,  
and

$$\alpha^{-1} = \begin{bmatrix} \alpha_{11}^{-1} & 0 \\ \tilde{\alpha}_{21} & \alpha_{22}^{-1} \end{bmatrix}.$$

The determinant condition becomes in this case

$$1 = \det(\alpha)^2 \det(x_{22}) \det(y_{22}) = \det(\alpha_{11})^2 \det(\alpha_{22})^3 \det(x_{22}).$$

This implies

$$\begin{aligned} \det(x_{22}) / \det(\alpha_{22}) &= \frac{1}{\det(\alpha_{11})^2 \det(\alpha_{22})^4} \\ &= \left( \frac{1}{\det(\alpha_{11}) \det(\alpha_{22})^2} \right)^2 \in (\mathbb{F}^*)^2 \end{aligned}$$

and

$$\begin{aligned} \det(x_{22}) / \det(\alpha_{11}) &= \frac{1}{\det(\alpha_{11})^3 \det(\alpha_{22})^3} \\ &= \left( \frac{1}{\det(\alpha_{11}) \det(\alpha_{22})} \right)^3 \in (\mathbb{F}^*)^3. \end{aligned}$$

Also, if  $B$  is not included in the first two cases, we always have four subcases:

(a)  $\text{rank}(B_{21}) = \text{rank}(B_{12})$ ,  $0 < \text{rank}((B_{21})^1) < s$ , and  $0 < \text{rank}((B_{21})^2) = m - s < s$ , i.e., even in the "squeezed" form of  $(B_{21})^1$  and  $(B_{21})^2$ , and we can still have zero columns in  $(B_{21})^1$ .

(b)  $\text{rank}(B_{21}) = \text{rank}(B_{12})$ ,  $0 < \text{rank}((B_{21})^1) < s$ , and  $0 < \text{rank}((B_{21})^2) < m - s$  (recall that  $B_{11} \in \mathbb{F}^{m \times m}$  and  $m = \text{rank}(C)$ ), i.e., even in the "squeezed" form of  $(B_{21})^1$  and  $(B_{21})^2$ , and we can always have zero columns in  $(B_{21})^1$  and  $(B_{21})^2$ .

(c)  $\text{rank}(B_{21}) < s$ ;

(d)  $\mathbb{F} = \overline{\mathbb{F}}$ , in which we needn't consider the determinant condition.

In case (a), first, using approach (I) we are able to make  $(B_{21})^1$  and  $(B_{21})^2$  into diagonal-like form and we let  $\alpha_{12} = 0$  in following steps. If any permutation is involved, we can transfer the sign 1 or -1 to the "diagonals" of  $B_{21}$ . Here is what I mean. For example, to commute the first and the second rows/columns, instead of using  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  we can use  $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$  or  $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ . From now on in this proof we only need  $x_{22}, y_{22}, \alpha_{11}$  and  $\alpha_{22}$  be diagonal matrices. In each step of following manipulations we only need to choose special  $x_{22}$  or  $\alpha_{11}$  (and therefore  $y_{22}$ ) or  $\alpha_{22}$ , and we assume other parts of  $\alpha$  to be identities, Therefore we only need to make the specified sub-matrix in  $\alpha$  have the determinant one to satisfy determinant condition. Then the first  $s$  columns and the last  $m - s$  columns of  $x_{22}B_{21}\alpha^{-1}$  are  $x_{22}(B_{21})^1\alpha_{11}^{-1}$  and  $x_{22}(B_{21})^2\alpha_{22}^{-1}$ . We claim that we are able to assume

$$\begin{aligned} B_{21} &= \begin{bmatrix} (B_{21})^1 & (B_{21})^2 \end{bmatrix} \\ (B_{21})^1 &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ (B_{21})^2 &= \begin{bmatrix} 0 \\ I \end{bmatrix}. \end{aligned}$$

Here is how to reduce to the above form. Let the diagonals of  $x_{22}$  be the reciprocals of the corresponding "diagonals" in  $B_{21}$  except the first diagonal, which is the reciprocal of the product of other diagonals in  $x_{22}$  and makes  $\det(x_{22}) = 1$ . Now we have following form  $(B_{21}$

has full row-rank, but always has zero columns)

$$\begin{aligned}
B_{21} &= \begin{bmatrix} (B_{21})^1 & (B_{21})^2 \end{bmatrix} \\
(B_{21})^1 &= \left[ \begin{array}{ccc|c} \det(B_{21}) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ \hline & 0 & & 0 \end{array} \right] \\
(B_{21})^2 &= \begin{bmatrix} 0 \\ I \end{bmatrix},
\end{aligned}$$

(We have similar form/conclusion in the following Theorem 6.1.25 and in that theorem by  $\det(B_{21})$  we mean the determinant of the diagonal matrix formed from  $B_{21}$ , the "squeezed and cut"  $B_{21}$ ). Now we let the first diagonal of  $\alpha_{11}$  be  $\det(B_{21})$ , the last of  $\alpha_{11}$  be  $\frac{1}{\det(B_{21})}$ , and we then have the desired form. Here we use the proper that  $\frac{1}{\det(B_{21})}$  goes to the last row of  $(B_{21})^1$ , which is zero, and we still have zero.

We can do same thing to  $(\overline{B}_{21})^1$  and  $(\overline{B}_{21})^2$ . Since those forms are determined by corresponding ranks only and those ranks for  $B$  and  $\overline{B}$  are same, we have  $B_{21} = \overline{B}_{21}$ ,  $B_{12} = \overline{B}_{12}$ , in other words,  $B = \overline{B}$ . We therefore of course conclude that those four equations have solution since  $g \stackrel{K_{\mathbb{F}}}{\sim} \overline{g}$ .

In case (b), similarly we are able to have following form

$$\begin{aligned}
B_{21} &= \begin{bmatrix} (B_{21})^1 & (B_{21})^2 \end{bmatrix} \\
(B_{21})^1 &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\
(B_{21})^2 &= \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.
\end{aligned}$$

Here is how. It's almost same as in the case (a). We first make  $B_{21}$  into diagonal-like form. Then by choosing special  $\alpha_{11}$  and  $\alpha_{22}$ , we can make  $(B_{21})^1$  and  $(B_{21})^2$  have the above special form. The only difference is that diagonals of  $\alpha_{11}$  and  $\alpha_{22}$  are determined by nonzero "diagonals" of  $B_{21}$  directly. Again, the form of  $(B_{21})^1$  and  $(B_{21})^2$  are determined by the corresponding ranks only.

In case (c), if  $\text{rank}(B_{21}) < s$ , we then always have some zero row in  $B_{21}$ . After making  $B_{21}$  to be diagonal-like, we can use  $\alpha$  to make all nonzero diagonals of  $B_{21}$  to be 1, and use  $x_{22}$  to

make  $(\det(\alpha))^2 \det(x_{22}) \det(y_{22}) = 1$ : the diagonal of  $x_{22}$  corresponding to the zero row of  $B_{21}$  can be any nonzero number in the field  $\mathbb{F}$ .

In case (d), we can always make nonzero diagonals of  $B_{21}$  to be one, since we don't need to care about the determinant in this case. We have proved (IV) in our theorem.

Now we prove (I). If  $B_{12}$  is not square, i.e.,  $B_{12}$  has more rows than columns, then  $\alpha_{22}$  exists in  $\alpha$ . By taking proper  $\alpha$  ( $x_{22}$  can be used to make the determinant condition to be satisfied) we are able to make

$$(B_{21})^2 = \begin{bmatrix} \det((B_{21})^1) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$(\overline{B}_{21})^2 = \begin{bmatrix} \det((\overline{B}_{21})^1) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

and  $(B_{21})^1 = (\overline{B}_{21})^1 = 0$  in this case. We let  $\alpha_{12} = \tilde{\alpha}_{12} = 0$  and use  $\alpha_{22}$  to make determinant condition satisfied.

From equation (6.1.8)

$$x_{21}B_{11} + x_{22}B_{21} = \overline{B}_{21}\alpha$$

or  $(B_{11} = 0)$

$$x_{22}B_{21} = \overline{B}_{21}\alpha$$

we have  $((B_{21})^1 = (\overline{B}_{21})^1 = 0)$

$$x_{22} \cdot (B_{21})^2 = (\overline{B}_{21})^2 \cdot \alpha_{22}$$

$$0 = (B_{21})^2 \alpha_{12}$$

and therefore

$$\alpha_{12} = 0$$

$$\det(x_{22}) \det((B_{21})^2) = \det((\overline{B}_{21})^2) \det(\alpha_{22})$$

$$\det((\overline{B}_{21})^2) / \det((B_{21})^2) = \det(x_{22}) / \det(\alpha_{22})$$

$$= \left( \frac{1}{\det(\alpha_{11}) \det(\alpha_{22})^2} \right)^2 \in (\mathbb{F}^*)^2$$



On the other hand, if  $\det((\overline{B}_{21})^1)/\det((B_{21})^1) \in (\mathbb{F}^*)^2$ , we can simply let

$$\begin{aligned}\alpha_{22} &= I \\ \alpha_{12} &= 0 \\ \alpha_{21} &= 0 \\ x_{22} &= \begin{bmatrix} \det((\overline{B}_{21})^2)/\det((B_{21})^2) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \\ \det(\alpha_{11}) &= \frac{1}{\sqrt{\det(x_{22})}}\end{aligned}$$

to satisfy the equation (6.1.8). Notice that we have no condition on  $\alpha_{22}$  except the determinant condition.

If  $(B_{21})^1$  does not exist, which is part (b) in Case (I), we can see that  $B_{12} = I$  and  $\alpha_{11}$  does not exist in  $\alpha$  either. Repeating proof in Part (I) we can prove the necessity. For sufficiency, let  $\det((\overline{B}_{21})^2)/\det((B_{21})^2) = p^4$  and

$$\alpha_{22} = \begin{bmatrix} p^{-1} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \text{ and } x_{22} = \begin{bmatrix} p^3 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

Then equation (6.1.6)-(6.1.9) and the determinant conditions are all satisfied.

The proof for (II) is similar to the proof of (I). Notice that we always have  $\alpha_{11}$  here since there is no way for  $B_{12}$  or  $B_{21}$  to be square. Also  $(B_{12})^2$  always exists in this case.

First we able to make

$$(B_{21})^1 = \begin{bmatrix} \det((B_{21})^1) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$(\overline{B}_{21})^1 = \begin{bmatrix} \det((\overline{B}_{21})^1) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

and we instead have equation

$$x_{22} \cdot (B_{21})^1 = (\overline{B}_{21})^1 \cdot \alpha_{11}$$

and we get

$$\begin{aligned} \det(x_{22}) \det((B_{21})^1) &= \det((\overline{B}_{21})^1) \det(\alpha_{11}) \\ \det((\overline{B}_{21})^1) / \det((B_{21})^1) &= \det(x_{22}) / \det(\alpha_{11}) \\ &= \left( \frac{1}{\det(\alpha_{11}) \det(\alpha_{22})} \right)^3 \in (\mathbb{F}^*)^3. \end{aligned}$$

On the other hand, if  $\det((\overline{B}_{21})^1) / \det((B_{21})^1) \in (\mathbb{F}^*)^3$ , we simply let

$$\begin{aligned} \alpha_{11} &= I \\ \alpha_{12} &= 0 \\ \alpha_{21} &= 0 \\ x_{22} &= \begin{bmatrix} \det((\overline{B}_{21})^1) / \det((B_{21})^1) & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \\ \det(\alpha_{22}) &= \frac{1}{\sqrt[3]{\det(x_{22})}} \end{aligned}$$

□

Now we assume  $B_{11}$  has nonzero Jordan blocks, i.e.,  $\alpha_{00}$  is at most a proper submatrix of  $\alpha$ .

We try to reduce  $B_{12}, B_{21}$  (and  $\overline{B}_{12}, \overline{B}_{21}$  in a same way) to a special/canonical form, i.e., we are using approach (I): act on  $B_{12}, B_{21}$  by  $\alpha \cdot B_{12} y_{22}^{-1}, x_{22} \cdot B_{21} \alpha^{-1}$ .

Let's first consider rows in  $B_{12}$  corresponding to the last rows of  $J_i, i \geq 1$ , and columns in  $B_{21}$  corresponding to the first columns of  $J_i, i \geq 1$ .

To see what's going on explicitly, we write down, for example, following two equations

$$\begin{aligned}\alpha \cdot B_{12} &= \overline{B}_{12} \cdot y_{22} \\ x_{22} \cdot B_{21} &= \overline{B}_{21} \cdot \alpha.\end{aligned}$$

We can write

$$\alpha \cdot B_{12} = \overline{B}_{12} \cdot y_{22}$$

as

76

and

$$x_{22} \cdot B_{21} = \overline{B}_{21} \cdot \alpha$$

as

$$x_{22} * \left[ \begin{array}{c|c|c|c} c^{00} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$= \left[ \begin{array}{c|c|c|c} \overline{c}^{00} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$\left[ \begin{array}{c|c|c|c} & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * \\ \hline * & * & * & * & * & * & * & * & * & * & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & 0 & * \\ \hline * & * & * & * & * & * & * & * & * & * & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{array} \right]$$

Since we are simplifying problem, we just take  $\alpha, x_{22}, y_{22}$  in a way we want:

We first let all entries in  $\alpha_{ij}$ ,  $i \cdot j \neq 0$ , be zeros except those being on the main diagonals of

each block, i.e.,  $\alpha_{ij} = \omega_{ij}I$ ,  $i \cdot j \neq 0$ ,  $\omega_{ij} \in \mathbb{F}$ ;

$\alpha_{ij} = 0$  if  $i = 0$  or  $j = 0$  but  $i + j \neq 0$ .

Also, keeping in mind that we only consider the possibly nonzero rows in  $B_{12}$  and possibly nonzero columns in  $B_{21}$ , and observing the above multiplication, it's clear that only need to study two cases here:

- (1)  $J_i, 1 \leq i \leq r$  have same size;
- (2)  $J_i, 1 \leq i \leq r$  have different sizes.

In the first case, we have  $(\alpha^0$  is the squeezed  $(\alpha_{ij})_{i,j \geq 1}$  without (i.e., deleting) the first rows and first columns (the rows and the columns where  $\alpha_{00}$  lies in).  $b^{11} = b = (b_{ij}), c^{11} = c = (c_{ij}), \bar{b}^{11} = \bar{b} = (\bar{b}_{ij}), \bar{c}^{11} = \bar{c} = (\bar{c}_{ij})$  are squeezed here too)

$$\left\{ \begin{array}{l} \sum_{1 \leq k \leq r} \alpha_{ik}^0 * b_{kj} = \sum_{1 \leq k \leq r} \bar{b}_{ik} * (y_{22})_{kj} \\ \sum_{1 \leq k \leq r} (x_{22})_{ik} * c_{kj} = \sum_{1 \leq k \leq r} \bar{c}_{ik} * \alpha_{kj}^0 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \alpha^0 \cdot b = \bar{b} \cdot y_{22} \\ x_{22} \cdot c = \bar{c} \alpha^0 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \alpha^0 \cdot b \cdot y_{22}^{-1} = \bar{b} \\ x_{22} \cdot c \cdot (\alpha^0)^{-1} = \bar{c} \end{array} \right. \quad (*)$$

*Remark 6.1.19.* Using Lemma 6.1.4 we can see that  $\alpha^0$  is always invertible.

Therefore, again, we are able to make  $b = \begin{bmatrix} I \\ 0 \end{bmatrix}$  which gives us

$$\alpha^0 = \begin{bmatrix} \alpha_{11}^0 & \alpha_{12}^0 \\ 0 & \alpha_{22}^0 \end{bmatrix} \text{ (to make } \bar{b} = b \text{ keep same form/shape).}$$

*Remark 6.1.20.* The determinant condition goes to  $B_{21}$  again if  $\mathbb{F} \neq \bar{\mathbb{F}}$ .

The classification of unipotent elements is same as in the case  $\alpha = \alpha_{00}$ : Theorem 6.1.18.

In the second case that  $J_i, 1 \leq i \leq r$ , have different sizes/orders. We assume  $\mathbb{F} = \bar{\mathbb{F}}$  first.

We cannot take all elements in  $\alpha^0$  to act on  $B_{12}, B_{21}$  (or  $b, c$  here):

$$\begin{aligned} \text{on } b : \quad \alpha^0 &= \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & & * & * & * & * \\ & & * & * & * & * \\ & & & & * & * \\ & & & & * & * \end{bmatrix} \\ \text{on } c : \quad \alpha^0 &= \begin{bmatrix} * & * & & & & \\ * & * & & & & \\ * & * & * & * & & \\ * & * & * & * & & \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{bmatrix} \end{aligned}$$

From (\*) we are still able to make  $b$  and  $c$  to be identity-like shape, but we need more delicate conditions here.

Abusing notations again, we assume that  $B_{11}$  (and  $\alpha$  correspondingly) has  $m_r$  nonzero Jordan blocks of largest sizes,  $m_{r-1}$  nonzero Jordan blocks of the second largest size plus largest size,  $\dots$ ,  $m_1$  nonzero Jordan blocks, i.e., we have for  $B_{11}$

$$B_{11} =$$



and the way of  $\alpha^0$  acting on  $b$  and  $c$  should have same pattern as the above  $B_{11}$ .

Now we partition  $B_{12}, B_{21}$  (and therefore  $b, c$ ) correspondingly, i.e., the sizes of blocks in  $b, c$  are, in order,  $m_1 - m_2, m_2 - m_3, \dots, m_{r-1} - m_r, m_r$ .

Let's look at actions on  $b$  first:  $\alpha^0 \cdot b \cdot y_{22}^{-1}$ .

We assume  $b = \begin{bmatrix} b^1 \\ \vdots \\ b^r \end{bmatrix}$ . We are still able to make

$$b_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \end{bmatrix}, 1 \leq i \leq r \quad (6.1.11)$$

(first make  $b$  upper-triangular-like using  $y_{22}^{-1}$ , then make  $b$  identity-like using  $\alpha^0$ ). For example,  $b$  could be like

$$\begin{bmatrix} & & & & & \\ & & & & & \\ & \boxed{\diagdown} & & & & \\ & & \boxed{\diagdown} & & & \\ & & & \boxed{\diagdown} & & \\ & & & & \boxed{\diagdown} & \\ & & & & & \boxed{\diagdown} \end{bmatrix} \text{ or } \begin{bmatrix} \boxed{\diagdown} & & & & & \\ & \boxed{\diagdown} & & & & \\ & & \boxed{\diagdown} & & & \\ & & & \boxed{\diagdown} & & \\ & & & & \boxed{\diagdown} & \\ & & & & & \boxed{\diagdown} \end{bmatrix}.$$

To make  $b$  unchanged, we get some new conditions on each column of  $\alpha^0$  from  $\alpha^0 \cdot b = by_{22}$  :

$$\alpha^0 = \begin{bmatrix} * & * & * & * & 0 & 0 \\ * & * & * & * & 0 & 0 \\ & & * & * & 0 & 0 \\ & & * & * & 0 & 0 \\ & & & * & * & \\ & & & * & * & \end{bmatrix} \text{ or } \alpha^0 = \begin{bmatrix} * & * & * & * & * & 0 \\ * & * & * & * & * & 0 \\ & & * & * & * & 0 \\ & & * & * & * & 0 \\ & & & * & 0 & \\ & & & * & * & \end{bmatrix}$$

where the first type of above has no effect on  $\alpha^0$  (on  $c$ ) since, we actually don't use the new

zero part in  $\alpha^0$  acting on  $c$ :

$$\alpha^0 = \begin{bmatrix} * & * & & & & \\ * & * & & & & \\ * & * & * & * & & \\ * & * & * & * & & \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{bmatrix}, \quad (6.1.12)$$

but the second type affects the structure of  $\alpha^0$  acting on  $c$ :

$$\alpha^0 = \begin{bmatrix} * & * & & & & \\ * & * & & & & \\ * & * & * & * & & \\ * & * & * & * & & \\ * & * & * & * & * & 0 \\ * & * & * & * & * & * \end{bmatrix}. \quad (6.1.13)$$

It makes part of  $\alpha_{ii}$  be zero, but it never affect entries under the main diagonal line.

Now we consider action on  $c : x_{22} \cdot c \cdot (\alpha^0)^{-1}$  and  $(\alpha^0)^{-1}$  has same shape as  $\alpha^0$ .

We partition  $c_i$  according to the form (6.1.12) or (6.1.13) of  $\alpha^0$ :

$c^i = c_0^i$  in the case of (6.1.12);

$c^i = [c_1^i, c_2^i]$  in the case of (6.1.13), and sizes of  $c_1^i, c_2^i$  are determined by  $m_{i+1} - m_i$  and  $\text{rank}(b^i)$ .

Now using  $x_{22}, (\alpha^0)^{-1}$  we are able to make  $c_j^i, j = 0, 1, 2$  look like

$$\begin{bmatrix} 0 & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix} \quad (**)$$

This is because we are able to make  $c$  to be lower-triangular-like first, then we are able to use  $(\alpha^0)^{-1}$  to "clear up" extra nonzero elements in each row (we can do this because we are using elements in  $\alpha$  under the main diagonal line only), then within each block, we are able to re-arrange  $c_0^i, c_1^i, c_2^i$  into the above form.

What's more: the canonical form in (\*\*) is determined by  $m_i, \text{rank}(b^i), \text{rank}(c^i)$  or  $\text{rank}(c_1^i), \text{rank}(c_2^i)$ ,  $1 \leq i \leq r$ .

To determine  $b$  and  $c$  we should always "put nonzero elements into their positions start from right bottom corner".

Now let's state our first part of classification of  $K_{\mathbb{F}}$ -orbits for the key involution:

**Theorem 6.1.21.** *If  $\mathbb{F} = \overline{\mathbb{F}}$ , equations (6.1.6)-(6.1.9) with  $B$  and  $\overline{B}$  being special as pre-described have solutions if and only if following conditions hold for  $1 \leq i \leq r$ :*

$$\text{rank}(B_{12}) = \text{rank}(\overline{B}_{12}) \quad (6.1.14)$$

$$\text{rank}(B_{21}) = \text{rank}(\overline{B}_{21}) \quad (6.1.15)$$

$$\text{rank}(b^{00}) = \text{rank}(\overline{b}^{00}) \quad (6.1.16)$$

$$\text{rank}(b^i) = \text{rank}(\overline{b}^i) \quad (6.1.17)$$

$$\text{rank}(c_j^i) = \text{rank}(\overline{c}_j^i), j = 0, 1, 2, \quad (6.1.18)$$

$$\text{rank}(c^{00}) = \text{rank}(\overline{c}^{00}) \quad (6.1.19)$$

*Proof.* In this case,  $\mathbb{F} = \overline{\mathbb{F}}$ , we don't need to consider the determinant condition here.

$\Leftarrow$ : We have reduced everything into canonical forms, which are determined by corresponding ranks only. Here we simply take  $\alpha = I, x = I, y = I$  to make equations (6.1.6)-(6.1.9) satisfied.

$\Rightarrow$ :  $\text{rank}(B_{12}) = \text{rank}(\overline{B}_{12})$  and  $\text{rank}(B_{21}) = \text{rank}(\overline{B}_{21})$  are obvious.

Next we will show that  $\text{rank}(b^i) = \text{rank}(\overline{b}^i)$ .

Starting from  $r$ , we take the last  $m_i$  ( $1 \leq i \leq r$ ) rows of  $\alpha^0 \cdot b = \overline{b} \cdot y_{22}$ :

$$(b_{m_i} = \begin{bmatrix} b^i \\ \vdots \\ b^r \end{bmatrix}, \overline{b}_{m_i} = \begin{bmatrix} \overline{b}^i \\ \vdots \\ \overline{b}^r \end{bmatrix})$$

$$\left[ \begin{array}{c|c|c} \begin{array}{c} \diagdown \\ \diagup \end{array} & & * \\ \hline 0 & \begin{array}{c} \diagdown \\ \diagup \end{array} & \\ \hline \end{array} & \begin{array}{c} * \\ \\ \end{array} & \left. \begin{array}{c|c} \begin{array}{c} \diagdown \\ \diagup \end{array} & * \\ \hline 0 & \begin{array}{c} \diagdown \\ \diagup \end{array} \end{array} \right\} m_i \right] \cdot b = \overline{b} \cdot y_{22} \Rightarrow$$

$$\begin{aligned}
& \left[ \begin{array}{c} 0, \\ \left[ \begin{array}{|c|c|} \hline \text{diag} & * \\ \hline 0 & \text{diag} \\ \hline \end{array} \right] \end{array} \right] \cdot b = \bar{b}^{m_i} \cdot y_{22} \\
& \left[ \begin{array}{c} \left[ \begin{array}{|c|c|} \hline \text{diag} & * \\ \hline 0 & \text{diag} \\ \hline \end{array} \right] \end{array} \right] \cdot b_{m_i} = \bar{b}_{m_i} \cdot y_{22} \\
& \alpha_{m_i}^0 \cdot b_{m_i} = \bar{b}_{m_i} \cdot y_{22}
\end{aligned}$$

$\alpha_{m_i}^0$  and  $y_{22}$  are both nonsingular  $\Rightarrow$ :

$$\begin{aligned}
\text{rank}(b_{m_i}) &= \text{rank}(\bar{b}_{m_i}) \\
\text{rank}(b_{m_i}) - \text{rank}(b_{m_{i+1}}) &= \text{rank}(\bar{b}_{m_i}) - \text{rank}(\bar{b}_{m_{i+1}}) \\
\text{rank}(b^i) &= \text{rank}(\bar{b}^i)
\end{aligned}$$

The proof for  $\text{rank}(c_j^i) = \text{rank}(\bar{c}_j^i)$  is similar, noticing that we have to partition  $\alpha_{m_i}^0$  according to form (6.1.12) or (6.1.13) again, if  $j \geq 1$ .

(6.1.14) and (6.1.17)  $\Rightarrow$  (6.1.16).

(6.1.15) and (6.1.18)  $\Rightarrow$  (6.1.19). □

*Remark 6.1.22.* If  $\mathbb{F} = \overline{\mathbb{F}}$ , from the above theorem we can see that the most important case is that  $B_{11}$  has no zero blocks, and sizes of Jordan blocks can be different. In the worst case that sizes of Jordan blocks are all different from each other, we actually require  $B_{12}$  and  $B_{21}$  have exactly same shape.

*Remark 6.1.23.* If  $\mathbb{F} = \overline{\mathbb{F}}$ , we can see that actually we can find corresponding part of Theorem 6.1.21 in partitions of natural numbers. The sizes of Jordan blocks correspond to summands of integer  $m$ : (6.1.20)

$$m = \beta_0 n_0 + \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_r n_r, \quad (6.1.20)$$

It's obvious that  $n_i, i = 0, 1, \dots, s$ , correspond to the sizes of Jordan blocks of  $B_{11}$ . We also have

$$\begin{aligned}
m_r &= \beta_r \\
m_{r-1} &= \beta_{r-1} + \beta_r \\
&\dots \\
m_i &= \sum_{i \leq k \leq r} \beta_k \\
&\dots \\
m_1 &= \sum_{1 \leq k \leq r} \beta_k,
\end{aligned}$$

i.e.,  $\beta_i, i = 1, \dots, r$ , correspond to partitions of  $b$  and  $c$  (or of  $b^{11}$  and  $c^{11}$ , parts in  $B_{12}$  and  $B_{21}$  corresponding to the nonzero Jordan blocks in  $B_{11}$ ). We therefore have  $\text{rank}(b^i) \leq \beta_i$ . If  $\text{rank}(b^i) = 0$  or  $\beta_i$ , then  $c^i = c_0^i$ , i.e., no further partition is needed. If  $0 < \text{rank}(b^i) < \beta_i$ , we need to partition  $c^i$  into  $[c_1^i, c_2^i]$  with  $c_2^i$  has  $\text{rank}(b^i)$  columns, and consider all subcases  $0 \leq \text{rank}(c_1^i) \leq \beta_i - \text{rank}(b^i)$  combining with subcases  $0 \leq \text{rank}(c_2^i) \leq \text{rank}(b^i)$ . Now if we want to give all representatives of canonical forms, we need to determine  $m = \text{rank}(C)$  first, then partition  $m$  as in (6.1.20), give all possible  $\text{rank}(b^i)$  and correspondingly  $\text{rank}(c_j^i), j = 0$  or  $1$  or  $2$ . For  $b^{00}$ , which is corresponding to the zero Jordan blocks and  $\beta_0$  in partition (6.1.20), we have  $\text{rank}(b^{00}) \leq \beta_0$  and we add this part if we consider all cases with  $B$  in the form in part (c) in Theorem 6.1.16. Following is obvious

$$s = \sum_{0 \leq i \leq r} \text{rank}(b^i) \leq \sum_{0 \leq k \leq r} \beta_k$$

since  $B_{12}$  has full column rank.

To be explicit, we can express in the following form:

$$(\text{rank}(C), (n_i, \beta_i), \text{rank}(b^i), \text{rank}(c_j^i), \text{rank}(b^{00}), \text{rank}(c^{00}), \text{rank}(B_{21}^{22})), \quad (6.1.21)$$

where  $(n_i, \beta_i)$  comes from partition (6.1.20) of  $m = \text{rank}(C)$ ,  $\beta_i, i \geq 1$ , give upper bounds of  $\text{rank}(b^i)$ , and we need further partition  $\text{rank}(c_j^i)$  if  $\text{rank}(b^i) \neq 0$  or  $\beta_i$ ;  $\text{rank}(b^{00}), \text{rank}(c^{00})$  (upper bounded by  $\beta_0, \text{rank}(B_{21}^{22})$ ) only give up about the "corners" of  $B$ .

If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , we are still able to make  $B_{12}$  or  $b$  to have the above special form, since we can move the determinant condition to  $B_{21}$  or  $c$ . Conclusion and proof will be similar to cases I and II in Theorem 6.1.18. In case I or II of Theorem 6.1.18, we can see that we can actually

take

$$B_{21} = \left[ \begin{array}{ccc|c} p & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{array} \right] 0 \text{ with } p \in (\mathbb{F}^*)^2$$

or

$$B_{21} = \left[ \begin{array}{ccc} p & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{array} \right] \text{ with } p \in (\mathbb{F}^*)^4$$

or

$$B_{21} = \left[ \begin{array}{c|ccc} & p & & \\ 0 & & 1 & \\ & & & \ddots \\ & & & & 1 \end{array} \right] \text{ with } p \in (\mathbb{F}^*)^3$$

as representatives in  $K_{\mathbb{F}}$ -orbits. They correspond to the case of  $\text{rank}((B_{21})^1) = \text{rank}(\overline{B}_{12})$  or  $\text{rank}((B_{21})^2) = \text{rank}(\overline{B}_{12})$  respectively. In either of the case, we have  $\text{rank}(B_{21}) = \text{rank}(B_{12})$ , and from the proof of Theorem 6.1.18 we can see that, if  $\text{rank}(B_{21}) \neq \text{rank}(B_{12})$ , we can get rid of the determinant condition by making all diagonals of  $B_{21}$  ( $\overline{B}_{12}$ ) to be 1 or 0 and within the same block (in the case of Theorem 6.1.18, we have only one block:  $\alpha = \alpha_{00}$ ), i.e., "det( $B_{21}$ )" or "det( $\overline{B}_{21}$ )" being always 1 or 0. The version of Theorem 6.1.21 for the general field, especially for the case  $\mathbb{F} \neq \overline{\mathbb{F}}$  comes simply from combining Theorem 6.1.18 and Theorem 6.1.21. One thing we need to be careful with is that, in Theorem 6.1.18 we always try to align nonzero diagonal elements of  $b^{00}$  starting from the left top corner, but for  $b^{11}$  we try to stack onto the right bottom corner. Therefore we need to be careful when we pick up submatrices from diagonal blocks of  $\alpha$ .

If we want the determinant condition to make a difference, i.e., the determinant condition cannot be dismissed by being transferred to some zero row/column, we must have  $\mathbb{F} \neq \overline{\mathbb{F}}$  and  $B_{12}, B_{21}, B_{22}$  must satisfy (recall that  $m = \text{rank}(C)$  is the size of  $B_{11}$ :  $B_{11} \in \mathbb{F}^{m \times m}$ )

$$\text{rank}(B_{12}^1) + \text{rank}(B_{22}^{22}) = \frac{n}{2} - m \quad (6.1.22)$$

$$\text{rank}(B_{21}^1) + \text{rank}(B_{22}^{22}) = \frac{n}{2} - m \quad (6.1.23)$$

$$\text{rank}(b^{00}) = \text{rank}(c^{00}) = \text{rank}(\alpha_{00}) = n_0 \quad (6.1.24)$$

and one of the following two:  $(b = b^{11} = \begin{bmatrix} b^1 \\ \vdots \\ b^r \end{bmatrix} \text{ in the squeezed } B_{21}: \left[ \begin{array}{c|c} b^{00} & 0 \\ \hline 0 & b^{11} \end{array} \right])$  for  $i = 1, \dots, s$ , either

$$\text{rank}(b^i) = \text{rank}(c_j^i), j = 0 \text{ or } 2, \text{ and } c_1^i = 0, \quad (6.1.25)$$

or

$$\text{rank}(b^i) = \text{rank}(c_1^i), \text{ and } c_2^i = 0. \quad (6.1.26)$$

(6.1.24)-(6.1.25) or (6.1.24)-(6.1.26) must be satisfied (otherwise all "diagonals" of  $B_{21}$  and  $\overline{B}_{21}$  can be made into 1 or 0). In this case, what matters is only the "shape" of  $B_{21}$  and  $\overline{B}_{21}$  and therefore only corresponding ranks can make difference, which has been included in Theorem 6.1.21.

Also, we can see that, the condition (6.1.25) corresponds to the case I in Theorem 6.1.18 and the condition (6.1.26) corresponds to the case II in Theorem 6.1.18. They are not exactly same since we start aligning nonzero "diagonals" from the right bottom corner here.

From Lemma 6.1.4, we see that  $\det(\alpha^0) \neq 0$ , but we need a more precise result here, which is stated in following lemma:

**Lemma 6.1.24.**  $\det(\alpha) = \det(\tilde{\alpha})$ , where  $\tilde{\alpha} = [\tilde{\alpha}_{ij}]_{i \geq 0, j \geq 0}$  is formed from  $\alpha$ :

1. The size of  $\tilde{\alpha}_{ij}$  is same as the size of  $\alpha_{ij}$  for  $i, j = 0, 1, \dots, r$ ,
2.  $\tilde{\alpha}_{00} = \alpha_{00}$ ,
3.  $\tilde{\alpha}_{ij} = 0$  if  $i \cdot j = 0$  but  $i + j \neq 0$ , or  $\alpha_{ij}$  is not square,
4.  $\tilde{\alpha}_{ij} = \omega_{ij} I$  if  $\alpha_{ij}$  is square, where  $\omega_{ij}$  is the diagonal element of  $\alpha_{ij}$  (remember that all elements on the same (off-)diagonal are same).

*Proof.* We can assume  $\alpha$  in a standard form like in (6.1.10):  $\alpha = [\alpha_{ij}]_{i \geq 0, j \geq 0}$ . Also, we know that  $\alpha^0$  is the "squeezed"  $[\alpha_{ij}]_{i \geq 1, j \geq 1}$  and formed from the diagonals of each submatrices  $\alpha_{ij}$  (in each submatrices, elements in the same diagonal/off-diagonal are same). We first proved the

lemma for the case that all nonzero Jordan blocks have same size, i.e., in Remark 6.1.23, we have

$$m = \beta_0 n_0 + \beta_1 n_1,$$

which tells us  $\alpha^0 \in \mathbb{F}^{\beta_1 \times \beta_1}$ .

We are able to choose proper  $P = [p_{ij}I]_{i \geq 0, j \geq 0}$  with  $p_{00} = 1$ , and  $p_{ij} = 0$  if  $i \cdot j = 0$  but  $i + j \neq 0$ , such that the  $\alpha^1 = P^0 \cdot \alpha^0$  is a diagonal matrix, where  $P^0 = [p_{ij}]_{i \geq 0, j \geq 0}$ . From Lemma 6.1.4, we can see that all diagonals of  $\alpha^1$  are nonzero. By adding the columns of  $\alpha$ , in which the first diagonals at left top corner of each submatrix with  $i \cdot j \geq 1$  lie, to columns of  $\alpha$ , which belong to submatrices of  $\alpha$  with  $j = 0$  but  $i \geq 1$ , we can make those corresponding columns in  $P\alpha$  to be zero and these two steps do not change either the determinant of  $\alpha$  or the elements of  $\alpha_{00}$ . This is because

1.  $\det(P^0) = 1 \Rightarrow \det(P) = 1 \Rightarrow \det(P\alpha) = \det(\alpha)$ ;
2. those first diagonals are the only nonzero elements in those columns.

In the same way are able to make submatrices of  $P\alpha$  with  $i = 0$  and  $j \geq 1$  to be zero, without changing determinant of  $\alpha$  and elements of  $\alpha_{00}$ .

Now we'll use basic definition of determinant to reach our conclusion. In each diagonal submatrix of  $\alpha$  with  $i \cdot j \geq 1$ , the first (last) diagonal at left top (right bottom) corner of the submatrix with  $i \cdot j \geq 1$ , must be taken to get the determinant of  $\alpha$  since it's the only nonzero element in that column(row) of  $\alpha$ . After we cross out columns and rows of those first and last diagonals of submatrices with  $i \cdot j \geq 1$ , the remaining part of those submatrices have same shape and we have to again take the first and last diagonal of remaining part of those submatrices for the same reason. By repeating the above process we can see that actually elements appear in the expression of determinant of  $\alpha$  come from either  $\alpha_{00}$  or from diagonals of diagonal submatrices of  $P\alpha$  with  $i \cdot j \geq 1$  (or equivalently from diagonals of submatrices of  $\alpha$  with  $i \cdot j \geq 1$ ), which is what we want to prove. It's easy to check that in this case we have

$$\det(\alpha) = \det(\alpha_{00}) \cdot (\det(\alpha^0))^{n_1}.$$

If the sizes of nonzero Jordan blocks of  $B_{11}$  are different, i.e., in Remark 6.1.23 we have (6.1.20):

$$m = \beta_0 n_0 + \beta_1 n_1 + \beta_2 n_2 + \cdots + \beta_r n_r,$$

we know that the actually working part of  $\alpha$  is either upper triangular (acting on  $B_{12}^1$  or  $b$ )



or lower triangular (acting on  $B_{21}^1$  or  $c$ ). The only modification we need to make here is that the "elimination" of  $\alpha_{ij}$  with  $i \cdot j = 0$  but  $i + j \neq 0$  should be done in order. For example, in the case of acting on  $b$  which implies  $\alpha$  is upper triangular, when clearing the 0-th column except  $\alpha_{00}$ , we need to start from right bottom corner of diagonal submatrices, but when clearing the 0-th row except  $\alpha_{00}$ , we should start from the left top corner of diagonal submatrices. In the case of acting on  $c$  which implies  $\alpha$  is lower triangular, we need to reverse the above order.

Again, assuming  $\alpha^0 = [(\alpha_{ij}^0)]_{i,j \geq 1}$  being an upper/lower triangular block matrix, following is true:

$$\det(\alpha) = \det(\alpha_{00}) \prod_{1 \leq i \leq r} (\det(\alpha_{ii}^0))^{n_i}.$$

□

Suppose  $B_{11}$  has Jordan blocks as in partition (6.1.20), i.e., we have  $\beta_i$  Jordan blocks of size  $n_i$ . Notice that  $n_0 = 1$ , and if  $B_{11}$  does not have zero Jordan blocks, we don't have  $n_0$  in the partition (6.1.20). We then can get three sequences of integers for  $i = 0, \dots, r$ :

S1.  $(2 + 2n_i)\delta(\text{rank}(b_i)), 2n_i\delta(\text{rank}(\beta_i - \text{rank}(b_i)))$ .

S2.  $(1 + 2n_i)\delta(\text{rank}(b_i)), (1 + 2n_i)\delta(\text{rank}(\beta_i - \text{rank}(b_i)))$ .

S3.  $(2 + 2n_i)\delta(\text{rank}(b_i)), 2n_i\delta(\text{rank}(\beta_i - \text{rank}(b_i)))$  or  $(1 + 2n_i)\delta(\text{rank}(b_i)), (1 + 2n_i)\delta(\text{rank}(\beta_i - \text{rank}(b_i)))$ .

The meaning of the third sequence means that both terms in 1 and 2 could appear in the third sequence.

Now we can have our second part of classification of  $K_{\mathbb{F}}$ -orbits for the key involution:

**Theorem 6.1.25.** (I) If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , all conditions (6.1.22)-(6.1.24) and (6.1.25) are satisfied, if  $B_{12}$  and  $B_{21}$  are not square matrices, then all  $K_{\mathbb{F}}$ -orbits can choose following as representatives

$$c_0^i = \left[ \begin{array}{c|c} 0 & \\ \hline p & \\ & 1 \\ & \ddots \\ & 1 \\ \hline & 0 \end{array} \right] \text{ or } c_2^i = \left[ \begin{array}{c|c} 0 & 0 \\ \hline 0 & p \\ & 1 \\ & \ddots \\ & 1 \\ \hline 0 & 0 \end{array} \right], c_1^i = 0 \text{ or does not exist} \quad (6.1.27)$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence S1.

(II) If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , all conditions (6.1.22)-(6.1.24) and (6.1.26) are satisfied, if  $B_{12}$  and  $B_{21}$  are square matrices, then all  $K_{\mathbb{F}}$ -orbits can choose following as representatives

$$c_1^i = \left[ \begin{array}{c|c} 0 & 0 \\ \hline p & \\ & 1 \\ & \\ & \ddots \\ & 1 \\ \hline 0 & 0 \end{array} \right], c_2^i = 0 \quad (6.1.28)$$

where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence  $S2$ .

(III) If  $\mathbb{F} \neq \overline{\mathbb{F}}$ , all conditions (6.1.22)-(6.1.24), (6.1.25) and (6.1.26) are satisfied (there are some blocks in  $b_i$  satisfying (6.1.25) and some satisfying (6.1.26)), then all  $K_{\mathbb{F}}$ -orbits can choose as representatives from formula (6.1.27) and (6.1.28), where  $p \in \mathbb{F}^*/(\mathbb{F}^*)^h$  for  $i = 1$  and  $p = 1$  for  $i \geq 2$ ,  $h$  is the greatest common factor of sequence  $S3$ .

(IV) In all cases except the above three, all conditions in Theorem 6.1.21 have been good enough to serve as necessary and sufficient conditions.

*Proof.* It's clear that, to consider cases not included in Theorem 6.1.21, conditions (6.1.22)-(6.1.24), and one of (6.1.25) and (6.1.26) must be true. The possible exceptional cases here are almost same as discussion in Theorem 6.1.18 too. For each  $i = 1, \dots, s$ ,  $b^i$  and  $c^i$  are just like  $B_{12}$  and  $B_{21}$  in Theorem 6.1.18, so we only need to discuss cases similar to cases I and II in Theorem 6.1.18 only. We need to add discussions for  $\alpha^0$  and  $b^{00}, c^{00}$  correspondingly. Also, instead of having one block in Theorem 6.1.18, we have several blocks here and need to discuss all possible combinations. Case I in this theorem is the case that all blocks  $b^i$  and  $c^i$ ,  $i = 1, \dots, s$  have case I in Theorem 6.1.18. Instead of having  $\alpha_{11}$  and  $\alpha_{22}$  in the determinant condition, we have  $\alpha_{ii}^{11}$  and  $\alpha_{ii}^{22}$  in the determinant condition: first we can make  $B_{12}$  and  $B_{21}$  into diagonal-like form, make "diagonals" of  $b^{00}$ ,  $b^{11}$  (or  $b$  in notations in Theorem 6.1.21) and  $c^{00}$  to be one; then we are able to make all "diagonals" of  $c$  (or  $c^{11}$  in  $B_{21}$ ) to be one except the first diagonal. To keep  $b$  in the same form as in (6.1.11), from

$$\alpha \cdot B_{12} = B_{12}y_{22}$$

we have  $(y_{22} = ((y_{22})^{ij})_{i \geq 0, j \geq 0}, \alpha = [\alpha_{ij}]_{i, j \geq 0}$ , and  $\alpha^0 = (\alpha_{ij}^0)_{i \geq 1, j \geq 1}$  is squeezed  $[\alpha_{ij}]_{i, j \geq 1}$ : formed from diagonals of  $[\alpha_{ij}]_{i, j \geq 1}$ )

$$\begin{aligned}
(y_{22})^{ij} &= 0 \text{ if } i > j \geq 0, \\
\alpha_{00} &= \begin{cases} \alpha_{00}^{11}, & \text{if } \text{rank}(b^{00}) = 0, \\ \begin{bmatrix} \alpha_{00}^{11} & 0 \\ \alpha_{00}^{12} & \alpha_{00}^{22} \end{bmatrix}, & \text{if } 0 < \text{rank}(b^{00}) < \beta_0, \\ \alpha_{00}^{22}, & \text{if } \text{rank}(b^{00}) = \beta_0, \end{cases} \\
\alpha_{00}^{22} &= (y_{22})^{00} \text{ if } \text{rank}(b^{00}) \neq 0; \\
\alpha_{ii}^0 &= \begin{cases} (\alpha_{ii}^0)^{11}, & \text{if } \text{rank}(b^i) = 0 \\ \begin{bmatrix} (\alpha_{ii}^0)^{11} & 0 \\ (\alpha_{ii}^0)^{21} & (\alpha_{ii}^0)^{22} \end{bmatrix}, & \text{if } 0 < \text{rank}(b^i) < \beta_i = m_{i-1} - m_i, i \geq 1, \\ (\alpha_{ii}^0)^{22}, & \text{if } \text{rank}(b^i) = \beta_i = m_{i-1} - m_i, i \geq 1, \end{cases} \\
(\alpha_{ii}^0)^{22} &= (y_{22})^{ii}, i \geq 1, \text{ if } \text{rank}(b^i) \neq 0.
\end{aligned}$$

If  $\text{rank}(b^{00}) = 0$ , we have  $y_{22} = ((y_{22})^{ij})_{i \geq 1, j \geq 1}$ , i.e., we don't have the 0-th row and column in  $y_{22}$ . If  $\text{rank}(b_i) = 0$ , we don't have the  $i$ -column in  $y_{22}$ .

Now we assume  $0 < \text{rank}(b^{00}) < \beta_0$ , so  $\alpha_{00}$  and  $y_{22}^{00}$  both exists. We also assume first  $0 < \text{rank}(b_i) < \beta_i, i = 1, \dots, r$ . We then have

$$\begin{aligned}
1 &= \det(x_{22}) \det(y_{22}) (\det(\alpha))^2 \\
&= \det(x_{22}) \prod_{0 \leq i \leq r} \det((y_{22})^{ii}) (\det(\alpha_{00}))^2 \left( \prod_{1 \leq i \leq r} (\det((\alpha_{ii}^0)^{11}) \cdot \det((\alpha_{ii}^0)^{22}))^{n_i} \right)^2 \\
&= \det(x_{22}) \det(\alpha_{00}^{22}) \prod_{1 \leq i \leq r} \det((\alpha_{ii}^0)^{22}) (\det(\alpha_{00}^{11}))^2 (\det(\alpha_{00}^{22}))^2 \left( \prod_{1 \leq i \leq r} \det((\alpha_{ii}^0)^{11})^{2n_i} \det((\alpha_{ii}^0)^{22})^{2n_i} \right) \\
&= \det(x_{22}) \det(\alpha_{00}^{11})^2 \det(\alpha_{00}^{22})^{1+2} \left( \prod_{1 \leq i \leq r} \det((\alpha_{ii}^0)^{22})^{1+2n_i} \right) \cdot \left( \prod_{1 \leq i \leq r} \det((\alpha_{ii}^0)^{11})^{2n_i} \right). \\
&= \det(x_{22}) \left( \prod_{0 \leq i \leq r} \det((\alpha_{ii}^0)^{22})^{1+2n_i} \right) \cdot \left( \prod_{0 \leq i \leq r} \det((\alpha_{ii}^0)^{11})^{2n_i} \right).
\end{aligned}$$

where, by abusing notations, we assume  $\det(\alpha_{00}^{11}) = \det((\alpha_{00}^0)^{11}), \det(\alpha_{00}^{22}) = \det((\alpha_{00}^0)^{22})$ .

The most general form for the above condition is

$$1 = \det(x_{22}) \left( \prod_{0 \leq i \leq r} \det((\alpha_{ii}^0)^{22})^{(1+2n_i)\delta(\text{rank}(b_i))} \right) \cdot \left( \prod_{0 \leq i \leq r} \det((\alpha_{ii}^0)^{11})^{2n_i\delta(\beta_i - \text{rank}(b_i))} \right). \quad (6.1.29)$$

and we assume  $b_{00} = b_0$  here. The function  $\delta$ -function is to eliminate terms that might not appear in the above determinant condition. Notice that  $(1 + 2n_i)\delta(\text{rank}(b_i)) + \delta(\text{rank}(b_i))$  and

$2n_i\delta(\beta_i - \text{rank}(b_i))$  actually form our  $S1$  sequence.

They all have same pattern as that in case I in Theorem 6.1.18. Same as before, we can first make  $B_{21}$  ( $B_{21}^1$ ) into a diagonal-like form, we can use  $x_{22}$  to make all "diagonals" to be one except the first diagonal in the first diagonal block, which is "determinant of the squeezed and neatly cut  $B_{21}$ ". We then can deploy same analysis in Theorem 6.1.18 to get the conclusion. We take "determinants" of both sides of following equation (notice that only (sub)blocks of  $\alpha$  corresponding to  $x_{22}$  can appear in the determinant formula)

$$x_{22}B_{21} = \overline{B}_{21}\alpha \quad (6.1.30)$$

to show that  $\det(\overline{B}_{21})/\det(B_{21}) \in (\mathbb{F}^*)^h$ , which proves the necessity.

To prove the sufficiency, same as in the proof of Theorem 6.1.18, we can equivalently consider determinant condition and following equations, which correspond to the condition (6.1.30):

$$\begin{aligned} \omega_0 \det(B_{21}) &= \det(\overline{B}_{21})\gamma_0 \quad (\Leftrightarrow \omega_0 = \frac{\det(\overline{B}_{21})}{\det(B_{21})}\gamma_0 = p^h\gamma_0) \\ \omega_i &= \gamma_i, i = 1, \dots, r, \end{aligned}$$

where  $\det(x_{22}) = \prod_{0 \leq i \leq r} \omega_i$  and  $\gamma_0 = \det(\alpha_{00}^{22})$  (if it exists),  $\gamma_i = \det((\alpha_{ii}^0)^{22})$ .

Notice that  $B_{21}$  and  $\overline{B}_{21}$  have been in special form, and all "diagonal" are all 1 except the first "diagonal". Also  $\omega_i$  (or  $\gamma_i$ ) are to be determined.

If we plug  $\gamma_i$  (and  $p^h = \det(\overline{B}_{21})/\det(B_{21})$ ) into the determinant condition, we can see that the determinant condition is satisfied if and only if  $\det(\overline{B}_{21})$  and  $\det(B_{21})$  are  $S1$ -equivalent and using Lemma 5.1.3 if and only if  $\det(\overline{B}_{21})/\det(B_{21}) \in (\mathbb{F}^*)^h$ .

The proof for the general case of Case I is almost same. The proof of Case II and Case III is same as the proof above too.

□

*Remark 6.1.26.* 1. The equation (6.1.29) is a key formula for our derivation.

2. When we take the determinant of both sides of (6.1.30), only diagonal (sub)blocks of  $\alpha$  corresponding to  $x_{22}$  can be taken, other diagonal (sub)blocks simply disappear.

3. In proof of sufficiency,  $B_{21}$  and  $\overline{B}_{21}$  have been in their special forms, i.e., diagonal-like forms, which makes our proof relatively straightforward.

*Remark 6.1.27.* 1. The case of  $B_{11} = 0$  is included in this formula, i.e., if we take  $n_0 = 1$  and  $\beta_0 = m$ , we would get Theorem 6.1.18.

2. If all sizes of Jordan blocks in  $B_{11}$  have same size, then classification of  $K_{\mathbb{F}}$ -orbits of unipotent elements would be same as Theorem 6.1.18.

*Remark 6.1.28.* For the key involution, we actually reduce the classification problem to a Jordan-canonical-form-like problem: we are trying to reduce a matrix pair/pencil simultaneously to some special/canonical form:  $B \rightarrow x^{-1}By, C \rightarrow y^{-1}Cx$ . It can almost be reduced to the usual JCM if  $C$  (especially when we can make  $C = I$ ) is nonsingular, but it's completely different if  $C$  is singular.

### 6.1.3 Examples

Here are some examples for the representative elements/canonical forms we can take from  $K_{\mathbb{F}}$ -orbits of unipotent elements in  $P = G_{\mathbb{F}}/K_{\mathbb{F}}$  for  $G_{\mathbb{F}} = \text{SL}(n, \mathbb{F})$ . To make it typical, we have to make  $n$  to be large enough, and at same time this would make us hard to write out all representatives for that specific  $n$ . Therefore in examples here we will fix both  $n$  and  $\text{rank}(C) = m$  and assume  $\text{rank}(B_{12}) \geq \text{rank}(B_{21})$  (as we have done in our analysis and proofs) and other condition(s) to keep our examples short.

*Example 6.1.29* (Case  $B_{11} = 0$ ). Let  $n = 18, m = \text{rank}(C) = 6$ , so

$$B_{11} = 0, B_{12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, B_{21} \in \mathbb{F}^{3 \times 6}.$$

Then all  $B_{21}$  we can take here are

- $\text{rank}(B_{21}) = 0$ :  $B_{21} = 0$ .

- $\text{rank}(B_{21}) = 1$ :

$$- \text{rank}((B_{21})^2) = 0, B_{21} = \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array} \right]$$

$$- \text{rank}((B_{21})^2) = 1, B_{21} = \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

- $\text{rank}(B_{21}) = 2$ :

$$\begin{aligned}
- \text{rank}((B_{21})^2) = 0, B_{21} &= \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array} \right] \\
- \text{rank}((B_{21})^2) = 1, B_{21} &= \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \\
- \text{rank}((B_{21})^2) = 2, B_{21} &= \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right]
\end{aligned}$$

•  $\text{rank}(B_{21}) = 3$ :

$$\begin{aligned}
- \text{rank}((B_{21})^2) = 1, B_{21} &= \left[ \begin{array}{ccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \\
- \text{rank}((B_{21})^2) = 2, B_{21} &= \left[ \begin{array}{ccc|ccc} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \\
- \text{rank}((B_{21})^2) = 0, B_{21} &= \left[ \begin{array}{ccc|ccc} p & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array} \right], p \in \mathbb{F}^*/(\mathbb{F}^*)^3
\end{aligned}$$

The nonzero part of  $B_{21}$  does not lie in the same corner as  $B_{12}$ , which corresponds to case (II).

$$- \text{rank}((B_{21})^2) = 3, B_{21} = \left[ \begin{array}{ccc|ccc} 0 & 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right], p \in \mathbb{F}^*/(\mathbb{F}^*)^2$$

The nonzero part of  $B_{21}$  lies in the same corner as  $B_{12}$ , which corresponds to case (I).

In the following example, we work on a relatively general case though it actually does not have zero Jordan blocks in  $B_{11}$ , which makes it a little simpler. Under conditions we mentioned at the beginning of this subsection, we list all possible (representative elements of)  $K_{\mathbb{F}}$ -orbits for a given JCM of  $B_{11}$  only.

*Example 6.1.30* (JCM of  $B_{11}$  is given and has no zero Jordan blocks). Let  $n = 20, m = \text{rank}(C) = 8$  and  $8 = 2 \cdot 2 + 1 \cdot 4 = \beta_1 n_1 + \beta_2 \cdot n_2$ , i.e.,  $n_1 = 2, \beta_1 = 2, n_2 = 4, \beta_2 = 1$ , so  $B_{11}$  has two Jordan blocks of size 2 and one Jordan block of size 4:  $(B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix})$  with

$B_{11} \in \mathbb{F}^{8 \times 8}$  and  $B_{22} \in \mathbb{F}^{2 \times 2}$ , only consider the case  $\text{rank}(B_{12}) \geq \text{rank}(B_{21})$

$$B_{11} = \begin{bmatrix} J_1 & & \\ & J_1 & \\ & & J_2 \end{bmatrix}, J_1 = \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}, J_2 = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

We have two sizes of Jordan blocks:

$$r = 2 \Rightarrow b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, c = \begin{bmatrix} c_1 & c_2 \end{bmatrix}.$$

- $\text{rank}(B_{22}) = 2$ :  $B_{22} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

- $\text{rank}(B_{12}) = \text{rank}(B_{21}) = 0$ :  $B_{12} = 0, B_{21} = 0$ .

- $\text{rank}(B_{22}) = 1$ :  $B_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

- $\text{rank}(B_{12}) = 0, \text{rank}(B_{21}) = 0$ :  $B_{12} = 0, B_{21} = 0$ .

- $\text{rank}(B_{12}) = 1 = \text{rank}(b)$ .

$\text{rank}(b_1) = 0, \text{rank}(b_2) = 1$ :

$$\Leftrightarrow B_{12} = \begin{bmatrix} \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{0}{0} \\ \frac{0}{0} & \frac{1}{1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = b$$

- \*  $\text{rank}(B_{21}) = 0 = \text{rank}(c)$ :  $B_{21} = 0$

- \*  $\text{rank}(B_{21}) = 1 = \text{rank}(c)$ :

- $\text{rank}(c_1) = 0, \text{rank}(c_2) = 1$ :  $c = \begin{bmatrix} 0 & 0 \mid p \end{bmatrix}, p \in \mathbb{F}^*/(\mathbb{F}^*)^2$ :

The nonzero diagonals in  $b$  and  $c$  are both at the right bottom corner, so it corresponds to sequence  $S1$ :  $2 * 2, 2 + 2 * 4$ , which implies  $h = 2$  is the greatest common factor of sequence:  $4, 10$ .)

$$B_{21} = \begin{bmatrix} 0 \mid 0 \parallel 0 \parallel 0 \parallel 0 & 0 & 0 & 0 \mid p \\ 0 \mid 0 \parallel 0 \parallel 0 \parallel 0 & 0 & 0 & 0 \mid 0 \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \end{bmatrix} = c.$$

$$\circ \text{rank}(c_1) = 1, \text{rank}(c_2) = 0: c = \left[ \begin{array}{cc|c} 0 & 1 & 0 \end{array} \right].$$

Here we have 1 instead of  $p$  because  $\beta_1 = 2 > 1$ .

$$B_{21} = \left[ \begin{array}{c|c|c|c|c} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] = \left[ \begin{array}{cc} c_1 & c_2 \end{array} \right] = c.$$

$$\bullet \text{rank}(B_{22}) = 0: B_{22} = 0.$$

$$- \text{rank}(B_{12}) = 0 \text{ same as before.}$$

$$- \text{rank}(B_{12}) = 1 \text{ same as before but } p = 1 \text{ since we have } B_{22} = 0.$$

$$- \text{rank}(B_{12}) = 2 = \text{rank}(b)$$

$$* \text{rank}(b_1) = 2, \text{rank}(b_2) = 0:$$

$$b = \left[ \begin{array}{c|c} 1 & \\ \hline & 1 \\ 0 & 0 \end{array} \right] = \left[ \begin{array}{c} I \\ 0 \end{array} \right] = \left[ \begin{array}{c} b_1 \\ b_2 \end{array} \right] \quad (b_1 = I).$$

$$\circ \text{rank}(c) = 0 \text{ or } 1, \text{ we have same form for } c \text{ as before.}$$

$$\circ \text{rank}(c_1) = 2, \text{rank}(c_2) = 0: c = \left[ \begin{array}{c|c} p & 0 \\ & 1 \\ & 0 \end{array} \right] = \left[ \begin{array}{cc} I & 0 \end{array} \right] = \left[ \begin{array}{cc} c_1 & c_2 \end{array} \right], p \in \mathbb{F}^*/(\mathbb{F}^*)^2.$$

It is Part I in classification theorem ( $b$  and  $c$  are both at the left top corner this time) and  $h = 2$  is the greatest common facotor of the first sequence  $S1$ :  $2 + 2 * 2, 2 * 4$ , or  $6, 8$ , which implies its greatest common factor is 2.

$$\circ \text{rank}(c_1) = 1, \text{rank}(c_2) = 1: c = \left[ \begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right].$$

$$* \text{rank}(b_1) = 1, \text{rank}(b_2) = 1: b = \left[ \begin{array}{c|c} 0 & 0 \\ \hline 1 & 0 \\ 0 & 1 \end{array} \right].$$

$$\circ \text{rank}(c_1) = 2, \text{rank}(c_2) = 0: c = \left[ \begin{array}{c|c} 1 & 0 \\ & 1 \\ & 0 \end{array} \right] = \left[ \begin{array}{cc} I & 0 \end{array} \right].$$

$$\circ \text{rank}(c_1) = 1, \text{rank}(c_2) = 1: c = \left[ \begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right].$$

## 6.2 The first type of inner involution with different sizes of $I_1$ and $I_2$

The inner involution we have to study here is



$$\theta(g) = \begin{bmatrix} I_1 & \\ & -I_2 \end{bmatrix}^{-1} g \begin{bmatrix} I_1 & \\ & -I_2 \end{bmatrix},$$

where  $I_1$  and  $I_2$  have different sizes. Without loss of generality, we assume the size  $n_1$  of  $I_1$  is greater than size  $n_2$  of  $I_2$ , and we can therefore assume that, there exists identity matrix  $I_3$  of size  $n_1 - n_2$  such that

$$I_1 = \begin{bmatrix} I_2 & \\ & I_3 \end{bmatrix}.$$

To study  $(G_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F}), \theta)$ , we can study a restriction of  $(\tilde{G}_{\mathbb{F}} = \mathrm{SL}(2n_1, \mathbb{F}), \sigma)$ , where

$$\sigma(\tilde{g}) = \begin{bmatrix} I_1 & \\ & -I_1 \end{bmatrix}^{-1} \tilde{g} \begin{bmatrix} I_1 & \\ & -I_1 \end{bmatrix},$$

and  $G_{\mathbb{F}}$  can be regarded as a restriction of  $\tilde{G}_{\mathbb{F}}$ :

$$g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \hookrightarrow \tilde{g} = \begin{bmatrix} g & \\ & I_3 \end{bmatrix} = \left[ \begin{array}{cc|c} g_{11} & g_{12} & 0 \\ g_{21} & g_{22} & 0 \\ \hline 0 & 0 & I_3 \end{array} \right] = \left[ \begin{array}{cc|c} g_{11} & g_{12} & 0 \\ g_{21} & g_{22} & 0 \\ \hline 0 & 0 & I_3 \end{array} \right],$$

where  $g_{11} \in \mathfrak{gl}(n_1, \mathbb{F})$ .

From analysis for the key inner involution, we know that we only need to study the case with  $g_{11} = I_1, g_{22} = I_2$ . For classification, we only need to include  $K_{\mathbb{F}}$ -orbits in key inner involution with the last  $n_1 - n_2$  rows of  $C$  zero, and the last  $n_1 - n_2$  columns of  $B$  zero.

## 6.3 Second type of inner involution

We first define a special type of power function, then translate the second type of inner involution to a special form of key inner involution, and classify  $K_{\mathbb{F}}$ -orbits of equivalent problem.

### 6.3.1 $\{\}$ power function

Given a field  $\mathbb{F}$ , suppose "conjugation operator"  $c : \mathbb{F} \rightarrow \mathbb{F}$  is defined:  $(a, b \in \mathbb{F})$

$$\begin{aligned}
c(a \pm b) &= c(a) \pm c(b) \\
c(a * b) &= c(a) * c(b) \\
c\left(\frac{1}{a}\right) &= \frac{1}{c(a)}, \text{ if } a \neq 0 \\
c^2(a) &= a.
\end{aligned}$$

We denote  $c(a) = \bar{a}$  for  $a \in \mathbb{F}$ . With conjugation operator we can define a "curly bracket" power function over  $\mathbb{F}^*$  in the following way:

**Definition 6.3.1.** Let  $c$  be a conjugation operator defined over field  $\mathbb{F}$ . We define "curly bracket" power function, denoted by " $\{\}$ ", over  $\mathbb{F}^*$  as following: for  $a \in \mathbb{F}^*$

$$\begin{aligned}
n = 0 : a^{\{0\}} &= 1, \\
n \in \mathbb{Z}^+ : a^{\{n\}} &= \prod_{i=1}^n c^i(a) = \overbrace{a\bar{a}a\bar{a}\cdots}^n = \begin{cases} (a\bar{a})^m, & \text{if } n = 2m; \\ (a\bar{a})^m * a, & \text{if } n = 2m + 1, \end{cases} \\
n \in \mathbb{Z}^- : a^{\{n\}} &= \frac{1}{a^{\{-n\}}}, \\
m, n \in \mathbb{Z} : a^{\{m\}*\{n\}} &= (a^{\{m\}})^{\{n\}}.
\end{aligned}$$

*Remark 6.3.2.* We do not define following for  $m, n \in \mathbb{Z}$  and  $a \in \mathbb{F}^*$

$$\begin{aligned}
a^{\{m\}+\{n\}} &= a^{\{m\}} * a^{\{n\}}, \\
a^{\{m\}-\{n\}} &= \frac{a^{\{m\}}}{a^{\{n\}}},
\end{aligned}$$

since it makes no sense/use in our work.

It's easy to check that  $\{\}$  power function has following properties: ( $a, b \in \mathbb{F}^*, n, m \in \mathbb{Z}$ )

$$a^{\{m\}*\{n\}} = (a^{\{m\}})^{\{n\}} = a^{\{mn\}} = a^{\{n\}*\{m\}} \quad (6.3.1)$$

$$(ab)^{\{n\}} = a^{\{n\}} * b^{\{n\}} \quad (6.3.2)$$

$$\left(\frac{a}{b}\right)^{\{n\}} = \frac{a^{\{n\}}}{b^{\{n\}}} \quad (6.3.3)$$

We have following lemma for  $\{\}$  power function:

**Lemma 6.3.3.** *Given  $n_i \in \mathbb{Z}, i = 1, \dots, r, a \in \mathbb{F}^*$ , then following equation has solution  $x_i, i = 1, \dots, s$  over the field  $\mathbb{F}$ ,*

$$\prod_{i=1}^s x_i^{\{n_i\}} = a$$

*if and only  $a \in (F^*)^{\{l\}}$ , where  $l$  is the greatest common factor of  $n_i, i = 1, \dots, s$ , and  $(F^*)^{\{l\}} = \{a^{\{l\}} | a \in \mathbb{F}^*\}$ .*

*Proof.*  $\Rightarrow$ : Let  $n_i = lm_i$ , then

$$\begin{aligned} a &= \prod_{i=1}^s x_i^{\{n_i\}} \\ &= \prod_{i=1}^s x_i^{\{lm_i\}} \\ &\stackrel{6.3.1}{=} \prod_{i=1}^s (x_i^{\{m_i\}})^{\{l\}} \\ &\stackrel{6.3.2}{=} \left( \prod_{i=1}^s x_i^{\{m_i\}} \right)^{\{l\}} \in \mathbb{F}^{\{l\}}. \end{aligned}$$

$\Leftarrow$ : First, Without loss of generality, we can assume  $n_i$  are all positive here (we can replace  $x_i$  by  $\frac{1}{x_i}$  if  $n_{i_0} < 0$  for some  $i_0$ ).

Since  $l$  is the greatest common factor of  $n_i$ , using long division/Euclidean algorithm, we can find nonnegative integers  $m_i$  such that

$$l = \sum_{i=1}^s m_i n_i = \sum_{i=1}^{s_1} m_i n_i - \sum_{i=S_1+1}^s m_i n_i, \quad (6.3.4)$$

where  $1 \leq s_1 \leq s - 1$ .

We always have the negative part since  $l \leq \min_{1 \leq i \leq s} n_i$ . Without loss of generality, we assume  $m_i > 0, i = 1, \dots, s$ .

Since  $a \in (\mathbb{F}^*)^{\{l\}}$ , we can assume  $a = d^{\{l\}}$  for some  $d \in \mathbb{F}^*$ . Let  $x_1 = d^{\{m_1\}}, x_2 = (c^{m_1 n_1}(d))^{\{m_2\}}, x_{i+1} = (c^{\sum_{j=1}^i m_j n_j}(d))^{\{m_{i+1}\}}, 1 \leq i \leq s_1 - 1$ . First,  $x_1^{\{n_1\}} = (d^{\{m_1\}})^{\{n_1\}} \stackrel{6.3.1}{=} d^{\{m_1 n_1\}}$ . We know that in  $\{\}$  power function  $d$  and  $\bar{d}$  appear alternatively, so  $x_1^{\{n_1\}}$  ends with  $c^{m_1 n_1 - 1}(d)$ . Noticing  $x_2^{\{n_2\}} = (c^{m_1 n_1}(d))^{\{m_2 n_2\}}$  begins with  $c^{m_1 n_1}(d)$ , we can see that

$$x_1^{\{n_1\}} * x_2^{\{n_2\}} = d^{\{m_1 n_1 + m_2 n_2\}}.$$

In the same way,  $x_1^{n_1} * x_2^{n_2}$  ends with  $c^{m_1 n_1 + m_2 n_2 - 1}(d)$ . Since  $x_3^{\{n_3\}} = (c^{m_1 n_1 + m_2 n_2}(d))^{\{m_3 n_3\}}$  begins with  $c^{m_1 n_1 + m_2 n_2}(d)$ , we can see that

$$x_1^{\{n_1\}} * x_2^{\{n_2\}} * x_3^{\{n_3\}} = d^{\{m_1 n_1 + m_2 n_2 + m_3 n_3\}}.$$

Repeating the above process we get

$$\prod_{i=1}^{s_1} x_i^{\{n_i\}} = d^{\left\{ \sum_{i=1}^{s_1} m_i n_i \right\}} = \overbrace{d d d d d d d d}^{\sum_{i=1}^{s_1} m_i n_i} \dots t t t,$$

where  $\prod_{i=1}^{s_1} x_i^{\{n_i\}}$  ends with  $t = c^{\sum_{i=1}^{s_1} \{m_i n_i\} - 1}(d)$ .

Let  $x_{s_1+1} = t^{\{m_{s_1+1}\}}$ . We generate  $x_i, i = s_1 + 2, \dots, s$  in a way same as above to get

$$\begin{aligned} \prod_{i=s_1+1}^s x_i^{\{n_i\}} &= t^{\sum_{i=s_1+1}^s \{m_i n_i\}} \\ &= \overbrace{t t t t t t t t}^{\sum_{i=s_1+1}^s m_i n_i} \dots \\ &= \overbrace{\dots t t t t t t t t}^{\sum_{i=s_1+1}^s m_i n_i}, \end{aligned}$$

where in the last equality we simply reverse the order of multiplications.

Let

$$y_i = \begin{cases} x_i, & 1 \leq i \leq s_1 \\ \frac{1}{x_i} & s_1 + 1 \leq i \leq s. \end{cases}$$

We therefore have

$$\begin{aligned}
\prod_{i=1}^s y_i^{\{n_i\}} &= \prod_{i=1}^{s_1} y_i^{\{n_i\}} * \prod_{i=s_1+1}^s y_i^{\{n_i\}} \\
&= \prod_{i=1}^{s_1} x_i^{\{n_i\}} * \frac{1}{\prod_{i=s_1+1}^s x_i^{\{n_i\}}} \\
&= \frac{\overbrace{d\bar{d}d\bar{d}d\bar{d}d\bar{d}\dots\dots t\bar{t}t}^{\sum_{i=1}^{s_1} m_i n_i}}{\underbrace{\dots\dots\dots \bar{t}\bar{t}\bar{t}\bar{t}\bar{t}\bar{t}\bar{t}}_{\sum_{i=s_1+1}^s m_i n_i}} \\
&= \overbrace{d\bar{d}d\bar{d}d\bar{d}d\bar{d}\dots\dots\dots}^{\sum_{i=1}^{s_1} m_i n_i - \sum_{i=s_1+1}^s m_i n_i} \quad \left( \sum_{i=1}^{s_1} m_i n_i - \sum_{i=s_1+1}^s m_i n_i = l > 0 \right) \\
&= \overbrace{d\bar{d}d\bar{d}d\bar{d}d\bar{d}\dots\dots\dots}^l \\
&= d^{\{l\}} = a.
\end{aligned}$$

□

For the extension field  $\mathbb{F}(\sqrt{p})$  of  $\mathbb{F}$  with  $0 \neq p \in \mathbb{F} \setminus \mathbb{F}^2$ , we define the conjugation of  $a + b\sqrt{p} \in \mathbb{F}(\sqrt{p})$  as

$$c(a + b\sqrt{p}) = \overline{a + b\sqrt{p}} = a - b\sqrt{p},$$

and the curly bracket power function  $\{\}$  is defined over the extension field  $\mathbb{F}(\sqrt{p})$  accordingly.

### 6.3.2 Translation to key inner involution

Let's first review what do for the case  $n = 2$ , i.e., how we translate the original problem to the one with "key involution". For  $g \in \tilde{G}_{\mathbb{F}} = \text{SL}(2, \mathbb{F})$ ,  $p \in \mathbb{F}^* \setminus (\mathbb{F}^*)^2$ , i.e.,  $p$  is a non-square element, we have (abusing the letter  $p$ )

$$\begin{aligned}
p &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, p_1 = \begin{bmatrix} 1 & \\ & \sqrt{p} \end{bmatrix}, p_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, p_\sigma = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \tilde{p} = \begin{bmatrix} & 1 \\ p & \end{bmatrix} \\
\sigma(g) &= \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \\
\theta(g) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
\tilde{\theta}(g) &= \begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} 0 & 1 \\ q & 0 \end{bmatrix} = \tilde{p}^{-1} g \tilde{p} \\
\theta_{p_1}(g) &= p_1^{-1} g p_1 \\
\theta &= \theta_p^{-1} \circ \sigma \circ \theta_p \\
\tilde{\theta} &= \theta_{p_1}^{-1} \circ \theta \circ \theta_{p_1}.
\end{aligned}$$

Instead of studying  $(\tilde{G}_{\mathbb{F}}, \tilde{\theta})$ , we study  $(G_{\mathbb{F}}, \sigma)$  equivalently, where

$$G_{\mathbb{F}} = \{g = \begin{bmatrix} A & \overline{D} \\ D & \overline{A} \end{bmatrix} \mid A, D \in \mathbb{F}(\sqrt{p}), \det g = 1\}.$$

To be specific, the correspondence of elements in  $\tilde{G}_{\mathbb{F}}$  and  $G_{\mathbb{F}}$  is given by

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \rightarrow \tilde{g} = \begin{bmatrix} (a+d) + \sqrt{p}(b + \frac{c}{p}) & (-a+d) + \sqrt{p}(b - \frac{c}{p}) \\ (-a+d) - \sqrt{p}(b - \frac{c}{p}) & (a+d) - \sqrt{p}(b + \frac{c}{p}) \end{bmatrix}.$$

The fixed point subgroup  $K_{\mathbb{F}}$  for  $(\sigma, G_{\mathbb{F}})$  is also changed to

$$K_{\mathbb{F}} = \{k = \begin{bmatrix} z & \\ & \bar{z} \end{bmatrix} \mid z \in \mathbb{F}(\sqrt{p}), \det k = 1\}.$$

Now we can generalize this translation to the  $n \times n$  case. Let

$$\begin{aligned}
P &= \begin{bmatrix} p & & \\ & \ddots & \\ & & p \end{bmatrix}, P_1 = \begin{bmatrix} p_1 & & \\ & \ddots & \\ & & p_1 \end{bmatrix}, P_2 = \begin{bmatrix} p_2 & & \\ & \ddots & \\ & & p_2 \end{bmatrix}, \\
P_{\tilde{\sigma}} &= \begin{bmatrix} p_{\sigma} & & \\ & \ddots & \\ & & p_{\sigma} \end{bmatrix}, P_{\sigma} = \begin{bmatrix} I & \\ & -I \end{bmatrix}, \tilde{P} = \begin{bmatrix} \tilde{p} & & \\ & \ddots & \\ & & \tilde{p} \end{bmatrix}, \\
\sigma(g) &= \begin{bmatrix} I & \\ & -I \end{bmatrix}^{-1} \cdot g \cdot \begin{bmatrix} I & \\ & -I \end{bmatrix} = P_{\sigma}^{-1} g P_{\sigma}, \\
\tilde{\sigma}(g) &= \left[ \begin{array}{c|c|c} 1 & & \\ \hline & -1 & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & -1 \end{array} \right]^{-1} \cdot g \cdot \left[ \begin{array}{c|c|c} 1 & & \\ \hline & -1 & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & -1 \end{array} \right] = P_{\tilde{\sigma}}^{-1} g P_{\tilde{\sigma}}, \\
\theta(g) &= \left[ \begin{array}{c|c|c} 1 & & \\ \hline 1 & & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & 1 \end{array} \right]^{-1} \cdot g \cdot \left[ \begin{array}{c|c|c} 1 & & \\ \hline 1 & & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & 1 \end{array} \right] = P_2^{-1} \cdot g \cdot P_2, \\
\tilde{\theta}(g) &= \left[ \begin{array}{c|c|c} 1 & & \\ \hline p & & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & p \end{array} \right]^{-1} \cdot g \cdot \left[ \begin{array}{c|c|c} 1 & & \\ \hline p & & \\ \hline & & \ddots \\ \hline & & & 1 \\ & & & p \end{array} \right] = \tilde{P}^{-1} \cdot g \cdot \tilde{P}, \\
\theta_{p_1}(g) &= P_1^{-1} g P_1, \\
\theta_P(g) &= P^{-1} g P, \\
\theta_Q(g) &= Q^{-1} g Q, \\
\tilde{\sigma} &= \theta_Q^{-1} \circ \sigma \circ \theta_Q, \\
\theta &= \theta_P^{-1} \circ \tilde{\sigma} \circ \theta_P, \\
\tilde{\theta} &= \theta_{P_1}^{-1} \circ \theta \circ \theta_{P_1},
\end{aligned}$$

where  $p, p_1, p_2, p_{\sigma}, \tilde{p}$  are defined for the  $2 \times 2$  case,  $Q$  is the permutation matrix such that

$$P_{\tilde{\sigma}} = Q^{-1}P_{\sigma}Q.$$

Similarly, instead of studying  $(\tilde{G}_{\mathbb{F}} = \mathrm{SL}(n, \mathbb{F}), \tilde{\theta})$ , we study  $(G_{\mathbb{F}}, \sigma)$  equivalently, where (for subscripts we write  $\mathbb{F}$  instead of  $\mathbb{F}(\sqrt{p})$  for simplicity of notations)

$$G_{\mathbb{F}} = \{g = \begin{bmatrix} A & \overline{D} \\ D & \overline{A} \end{bmatrix} \mid A, D \in \mathrm{gl}(\frac{n}{2}, \mathbb{F}(\sqrt{p})), \det g = 1\},$$

$$K_{\mathbb{F}} = \{k = \begin{bmatrix} z & \\ & \overline{z} \end{bmatrix} \in G_{\mathbb{F}} \mid z \in \mathrm{GL}(\frac{n}{2}, \mathbb{F}(\sqrt{p})), \det k = 1\}.$$

Using conclusions from the case with "key involution", we can assume here  $A = I$  in the above  $g$ , so  $g = I + J$  with  $J$  nilpotent. To be precise, we are able to study following problem equivalently:

We know  $J = \begin{bmatrix} & \overline{D} \\ D & \end{bmatrix}$  is nilpotent and  $k = \begin{bmatrix} z & \\ & \overline{z} \end{bmatrix} \in K_{\mathbb{F}}$  ( $\det k = 1$ ). The problem we need to solve is that, given  $J, \tilde{J}$  nilpotent, find out if there exists  $k \in K_{\mathbb{F}}$  such that  $\tilde{J} = k^{-1}Jk$ , i.e., using  $k \in K_{\mathbb{F}}$ , what kind of canonical form of  $J$  can we reduce to?

### 6.3.3 Classification for the second type of inner involution

We have  $J \text{ nilp} \Leftrightarrow D\overline{D} \text{ nilp}$  and

$$\begin{aligned} \tilde{J} &= k^{-1}Jk \\ &= \begin{bmatrix} z & \\ & \overline{z} \end{bmatrix}^{-1} \begin{bmatrix} & \overline{D} \\ D & \end{bmatrix} \begin{bmatrix} z & \\ & \overline{z} \end{bmatrix} \\ &= \begin{bmatrix} & z^{-1}\overline{D}\overline{z} \\ \overline{z}^{-1}Dz & \end{bmatrix} \\ &= \begin{bmatrix} & \tilde{\overline{D}} \\ \tilde{D} & \end{bmatrix} \end{aligned}$$

Since  $\det \overline{D} = \overline{(\det D)}$  and  $D\overline{D}$  is nilpotent, we know that 0 is an eigenvalue of  $D$ . Therefore there exists  $0 \neq v_1 \in (\mathbb{F}(\sqrt{p}))^{\frac{n}{2}}$  s.t.  $Dv_1 = 0$ . Choose  $\frac{n}{2} - 1$  vectors  $v_2, \dots, v_{\frac{n}{2}}$  in  $(\mathbb{F}(\sqrt{p}))^{\frac{n}{2}}$  s.t.

$z = (v_1, v_2, \dots, v_{\frac{n}{2}}) \in \mathrm{SL}(\frac{n}{2}, \mathbb{F}(\sqrt{p})) \Rightarrow k = \begin{bmatrix} z & \\ & \overline{z} \end{bmatrix} \in K_{\mathbb{F}}$ , so we have



$$\begin{aligned}
Dz &= z \left[ \begin{array}{c|cc} 0 & * & * \\ \hline 0 & * & * \\ 0 & * & * \end{array} \right] \\
\Rightarrow \tilde{D} &= \bar{z}^{-1} Dz = \bar{z}^{-1} z \left[ \begin{array}{c|cc} 0 & * & * \\ \hline 0 & * & * \\ 0 & * & * \end{array} \right] = \left[ \begin{array}{c|c} 0 & t \\ \hline 0 & D_1 \end{array} \right] \\
\Rightarrow \tilde{D} \overline{\tilde{D}} &= \begin{bmatrix} 0 & t\overline{D_1} \\ 0 & D_1\overline{D_1} \end{bmatrix}
\end{aligned} \tag{6.3.5}$$

and

$$\tilde{J} \text{ nilp} \Leftrightarrow \tilde{D} \overline{\tilde{D}} \text{ nilp} \Leftrightarrow D_1 \overline{D_1} \text{ nilp}$$

therefore the above process can be repeated.

We define  $L_{i,j,s}$  as following

$$L_{i,j,s} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & s & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix},$$

where the entry  $s$  is in the  $(i,j)$ -position of  $L_{i,j,s}$ . It's obvious that  $\overline{L_{i,j,s}} = L_{i,j,\bar{s}}$  and  $k = \begin{bmatrix} L_{i,j,s} & \\ & \overline{L_{i,j,s}} \end{bmatrix} \in K_{\mathbb{F}}$ .

If  $n = 2$ , then  $D\overline{D} \in \mathfrak{gl}(\frac{n}{2}, \mathbb{F}(\sqrt{p})) = \mathfrak{gl}(1, \mathbb{F}(\sqrt{p})) \text{ nilp} \Leftrightarrow D = 0$ . This is why we have the identity orbit only when  $n = 2$ .

If  $n = 4$ , then we can see that  $J \text{ nilp} \Leftrightarrow D_1 \overline{D_1} \in \mathfrak{gl}(\frac{n}{2} - 1, \mathbb{F}(\sqrt{p})) = \mathfrak{gl}(1, \mathbb{F}(\sqrt{p})) \text{ nilp} \Leftrightarrow D_1 = 0$  and we can therefore assume

$$D = \begin{bmatrix} 0 & q \\ 0 & 0 \end{bmatrix} \text{ with } q \in \mathbb{F}(\sqrt{p}).$$

We assume it's true for  $n = 2m, 2 \leq m \in \mathbb{Z}^+$ , that  $D$  in  $J$  can be reduced to following form

$$D = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix} \quad (6.3.6)$$

which satisfies following conditions

1. The sizes of  $J_i, 1 \leq r \leq r$  are decreasing, especially if  $J_{i_0} = 0$ , then  $J_{i_0+1} = \cdots = J_r = 0$ ;
2. If  $J_i \neq 0$ , then

$$J_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & q \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (6.3.7)$$

where  $q \in \mathbb{F}(\sqrt{p})$  and  $q = 1$  if  $1 \leq i \leq r-1$ .

Assume the size of  $J_i$  for  $1 \leq i \leq r$  is  $n_i$ .

We now prove that the above conditions are true for  $n = 2m + 2$  too. We assume  $D$  to be in the form of (6.3.5) with  $D_1 \in \mathfrak{gl}(m, \mathbb{F}(\sqrt{p}))$ , so we can assume  $D_1$  has a form in (6.3.6) with the above two conditions satisfied ( $D_1 \bar{D}_1$  is nilpotent). The proof is put into following three cases.

**Case I.**  $D_1 = 0$ . Then

$$D = \left[ \begin{array}{c|cccc} 0 & t_1 & \cdots & t_m \\ \hline 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{array} \right].$$

If  $t_i = 0, i = 1, \dots, m$ , then  $D = 0$ . Done.

If there is  $t_{i_0} \neq 0, 1 \leq i_0 \leq m$ , using permutation matrix we can make  $i_0 = 1$ . By some diagonal matrix in  $K_{\mathbb{F}}$  we can make  $t_1 = 1$ . Then by choosing  $s$  in  $L_{2,3,s}$  and applying  $L_{2,3,s}$  to  $D$ :  $\bar{L}_{2,3,s} D L_{2,3,s}$  we can make  $t_2 = 0$  and make  $t_j = 0$  in the same way. Now  $D$  satisfies required conditions.

**Case II.**  $D_1$  has one Jordan block only. In this case we have

$$D = \left[ \begin{array}{c|cccccc} 0 & t_1 & t_2 & t_3 & \cdots & t_{m-1} & t_m \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & q \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array} \right].$$

with  $0 \neq \mathbb{F}(\sqrt{p})^*$  (or  $D$  has two Jordan blocks).

By applying  $L_{1,j,s_j}, j = 3, 4, \dots, m+1$  to  $D$  with  $s_j = \bar{t}_{j-1}, j = 3, \dots, m$  and  $s_{m+1} = \frac{\bar{t}_m}{q}$ , we can make  $t_i, i = 2, \dots, m$  in the new  $D$ , i.e.,  $\tilde{D}$ , to be zero. We still write  $\tilde{D}$  as  $D$ .

If  $t_1 \neq 0$ , let  $s_1 = 1, s_2 = t_1^{-1}, s_{i+1} = \bar{s}_i, i = 2, \dots, m-1, s_{m+1} = (\prod_{i=1}^m s_i)^{-1}$ , and  $z = \text{diag}(s_1, \dots, s_{m+1})$ . Then

$$\bar{z}^{-1} D z = \left[ \begin{array}{c|cccccc} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \bar{q} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array} \right],$$

which is till one Jordan block and has the expected form.

If  $t_1 = 0$ , then

$$D = \left[ \begin{array}{c|cccccc} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & q \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array} \right].$$

Using permutation matrix we can make  $D$  into

$$D = \left[ \begin{array}{cccccc|c} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array} \right],$$

Same as before we can make  $q = 1$ .

**Case III.**  $D$  has two or more Jordan blocks. In this case we have

$$D = \left[ \begin{array}{c|ccc} 0 & t_1 & \cdots & t_m \\ \hline 0 & J_1 & & \\ 0 & & \ddots & \\ 0 & & & J_r \end{array} \right].$$

If we have at most one zero Jordan block in  $D$ , i.e.,  $J_i \neq 0, 1 \leq i \leq r - 1$ . Like in Case II, we can make  $t_i = 0$  unless  $t_i$  corresponds to the first column of some Jordan block. Suppose the first columns of each Jordan blocks are the  $i_1 = 2, i_2, \dots, i_r$ -th columns of  $D$ .

If  $t_{i_j} = 0, j = 1, \dots, r$ , i.e.,  $t_i = 0$ , we simply need to permute 0 on the left top corner to the right bottom corner, and make the possible  $q$  in the last Jordan block, if this is the case, to be 1 to finish this case.

If  $t_{i_j} \neq 0, j = 1, \dots, r$ , same as in Case II, we can make  $t_1 = 1$  (notice that the possible  $q$  in the last Jordan block might change correspondingly too, if we have no zero Jordan block).

Let  $t_{i_2} = -u \neq 0$ , by applying  $L_{j+1, i_2+j, c^j(u)}, j = 0, 1, \dots, n_2 - 1$  to  $D$  in order, we get

$$\begin{aligned} D &\rightarrow \overline{L}_{1, i_2, u}^{-1} D L_{1, i_2, u} \\ &\rightarrow \overline{L}_{2, i_2+1, c(u)}^{-1} (\overline{L}_{1, i_2, u}^{-1} D L_{1, i_2, u}) L_{2, i_2+1, c(u)} \\ &\rightarrow \cdots \\ &\rightarrow \overline{L}_{n_2, i_2+n_2-1, c^{n_2-1}(u)}^{-1} \cdots (\overline{L}_{1, i_2, u}^{-1} D L_{1, i_2, u}) \cdots L_{n_2, i_2+n_2-1, c^{n_2-1}(u)}. \end{aligned}$$

Visually, it'd be like this

$D =$







$$\rightarrow \left[ \begin{array}{c|cccccc|ccccc|c} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & & & & & & \cdots \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & & & & & & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & & & & & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & & & & & & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & & & & & & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & & & & & \cdots \\ \hline 0 & & & & & & & 0 & 1 & \cdots & 0 & 0 & \cdots \\ \cdots & & & & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 0 & 1 & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 0 & 0 & \cdots \\ \hline \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{array} \right]$$

$$\rightarrow \left[ \begin{array}{c|cccccc|ccccc|c} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & & & & & & \cdots \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & & & & & & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & & & & & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & & & & & & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & & & & & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & & & & & \cdots \\ \hline 0 & & & & & & & 0 & 1 & \cdots & 0 & 0 & \cdots \\ \cdots & & & & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 1 & 0 & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 0 & 1 & \cdots \\ 0 & & & & & & & 0 & 0 & \cdots & 0 & 0 & \cdots \\ \hline \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{array} \right].$$

*Remark 6.3.4.* 1. Since the size of  $J_1$  is greater than or equal to the size of  $J_2$ , we can continue the above process without break.

2. In the last step, when multiplying  $L_{n_2, i_2+n_2-1, c^{n_2-1}(u)}$  from right, we make  $c^{n_2-1}(u) \rightarrow 0$ . When multiplying  $\bar{L}_{n_2, i_2+n_2-1, c^{n_2-1}(u)}^{-1}$  from left, since the  $(i_2 + n_2 - 1)$ -th row of  $D$  is zero, we would not introduce anything nonzero.

We therefore make  $t_2 = 0$  without changing anything else. In this same way, we can make

all  $t_i = 0$  for  $3 \leq i \leq r$ , and  $D$  is then in the expected form. The difference between  $D$  and  $D_1$  is that the size of the largest Jordan block in  $D$  is one more than that of the largest Jordan block in  $D_1$ , and possibly  $q$ 's might be different in  $D$  and  $D_1$ .

If only some of  $t_i$  are nonzero, we then only need to consider those Jordan blocks with nonzero corresponding elements in the first row of  $D$ . Same as before we first by permutation matrices re-sort Jordan blocks in a way that

1.  $t_i, i = 1, \dots, r$ , begin with nonzero elements and all nonzero elements cluster together, i.e., if  $t_{i_0} = 0$ , then  $t_i = 0$  for all  $i \geq i_0$ ;
2. sizes of Jordan blocks with nonzero  $t_i$  are decreasing.

*Remark 6.3.5.* Permutation does not change the correspondence between Jordan blocks and elements in the first row corresponding to these columns of Jordan blocks.

We then make  $t_1 = 1$ , and make nonzero  $t_i$  into 0. Notice that here we only work with Jordan blocks with nonzero  $t_i$ , and  $J_1$  after re-sort has the largest size among them, so the above elimination process can be adopted. We are done if the sizes of  $J_i$  are in decreasing order and finish this subcase by permutation matrices if they are not in decreasing order.

Now let's consider the subcase that  $D$  could have more than one zero Jordan blocks.

If elements in the first row corresponding the zero Jordan blocks are all zero, then it's reduced the previous subcase.

If elements in the first row corresponding the zero Jordan blocks are not all zero, by permutation matrix we can assume the element in the first row corresponding to the first zero Jordan block is nonzero, we can then make all elements in the first row "after" this element to be zero by choosing a proper  $L_{1,j,s}$ . Now we reduce the first subcase of Case III again.

So the conditions we assume to be true for  $n = 2m$  are still true for  $n = 2m + 2$ , and by induction, those conditions are true for all positive even number  $n$ .

Let's discuss classification. First let's see what  $z = x + \sqrt{p}y$  we can take between  $D$ 's satisfying the above conditions. We have following lemma.

**Lemma 6.3.6.** *If  $D$  and  $\tilde{D}$  satisfy the above two conditions and if there is  $k$  such that  $\tilde{J} = k^{-1}Jk$ , where*

$$J = \begin{bmatrix} & \overline{D} \\ D & \end{bmatrix}, \tilde{J} = \begin{bmatrix} & \overline{\tilde{D}} \\ \tilde{D} & \end{bmatrix}, k = \begin{bmatrix} z & \\ & \bar{z} \end{bmatrix},$$

*then corresponding Jordan blocks in  $D$  and  $\tilde{D}$  must be of same sizes.*



*Proof.* Since

$$\widetilde{\widetilde{D}} \overline{\widetilde{D}} = \overline{z}^{-1} D z \cdot \overline{(\overline{z}^{-1} D z)} = \overline{z}^{-1} D z \cdot z^{-1} \overline{D} \overline{z} = \overline{z}^{-1} D \overline{D} \overline{z}$$

we can see that  $\widetilde{\widetilde{D}} \overline{\widetilde{D}}$  and  $D \overline{D}$  have same Jordan blocks, and then it's obvious that corresponding Jordan blocks of  $\widetilde{\widetilde{D}} \overline{\widetilde{D}}$  and  $D \overline{D}$  have same sizes. This implies that all Jordan blocks of size greater than or equal to three of  $\widetilde{\widetilde{D}}$  and  $D$  are same, except possible different  $q$ 's in the last Jordan blocks of  $\widetilde{\widetilde{D}}$  and  $D$ . Since  $\text{rank}(\widetilde{\widetilde{D}}) = \text{rank}(D)$ , we know that  $\widetilde{\widetilde{D}}$  and  $D$  have same number of Jordan blocks of size two too. The number of zero Jordan blocks of  $\widetilde{\widetilde{D}}$  and  $D$  are now same too.  $\square$

If  $D$  has zero Jordan blocks, then there is no  $q$  in  $D$ , and then  $D$  itself represents one  $K_{\mathbb{F}}$ -orbit.

If there is no zero Jordan block in  $D$ , then there might some  $q$  instead of 1 in the last Jordan block of  $D$ . We now want to see what kind of  $z$  we can take in  $\widetilde{\widetilde{D}} = \overline{z}^{-1} D z$  if  $D$  and  $\widetilde{\widetilde{D}}$  have same types of Jordan blocks.

Notice that

$$D = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix}, \widetilde{\widetilde{D}} = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & \widetilde{J}_r \end{bmatrix}$$

with

$$J_i = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & q \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \widetilde{J}_r = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & \widetilde{q} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

where  $q = 1$  if  $1 \leq i \leq r-1$ , and sizes of  $J_i(\widetilde{J}_r), i = 1, \dots, r$ , are decreasing.

Let  $z = (z_{ij})_{r \times r}$  be partitioned according to  $J_i$ , so  $z_{ij} \in \mathbb{F}(\sqrt{p})^{n_i \times n_j}$ .

From  $\widetilde{\widetilde{D}} = \overline{z}^{-1} D z$  or  $D z = \overline{z} \widetilde{\widetilde{D}}$ , we have

- (1)  $J_i z_{ij} = \overline{z}_{ij} J_j, 1 \leq i, j \leq r-1;$
- (2)  $J_r z_{rj} = \overline{z}_{rj} \widetilde{J}_j, j = 1, \dots, r-1;$
- (3)  $J_i z_{ir} = \overline{z}_{ir} J_r, i = 1, \dots, r-1;$

$$(4) J_r z_{rr} = \bar{z}_{rr} \tilde{J}_r.$$

For (1), we have

$$(a) z_{ij}^{k+1,l} = \bar{z}_{ij}^{k,l-1};$$

$$(b) z_{ij}^{k,1} = 0 = z_{ij}^{n_i,l}, k \geq 2, l \leq n_i - 1.$$

For (2), besides (a) for  $1 \leq k \leq n_r - 2$  and (b), we have for the last row of  $z_{ij}$ :

$$(c) z_{rj}^{n_r,l} = \frac{1}{q} \bar{z}_{rj}^{n_r-1,l-1}, l = 2, \dots, n_j.$$

For (3), besides (a) for  $2 \leq l \leq n_r - 1$  and (b) we have for the last column of  $z_{ij}$ :

$$(d) z_{ir}^{k+1,r} = \bar{z}_{ir}^{k,r-1} \tilde{q}, k = 1, \dots, n_i - 1.$$

For (4), besides (a) for  $1 \leq k \leq n_r - 2, 2 \leq l \leq n_r - 1$ , (b), (c) for  $l = 2, \dots, n_r - 1$  and (d) for  $k = 1, \dots, n_r - 2$  we have

$$(e) z_{rr}^{n_r,n_r} = \frac{1}{q} \bar{z}_{rr}^{n_r-1,n_r-1} \tilde{q}.$$

Let  $\alpha = (\alpha_{ij})_{r \times r}$  with  $\alpha_{ij} \in \mathbb{F}(\sqrt{p})^{n_i \times n_j}$  and satisfy conditions (a) and (b) for all valid  $k$  and  $l$ . Then it's not hard to see that  $z = D_q^{-1} \alpha D_{\tilde{q}}$  where  $D_q = \text{diag}(1, \dots, 1, q)$  and  $D_{\tilde{q}} = \text{diag}(1, \dots, 1, \tilde{q})$ . So

$$\det z = \det(D_q^{-1} \alpha D_{\tilde{q}}) = \det(D_q^{-1}) \det \alpha \det(D_{\tilde{q}}) = q^{-1} \det \alpha \tilde{q} = a + b\sqrt{p}. \quad (6.3.8)$$

The determinant condition  $1 = \det z \det \bar{z} \Leftrightarrow a + b\sqrt{p} \in S^1$ , where

$$S^1 = \{a + b\sqrt{p} \mid a^2 - b^2 p = 1, a, b \in \mathbb{F}\},$$

which is a multiplicative subgroup of  $\mathbb{F}(\sqrt{p})^*$ , namely, the unit circle.

It's obvious that the set of  $\alpha$  satisfying conditions (a) and (b) form a subgroup of  $\text{GL}(\frac{n}{2}, \mathbb{F}(\sqrt{p}))$ , so the set of  $\det \alpha$  form a subgroup of  $\mathbb{F}(\sqrt{p})^*$ . We have a result similar to that in the key inner involution, Lemma 6.1.24, without the 0-th row and column there, but conjugation involved here. Also, the order of Jordan blocks here are decreasing.

**Lemma 6.3.7.**  $\det \alpha = \det(\tilde{\alpha})$ , where  $\tilde{\alpha} = (\tilde{\alpha}_{ij})$  is formed from  $\alpha$ :

1. The size of  $\tilde{\alpha}_{ij}$  is same as the size of  $\alpha_{ij}$  for  $i, j = 1, \dots, r$ ;
2.  $\tilde{\alpha}_{ij}^{kl} = \begin{cases} \alpha_{ij}^{kk}, & \text{if } k = l; \\ 0, & \text{otherwise.} \end{cases}$

*Proof.* we use the basic definition of determinant. For  $r \times r$  minors on the columns of  $\sum_{j=0}^{i-1} (1 + n_j)$ ,  $i = 1, \dots, r, n_0 = 0$ , only elements in diagonal positions can be taken to obtain nonzero minors. In finding the corresponding co-minor, after crossing out the above rows and columns,

we are in the same situation again have to taken certain elements in diagonal positions only. By repeating this process, we can see that only elements in diagonal positions can possibly contribute to the determinant of  $\alpha$ , and other elements could not affect  $\det \alpha$  at all. We would get same determinant if we set those elements to be zero, i.e.,  $\det \alpha = \det(\tilde{\alpha})$ .  $\square$

$\tilde{\alpha}$  is partitioned into a lower triangular block matrix according to sizes of Jordan blocks  $J_i$ : assume the diagonal block of  $\tilde{\alpha}$  are  $\tilde{\alpha}_1, \dots, \tilde{\alpha}_s$ . Now we use number theory to describe our result. Let  $n = 2m \in 2\mathbb{Z}^+$  and the partition of  $m$  determined by sizes of  $J_i$  is given by

$$m = \beta_1 n_1 + \dots + \beta_s n_s.$$

where  $n_i \geq n_j$  if  $i \geq j$ .

So  $\tilde{\alpha}_i$  is the submatrix of  $\tilde{\alpha}$  formed by  $\beta_i \times \beta_i$  blocks of size  $n_i \times n_i$  in  $\tilde{\alpha}$ . Let  $\gamma_i$  be the determinant of the matrix formed from all the (1,1)-entry in each block of  $\tilde{\alpha}_i$ , i.e.,

$$\begin{aligned} U_i &= ((\tilde{\alpha}_i)_{kl}^{11}), \\ \gamma_i &= \det U_i. \end{aligned}$$

It's obvious that  $\det(\tilde{\alpha}_i) = \gamma_i^{\{n_i\}}$ , so  $\det \alpha = \det \tilde{\alpha} = \prod_{i=1}^s \gamma_i^{\{n_i\}}$ . From (6.3.8), we have

$$\begin{aligned} \det \alpha &= \frac{q}{\tilde{q}} (a + b\sqrt{p})^{-1} \\ &= \prod_{i=1}^s \gamma_i^{\{n_i\}} \end{aligned}$$

From Lemma 6.3.3 we can see that there exist a solution of  $\gamma_i$  of the above equation if and only  $\frac{q}{\tilde{q}}(a + b\sqrt{p})^{-1} \in (\mathbb{F}^*(\sqrt{p}))^{\{l\}}$  or  $\frac{q}{\tilde{q}} \in (a + b\sqrt{p}) * (\mathbb{F}^*(\sqrt{p}))^{\{l\}} \subseteq S^1 * (\mathbb{F}^*(\sqrt{p}))^{\{l\}}$ , where  $l$  is the greatest common factor of  $n_i, i = 1, \dots, s$ . So for the same types of Jordan blocks, we can take elements  $q \in \mathbb{F}^*(\sqrt{p}) / (S^1 * (\mathbb{F}^*(\sqrt{p}))^{\{l\}})$  in  $J$  as representatives of  $K_{\mathbb{F}}$ -orbits.

Summing up the above analysis, we get following result:

**Theorem 6.3.8.** *If we study  $(G_{\mathbb{F}}, \sigma)$  instead of  $(\tilde{G}_{\mathbb{F}}, \tilde{\theta})$ , then we can choose elements  $g = I + J$  as representatives of  $K_{\mathbb{F}}$ -orbits, where*

$$J = \begin{bmatrix} & \overline{D} \\ D & \end{bmatrix} \text{ with } D = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{bmatrix},$$

and  $J_i, i = 1, \dots, r$ , satisfy following conditions

1. The sizes of  $J_i$  are decreasing, especially if  $J_{i_0} = 0$ , then  $J_{i_0+1} = \dots = J_r = 0$ ;
2. If  $J_i \neq 0$ , then

$$J_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & q \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

where  $q = 1$  if  $1 \leq i \leq r - 1$ . If  $i = r$  ( $D$  has no zero Jordan block in this case), then  $q$  is representative element of  $\mathbb{F}^*(\sqrt{p})/(\mathbb{F}^*(\sqrt{p}))^{\{l\}}$  with  $\{\}$  being the "curly bracket" power function defined in Definition 6.3.1, and  $l$  is the greatest common factor of  $n_i$ ,  $n_i$  are from the partition of positive integer  $m = \frac{n}{2}$ :

$$m = \beta_1 n_1 + \dots + \beta_s n_s.$$

which corresponds to the Jordan blocks  $J_i$  of  $D$ .

*Remark 6.3.9.* Again, we are studying a JCM-like problem. Instead of studying  $D \rightarrow x^{-1}Dx$  directly, a conjugation operator is involved:  $D \rightarrow (\overline{x})^{-1}Dx$ . We get something similar canonical form too, and the classification is similar to its corresponding part in Theorem 5.2.1.

*Remark 6.3.10.* The classification problem is more or less same as the canonical-form problem. For inner involutions, the classifications are similar to Jordan decomposition, and tools we use to reduce matrices or matrix pairs/pencils are elementary matrices, i.e., adding/subtracting certain rows and/or columns to other rows and/or columns. We have the Jordan canonical form to use in reduction. For outer involutions, it's completely different. From the viewpoint of matrix theory, we use orthogonal matrices instead of elementary matrices to reduce symmetric matrices to simple/special/canonical forms.

## Chapter 7

# Cartan/Outer Involution

The key involution for outer involutions is Cartan involution and the model for the classification problem in this case would be the one for Cartan involution over algebraically closed fields, though the problem could be simple, for example, we have the identity orbit for the real field  $\mathbb{R}$ , or for any other field that satisfies

$$\sum_{i=1}^n x_i^2 = 0 \Rightarrow x_i = 0, i = 1, \dots, n, \text{ for } x_i \in \mathbb{F}.$$

We treat the cases of small  $n = 2, 3$  and  $4$ . We hope to complete the general case in future research

The canonical form here corresponds to the JCM exactly and the main difference is that we use here orthogonal matrices instead of elementary matrices.

To be explicit, Cartan involution and corresponding concepts are defined as following

$$\begin{aligned}\sigma(g) &= g^{-T}, \\ \tau(g) &= g(\sigma(g))^{-1} = gg^T, \\ K &= \{k \in G | \sigma(k) = k^{-T} = k \Leftrightarrow kk^T = I\},\end{aligned}$$

where  $g^T$  is the transpose of  $g$ .

We work on algebraically closed field only, i.e.,  $\mathbb{F} = \overline{\mathbb{F}}$ .  $G = \text{SL}(n, \mathbb{F})$ . It's obvious that  $g = I + J \in P = G/K = \{gg^T \in G\}$  is unipotent  $\Rightarrow J = J^T$  is nilpotent.

Same as before we will study  $J$  instead. We first redo the case  $n = 2$  and develop some basic properties we need for the general case. We then illustrate how to get the canonical form (here it's tri-diagonal matrices) for the general case, and give results for  $n = 3$  and  $4$  as examples.

## 7.1 Redo the case $n = 2$

First let's redo the case of  $n = 2$ .

$J = J^T = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$  is nilpotent if and only if  $J^2 = 0$ , so

$$\begin{aligned} J^2 &= \begin{bmatrix} a & b \\ b & d \end{bmatrix} \cdot \begin{bmatrix} a & b \\ b & d \end{bmatrix} \\ &= \begin{bmatrix} a^2 + b^2 & b(a + d) \\ b(a + d) & b^2 + d^2 \end{bmatrix} = 0. \end{aligned}$$

It's clear that  $b = 0 \Leftrightarrow a = d = 0$ .

If  $b \neq 0$ , then we have  $a + d = 0$ , or  $d = -a$ , and we therefore have  $a = b\sqrt{-1}$  and

$$J = \begin{bmatrix} b\sqrt{-1} & b \\ b & -b\sqrt{-1} \end{bmatrix} = b \begin{bmatrix} \sqrt{-1} & 1 \\ 1 & -\sqrt{-1} \end{bmatrix}.$$

When  $n = 2$ ,

$$K = \left\{ k = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \in G \mid c^2 + s^2 = 1 \right\}.$$

Therefore

$$\begin{aligned} k^{-1} J k &= \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \cdot b \begin{bmatrix} \sqrt{-1} & 1 \\ 1 & -\sqrt{-1} \end{bmatrix} \cdot \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= b \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} \sqrt{-1} & 1 \\ 1 & -\sqrt{-1} \end{bmatrix} \cdot \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= b \begin{bmatrix} c\sqrt{-1} - s & c + s\sqrt{-1} \\ s\sqrt{-1} + c & s - c\sqrt{-1} \end{bmatrix} \cdot \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\ &= b \begin{bmatrix} c^2\sqrt{-1} - 2sc - s^2\sqrt{-1} & -s^2 + 2sc\sqrt{-1} + c^2 \\ c^2 + 2cs\sqrt{-1} - s^2 & s^2\sqrt{-1} + 2sc - c^2\sqrt{-1} \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
&= b \begin{bmatrix} (c + s\sqrt{-1})^2\sqrt{-1} & (c + s\sqrt{-1})^2 \\ (c + s\sqrt{-1})^2 & (c + s\sqrt{-1})^2(-\sqrt{-1}) \end{bmatrix} \\
&= b(c + s\sqrt{-1})^2 \begin{bmatrix} \sqrt{-1} & 1 \\ 1 & -\sqrt{-1} \end{bmatrix}
\end{aligned}$$

If  $b \neq 0$ , by choosing  $c$  and  $s$  proper we can make  $b = \pm\sqrt{-1}$  because following equations always have a solution:

$$\begin{aligned}
\frac{\pm\sqrt{-1}}{b} &= (c + s\sqrt{-1})^2 \\
1 &= c^2 + s^2 = c^2 - (-s^2) = (c + s\sqrt{-1})(c - s\sqrt{-1}).
\end{aligned}$$

Notice that  $\sqrt{-1}$  and  $-\sqrt{-1}$  are two different elements, so we will have actually two different  $K$ -orbits  $\begin{bmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & -1 \end{bmatrix}$  and  $\begin{bmatrix} 1 & -\sqrt{-1} \\ -\sqrt{-1} & -1 \end{bmatrix}$ . What's more, given  $\sqrt{-1}$  or  $-\sqrt{-1}$ , we can always find a lower triangular matrix  $L$  for the above  $J$  such that  $LL^T = I + J$ . For example, for  $g = I + J = \begin{bmatrix} 2 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{bmatrix}$  we have

$$L = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{-1}}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } LL^T = g.$$

So beside the identity  $K$ -orbit, we have another two  $K$ -orbits of unipotent elements in  $P = G/K$ :  $\begin{bmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & -1 \end{bmatrix}$  and  $\begin{bmatrix} 1 & -\sqrt{-1} \\ -\sqrt{-1} & -1 \end{bmatrix}$ , which coincides with the result we get from inner involution case: they correspond to the Jordan canonical form (JCF)  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ .

Before we go the higher case of  $n$ , we show "the other side of the above calculation" is true:

**Lemma 7.1.1.**  $g = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$  with  $b \neq 0$  cannot be diagonalized by rotation matrices if and only if  $d - a = \pm 2\sqrt{-1}b$ .

*Proof.* Let  $k = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \in K$ , then

$$\begin{aligned}
k^{-1}gk &= \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} a & b \\ b & d \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\
&= \begin{bmatrix} ac - sb & bc - sd \\ as + bc & bs + cd \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \\
&= \begin{bmatrix} ac^2 - 2sbc + s^2d & acs + b(c^2 - s^2) - cds \\ asc + b(c^2 - s^2) - cds & as^2 + 2sbc + c^2d \end{bmatrix}
\end{aligned}$$

Let's set

$$\begin{cases} (a - d)cs + b(c^2 - s^2) = 0 \\ c^2 + s^2 = (c + \sqrt{-1}s)(c - \sqrt{-1}s) = 1 \end{cases} \quad (7.1.1)$$

Let  $u = c + \sqrt{-1}s, v = c - \sqrt{-1}s$ , then we have

$$\begin{aligned}
c &= \frac{u + v}{2} \\
s &= \frac{u - v}{2} \\
cs &= \frac{u^2 - v^2}{4\sqrt{-1}} \\
c^2 - s^2 &= \frac{u^2 + v^2}{2}
\end{aligned}$$

Plugging the above expressions into the first equation of (7.1.1) we get

$$\begin{aligned}
(a - d) \frac{u^2 - v^2}{4\sqrt{-1}} + b \frac{u^2 + v^2}{2} &= 0 \\
\frac{(a - d)}{b} \cdot \frac{u^2 - v^2}{4\sqrt{-1}} + \frac{u^2 + v^2}{2} &= 0 \\
\frac{(a - d)}{b} \cdot (u^2 - v^2) + 2\sqrt{-1}(u^2 + v^2) &= 0 \\
(2\sqrt{-1} + \frac{(a - d)}{b})u^2 + (2\sqrt{-1} - \frac{(a - d)}{b})v^2 &= 0
\end{aligned}$$

The last equation does not have solution if and only  $(2\sqrt{-1} + \frac{(a-d)}{b}) = 0$  or  $(2\sqrt{-1} - \frac{(a-d)}{b}) = 0$ , which gives us the desired relation of  $a, b, d$ .

□



## 7.2 Reduction

Let's look at the case of  $n \geq 3$ . First we have following lemma.

**Lemma 7.2.1.** *Any orthogonal matrix  $k \in K$  can be written as a product of permutation matrices, rotation matrices and diagonal matrices with  $\pm 1$  as diagonals.*

*Proof.* Let  $k = (k_{ij})$ . If we have  $q \geq 2$  nonzero elements in the first column of  $k$ , using permutation matrices we can make the first  $q$  elements in the first column of  $k$  to be nonzero. By choosing  $c$  and  $s$  carefully in

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} r \\ t \end{bmatrix} = \begin{bmatrix} cr + st \\ ct - sr \end{bmatrix} = \begin{bmatrix} \tilde{r} \\ \tilde{t} \end{bmatrix}$$

we can make at least  $\tilde{t} = 0$  (or possibly both of them are zeros). This means that we eliminate at least one nonzero elements by one rotation matrix. By repeating this process we can make the first column of  $k$  to have only one nonzero element. It's obvious that this nonzero elements must be  $\pm 1$ . Now we can re-write

$$k = \begin{bmatrix} \pm 1 & v \\ 0 & u \end{bmatrix}.$$

where  $u$  is nonsingular.

$$\begin{aligned} I &= k k^t \\ &= \begin{bmatrix} \pm 1 & v \\ 0 & u \end{bmatrix} \cdot \begin{bmatrix} \pm 1 & 0 \\ v^T & u^T \end{bmatrix} \\ &= \begin{bmatrix} 1 + vv^T & vu^T \\ uv^T & uu^T \end{bmatrix} \\ \Rightarrow uv^T &= 0 \quad \& \quad uu^T = I \\ \Rightarrow v &= 0 \end{aligned}$$

The proof is finished by induction. □

Now we claim that we only need to study tri-diagonal matrices. Here is a quick why. We are considering matrices in  $P = G/K$  consisting of symmetric matrices, so we assume naturally

$g = g^T = (g_{ij})$ . If the first column/row is zero, we then reduce to the case of  $n - 1$ . If there are more than two nonzero elements in the first column/row, using permutation matrices we can make the all nonzero elements in the first column "packed up" to the left top corner, and same for the first row since permutation matrices act on  $J$  from both sides. By working on  $\begin{bmatrix} g_{11} & g_{1i} \\ g_{i1} & g_{ii} \end{bmatrix}$  we can use rotation matrices to make  $g_{i1} = g_{1i} = 0, i \geq 2$  unless condition in Lemma 7.1.1 is satisfied. If there are some nonzero elements cannot be made into zeroes, i.e., condition in Lemma 7.1.1 is satisfied, we simply use  $\begin{bmatrix} g_{21} \\ g_{31} \end{bmatrix}$  to be  $\begin{bmatrix} r \\ t \end{bmatrix}$  in Lemma 7.2.1 and we can therefore use rotation matrix to make  $g_{31} = 0$ . Treat  $\begin{bmatrix} g_{21} \\ g_{41} \end{bmatrix}$  to be new  $\begin{bmatrix} r \\ t \end{bmatrix}$  and we can then make  $g_{41} = 0$ . Repeat this until  $g_{11}$  and  $g_{21}/g_{12}$  are the only nonzero elements in the first column/row. Use permutation matrices to make  $g_{32}$  to be nonzero if there is at least one nonzero element in  $g_{i2}, i \geq 3$ . In the same way we can clear up all nonzero elements following  $g_{32}$  in the second column/row. As consequence we get a tri-diagonal matrix. See following transformations on a  $5 \times 5$  matrix as illustration of the above process.

$$\begin{array}{ccccccc}
\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \rightarrow & \begin{bmatrix} * & * & 0 & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \rightarrow & \begin{bmatrix} * & * & 0 & 0 & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \rightarrow & \begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{bmatrix} \rightarrow \\
\begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & * & 0 & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & * & * & * & * \end{bmatrix} & \rightarrow & \begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{bmatrix} & \rightarrow & \begin{bmatrix} * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & * & * & * & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix} .
\end{array}$$

Also, we here assume  $y_i \neq 0, i = 1, \dots, n - 1$  since if one of them is zero, the problem has been reduced to the one of smaller size.

Let  $g = I + J = (g_{ij})$ .  $g = g^T$  is tri-diagonal and unipotent  $\Leftrightarrow J = J^T$  is tri-diagonal and unipotent. We classify  $J$ 's instead of  $g$ 's.  $J$  is nilpotent if and only if  $\det(\lambda I - J) = \lambda^n$  or equivalently  $\det(\lambda I + J) = \lambda^n$ . The reason that we use "plus" operation instead in the characteristic polynomial is that we like "plus" signs more in our calculations. We assume from now on

$$J = \begin{bmatrix} x_1 & y_1 & & & \\ y_1 & x_2 & y_2 & & \\ & y_2 & \ddots & \ddots & \\ & & \ddots & x_{n-1} & y_{n-1} \\ & & & y_{n-1} & x_n \end{bmatrix}.$$

### 7.3 Ideas

If there exists nilpotent matrix  $J$  with  $g = I + J \in P$ , we then reduce the problem (or find the canonical form of  $J$  or  $g$ ) in following steps:

1. We illustrate that we may assume diagonals of  $J$  appear in a pattern of  $\alpha, -\alpha, \alpha, -\alpha, \dots$ . Therefore we only need to classify  $J$  or  $g$  with alternating-pattern diagonals. Especially if  $n$  is odd,  $\alpha = 0$  here.  $\alpha \neq 0$  if  $n$  is even, and we can make  $\alpha$  into any nonzero number in this case. We'd like to make  $\alpha = 1$ . Also if  $n$  is even, only even terms, i.e., terms of even degree, appear in  $\det(\lambda I + J)$ . If  $n$  is odd, only odd terms (again, terms of odd degree) appear in  $\det(\lambda I + J)$ , and we don't have constant term automatically.

2. Assuming step 1, we can find solutions for equation  $\det(\lambda I + J) = \lambda^n$ . There are more than one solution for  $\alpha, y_i, i = 1, \dots, n-1$ . Given any two different specific solutions, say  $\alpha^1, y_i^1, i = 1, \dots, n-1$  and  $\alpha^2, y_i^2, i = 1, \dots, n-1$ , by apply rotation matrices on  $J$  in special patterns, we can then make  $\alpha^1 = \alpha^2$  and  $y_i^1 = y_i^2, i = 1, \dots, y_{n-1}$ .

3. Assume  $n = 2m$  is an even integer. Given  $\alpha$  from the first step we have  $n-1$  unknowns,  $y_i, i = 1, \dots, n-1$ . From  $\det(\lambda I + J) = \lambda^n$  we can get  $m$  equations and from step 2 will have  $m-1$  unknowns, so we should have only finitely many different canonical forms. Two different canonical forms are expected here since the canonical forms here correspond to JCFs of largest rank  $n-1$ , and we have two JCFs in  $\text{SL}(n, \mathbb{F})$  of rank  $n-1$  (we need an odd number of permutations to get the upper triangular JCF from the lower JCF, which gives the determinant -1). The correspondence is also true for the case that  $n = 2m+1$  is odd, and it's expected to have one canonical forms only (similarly we need even number permutations to change lower JCF to upper JCF).

I use an  $8 \times 8$  matrix as our example, and we will do the second step first, and then the first one.

For the second step, here is the way we can work on  $J$  while keeping  $J$  with same diagonals. First we can multiply a rotation matrix on the first and third rows and columns. This would not change diagonals since we have  $\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} = \alpha I$  as our submatrix from the first and third

rows and columns. This will introduce a nonzero element in (4,1) position. To keep resulting matrix tri-diagonal, we just need to clear it up use a series of rotations. We call this one as Case (1,3) since, in order to get a scalar submatrix we choose the first and third rows and columns. We introduce one free variable in this way, but we get no more free INDEPENDENT variable by introducing different rotations. Luckily in the similar way we can choose the first and fifth rows and columns (Case (1,5)), and the first and seventh rows and columns (Case (1,7)), and clear up those nonzero elements introduced by rotations. It's also true that in the way of using different combinations to get scalar submatrix  $\alpha I$  we can introduce  $m - 1$  independent free variables, which allows us to make  $y_i, i = 1, \dots, n - 1$  into exactly the solution(s) we want. See following manipulations as what we mean.

$\alpha$	$y_1$	0	$0^1$	0	0	0	0
$y_1$	$-\alpha$	$y_2$	0	$0^2$	0	0	0
0	$y_2$	$\alpha$	$y_3$	0	$0^3$	0	0
$0^1$	0	$y_3$	$-\alpha$	$y_4$	0	$0^4$	0
0	$0^2$	0	$y_4$	$\alpha$	$y_5$	0	$0^5$
0	0	$0^3$	0	$y_5$	$-\alpha$	$y_6$	0
0	0	0	$0^4$	0	$y_6$	$\alpha$	$y_7$
0	0	0	0	$0^5$	0	$y_7$	$-\alpha$

Case (1,3). By multiplying a rotation matrix on the first and third rows and columns, we can change  $y_1$  and  $y_2$ . At same time we introduce nonzero elements in (4,1) and (1,4) positions, marked as  $0^1$ 's in the above matrix. To keep the matrix in the tri-diagonal form we need to make  $0^1$  into zero, so we multiply another rotation matrix on the second and fourth rows and columns. We introduced nonzero elements in (5,2) and (2,5) positions, marked as  $0^2$  in the matrix. Repeating this process we will get  $0^3, 0^4, 0^5$  before we get a tri-diagonal form again. During this process, diagonals,  $\alpha, -\alpha, \alpha - \alpha, \dots$ , do not change, and we only change  $y_i, i = 1, \dots, 8$ . We actually introduce one free variable.

Precisely we can express the above process in following calculations: (we write only involved elements)

$$\begin{aligned}
y_i^1 &= y_i, i = 1, \dots, 7 \\
\begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} y_1^1 & y_2^1 \\ 0 & y_3^1 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_1 & * \\ 0 & * \end{bmatrix} \\
y_2^2 &= s_1 y_1^1 + c_1 y_2^1 \\
y_3^2 &= c_1 y_3^1 \\
\begin{bmatrix} c_3 & -s_3 \\ s_3 & c_3 \end{bmatrix} \begin{bmatrix} y_2^2 & y_3^2 \\ 0 & y_4^1 \end{bmatrix} \begin{bmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_2 & * \\ 0 & * \end{bmatrix} \\
y_3^3 &= s_2 y_2^1 + c_2 y_3^2 \\
y_4^2 &= c_2 y_4^1 \\
\begin{bmatrix} c_4 & -s_4 \\ s_4 & c_4 \end{bmatrix} \begin{bmatrix} y_3^3 & y_4^2 \\ 0 & y_5^1 \end{bmatrix} \begin{bmatrix} c_3 & s_3 \\ -s_3 & c_3 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_3 & * \\ 0 & * \end{bmatrix} \\
y_4^3 &= s_3 y_3^2 + c_3 y_4^2 \\
y_5^2 &= c_3 y_5^1 \\
\begin{bmatrix} c_5 & -s_5 \\ s_5 & c_5 \end{bmatrix} \begin{bmatrix} y_4^3 & y_5^2 \\ 0 & y_6^1 \end{bmatrix} \begin{bmatrix} c_4 & s_4 \\ -s_4 & c_4 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_4 & * \\ 0 & * \end{bmatrix} \\
y_5^3 &= s_4 y_4^2 + c_4 y_5^2 \\
y_6^2 &= c_4 y_6^1 \\
\begin{bmatrix} c_6 & -s_6 \\ s_6 & c_6 \end{bmatrix} \begin{bmatrix} y_5^3 & y_6^2 \\ 0 & y_7^1 \end{bmatrix} \begin{bmatrix} c_5 & s_5 \\ -s_5 & c_5 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_5 & * \\ 0 & * \end{bmatrix}
\end{aligned}$$

$c_1$  is free and  $c_i, i = 2, 3, 4, 5, 6$  are determined in a way that makes the resulting matrices

are upper tri-anglular matrices. The resulting  $\tilde{y}_1$  is a function of  $y_1, y_2$  and  $y_3, y_2$  a function of  $y_1, y_2, y_3$  and  $y_4$ , etc.

Generally, we can express calculations above to get the new  $\tilde{y}_i, i = 1, 2, 3, 4, 5$  in following equations:

$$\begin{aligned} \begin{bmatrix} u & -v \\ v & u \end{bmatrix} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} &= \begin{bmatrix} \tilde{a} & * \\ 0 & * \end{bmatrix} \\ u(ac - sb) + vsd &= \tilde{a} \\ v(ac - sb) - usd &= 0 \\ u^2 + v^2 &= 1 \\ c^2 + s^2 &= 1 \\ \Rightarrow \sqrt{(ac - sb)^2 + (sd)^2} &= \tilde{a} \end{aligned}$$

We can repeat the above process using different  $\bar{c}$  but it gives us nothing new but a combination  $c\bar{c}$  and we still have only one free variable. Noticing the first, third, fifth and seventh diagonals are all  $\alpha$ , we would try the combinations of (1,5), (1,7).

Case (1,5). If we try (1,5), following shows the process we might have

$\alpha$	$y_1$	0	$0^1$	0	$0^*$	0	0
$y_1$	$-\alpha$	$y_2$	0	$0_{**}^\vee$	0	$0^{**}$	0
0	$y_2$	$\alpha$	$y_3$	0	$0_{***}^{**}$	0	$0^{***}$
$0^1$	0	$y_3$	$-\alpha$	$y_4$	0	$0_{****}^{***}$	0
0	$0_{**}^\vee$	0	$y_4$	$\alpha$	$y_5$	0	$0^{****}$
$0^*$	0	$0_{***}^{**}$	0	$y_5$	$-\alpha$	$y_6$	0
0	$0^{**}$	0	$0_{****}^{***}$	0	$y_6$	$\alpha$	$y_7$
0	0	$0^{***}$	0	$0_{****}^{***}$	0	$y_7$	$-\alpha$

Similarly to previous case, when we multiply a rotation matrix to the first and fifth rows and columns, we introduce nonzero elements in (4,1), (5,2) and (6,1) positions, marked by " $^1, \vee, *$ ". We first try to clear the nonzero element in (6,1). This would introduce new elements in (6,3) and (7,2), and it also affects (5,2) entry. We mark this step using "\*\*\*". We still try to clear the "outside" nonzero element(s) first. Using rotation matrix to make (7,2) entry into zero, we get nonzero entries in (7,4) and (8,3), and affect (6,3) entry, marked by "\*\*\*\*". When we clear (8,3) entry, we introduce nonzero entry in (8,5) and affect (7,4) entry, marked by "\*\*\*\*\*". No "outside" entry is introduced since we reach the bottom of the matrix. Now using the method we introduce in the above case, we can clear all nonzero elements in the third off-diagonals (i.e., (4,1), (5,2), (6,3), (7,4) and (8,5) entries).

The calculations in clearing the fifth off-diagonals are as following:

$$\begin{aligned} \begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} y_1^1 & 0 \\ 0 & y_5^1 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_1 & * \\ 0 & * \end{bmatrix} \\ \begin{bmatrix} c_3 & -s_3 \\ s_3 & c_3 \end{bmatrix} \begin{bmatrix} y_2^2 & 0 \\ 0 & y_6^1 \end{bmatrix} \begin{bmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_2 & * \\ 0 & * \end{bmatrix} \\ \begin{bmatrix} c_4 & -s_4 \\ s_4 & c_4 \end{bmatrix} \begin{bmatrix} y_3^3 & 0 \\ 0 & y_7^1 \end{bmatrix} \begin{bmatrix} c_3 & s_3 \\ -s_3 & c_3 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_3 & * \\ 0 & * \end{bmatrix} \end{aligned}$$

$\tilde{y}_1$  is not the final  $y_1$ . The final one, denoted by  $\tilde{\tilde{y}}_1$ , is a function of  $y_1, y_2, y_3$  and  $y_5, \tilde{\tilde{y}}_2$  a function of  $y_1, y_2, y_3, y_4, y_5$  and  $y_6$ , etc.

Though it shows the idea, it is not representative, since if the order of the matrix is bigger, and when we use entries  $y_5, y_6$  and  $y_7$  again, they have been modified.

We introduce a new free variable in this case, and it's clear that this variable cannot be generated by working on Case (1,3).

Case (1,7). Similarly we can work on the first and seventh rows and columns.



$\alpha$	$y_1$	0	0	0	$0^1$	0	$0^*$
$y_1$	$-\alpha$	$y_2$	0	0	0	$0_{**}^\vee$	0
0	$y_2$	$\alpha$	$y_3$	0	0	0	$0^{**}$
0	0	$y_3$	$-\alpha$	$y_4$	0	0	0
0	0	0	$y_4$	$\alpha$	$y_5$	0	0
$0^1$	0	0	0	$y_5$	$-\alpha$	$y_6$	0
0	$0_{**}^\vee$	0	0	0	$y_6$	$\alpha$	$y_7$
$0^*$	0	$0^{**}$	0	0	0	$y_7$	$-\alpha$

We will first produce nonzero elements in the slots of (6,1), (7,2) and (8,1). We clear (8,1) first and this will give us a nonzero element in (8,3) and affect (7,2) entry. Now we reduce the situation similar to Case (1,5). The free variable introduced in this case is independent of the previous cases since  $y_7$  would definitely be involved in the expressions of final  $y_i, i = 1, \dots, 7$ .

*Remark 7.3.1.* 1. We choose/start to work on  $\alpha$  parts here. We can start work on  $-\alpha$  parts, but when we clear up unneeded nonzero elements, we would go to one of the above three cases, so it's equivalent to start from  $\alpha$  parts.

2. We choose to clear farthest off-diagonals first because during clearing up nonzero entries in the line we only introduce/affect entries in nearer off-diagonals, and we can therefore do it in an inductive way.

3. In clearing each nonzero elements, there are always usually and at most three entries involved and they always keep in a shape like " $\triangleright$ ". Only odd-numbered off-diagonals are involved in all calculations.

We introduce three independent free variables into the system. We have 4 equations from nilpotency ( $\det(\lambda I + J) = \lambda^n$ ). If  $\alpha$  is determined, say  $\alpha = 1$ , we can almost uniquely determine the values of  $y_i, i = 1, \dots, n - 1 = 7$ , which actually give us the canonical forms of  $J$ . They correspond to JCF of rank 7.

Now we consider how to make our diagonals into alternating pattern. Suppose we have an 8 by 8 matrix.

$x_1$	$y_1$	$0^1$	0	0	0	0	0
$y_1$	$x_2$	$y_2$	$0^2$	0	0	0	0
$0^1$	$y_2$	$x_3$	$y_3$	$0^3$	0	0	0
0	$0^2$	$y_3$	$x_4$	$y_4$	$0^4$	0	0
0	0	$0^3$	$y_4$	$x_5$	$y_5$	$0^5$	0
0	0	0	$0^4$	$y_5$	$x_6$	$y_6$	$0^6$
0	0	0	0	$0^5$	$y_6$	$x_7$	$y_7$
0	0	0	0	0	$0^6$	$y_7$	$x_8$

We multiply a rotation matrix on the first and second rows and columns, and we will get a nonzero entry in (3,1), marked by  $0^1$ . We want to keep it in tri-diagonal shape, so we use the new  $y_1$  to clear it (the reason that we don't use the new  $x_1$  is that, if we did so, we would introduce more nonzero entries in the first row and column). We get another nonzero entry  $0^2$  in (4,2), and if we keep going on, we would have  $0^3$  in (5,3) before we go back to the tri-diagonal form.

The calculations are like Case (1,3):

$$\begin{aligned}
y_i^1 &= y_i, i = 1, \dots, 7 \\
x_i^1 &= x_i, i = 1, \dots, 8 \\
\begin{bmatrix} c_1 & -s_1 \\ s_1 & c_1 \end{bmatrix} \begin{bmatrix} x_1^1 & y_1^1 \\ y_1^1 & x_2^1 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} &= \begin{bmatrix} \tilde{x}_1 & y_1^2 \\ y_1^2 & x_2^2 \end{bmatrix} \\
\begin{bmatrix} 0 & y_2^1 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} &= \begin{bmatrix} * & y_2^2 \end{bmatrix} \\
\begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} y_1^1 & x_2^1 \\ 0 & y_2^1 \end{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_1 & * \\ 0 & * \end{bmatrix} \\
\begin{bmatrix} c_2 & -s_2 \\ s_2 & c_2 \end{bmatrix} \begin{bmatrix} x_2^2 & y_2^2 \\ y_2^2 & x_3^1 \end{bmatrix} \begin{bmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{bmatrix} &= \begin{bmatrix} \tilde{x}_2 & y_2^2 \\ y_2^2 & x_3^2 \end{bmatrix} \\
\begin{bmatrix} 0 & y_3^1 \end{bmatrix} \begin{bmatrix} c_2 & s_2 \\ -s_2 & c_2 \end{bmatrix} &= \begin{bmatrix} z_3 & y_3^2 \end{bmatrix} \\
\begin{bmatrix} c_3 & -s_3 \\ s_3 & c_3 \end{bmatrix} \begin{bmatrix} y_2^2 & x_3^2 \\ z_3 & y_3^2 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_2 & * \\ 0 & * \end{bmatrix} \\
\begin{bmatrix} c_3 & -s_3 \\ s_3 & c_3 \end{bmatrix} \begin{bmatrix} x_3^2 & y_3^2 \\ y_3^2 & x_4^1 \end{bmatrix} \begin{bmatrix} c_3 & s_3 \\ -s_3 & c_3 \end{bmatrix} &= \begin{bmatrix} \tilde{x}_3 & y_3^3 \\ y_3^3 & x_4^2 \end{bmatrix} \\
\begin{bmatrix} 0 & y_4^1 \end{bmatrix} \begin{bmatrix} c_3 & s_3 \\ -s_3 & c_3 \end{bmatrix} &= \begin{bmatrix} z_4 & y_4^2 \end{bmatrix} \\
\begin{bmatrix} c_4 & -s_4 \\ s_4 & c_4 \end{bmatrix} \begin{bmatrix} y_3^3 & x_4^2 \\ z_4 & y_4^2 \end{bmatrix} &= \begin{bmatrix} \tilde{y}_3 & * \\ 0 & * \end{bmatrix} \\
\begin{bmatrix} c_4 & -s_4 \\ s_4 & c_4 \end{bmatrix} \begin{bmatrix} x_4^2 & y_4^2 \\ y_4^2 & x_5^1 \end{bmatrix} \begin{bmatrix} c_4 & s_4 \\ -s_4 & c_4 \end{bmatrix} &= \begin{bmatrix} \tilde{x}_4 & y_4^3 \\ y_4^3 & x_5^2 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
& \begin{bmatrix} 0 & y_5^1 \end{bmatrix} \begin{bmatrix} c_4 & s_4 \\ -s_4 & c_4 \end{bmatrix} = \begin{bmatrix} z_5 & y_5^2 \end{bmatrix} \\
& \begin{bmatrix} c_5 & -s_5 \\ s_5 & c_5 \end{bmatrix} \begin{bmatrix} y_4^3 & x_5^2 \\ z_5 & y_5^2 \end{bmatrix} = \begin{bmatrix} \tilde{y}_4 & * \\ 0 & * \end{bmatrix} \\
& \begin{bmatrix} c_5 & -s_5 \\ s_5 & c_5 \end{bmatrix} \begin{bmatrix} x_5^2 & y_5^2 \\ y_5^2 & x_6^1 \end{bmatrix} \begin{bmatrix} c_5 & s_5 \\ -s_5 & c_5 \end{bmatrix} = \begin{bmatrix} \tilde{x}_5 & y_5^3 \\ y_5^3 & x_6^2 \end{bmatrix} \\
& \begin{bmatrix} 0 & y_6^1 \end{bmatrix} \begin{bmatrix} c_5 & s_5 \\ -s_5 & c_5 \end{bmatrix} = \begin{bmatrix} z_6 & y_6^2 \end{bmatrix} \\
& \begin{bmatrix} c_6 & -s_6 \\ s_6 & c_6 \end{bmatrix} \begin{bmatrix} y_5^3 & x_6^2 \\ z_6 & y_6^2 \end{bmatrix} = \begin{bmatrix} \tilde{y}_5 & * \\ 0 & * \end{bmatrix} \\
& \begin{bmatrix} c_6 & -s_6 \\ s_6 & c_6 \end{bmatrix} \begin{bmatrix} x_6^2 & y_6^2 \\ y_6^2 & x_7^1 \end{bmatrix} \begin{bmatrix} c_6 & s_6 \\ -s_6 & c_6 \end{bmatrix} = \begin{bmatrix} \tilde{x}_6 & y_6^3 \\ y_6^3 & x_7^2 \end{bmatrix} \\
& \begin{bmatrix} 0 & y_7^1 \end{bmatrix} \begin{bmatrix} c_6 & s_6 \\ -s_6 & c_6 \end{bmatrix} = \begin{bmatrix} z_7 & y_7^2 \end{bmatrix} \\
& \begin{bmatrix} c_7 & -s_7 \\ s_7 & c_7 \end{bmatrix} \begin{bmatrix} y_6^3 & x_7^2 \\ z_7 & y_7^2 \end{bmatrix} = \begin{bmatrix} \tilde{y}_6 & * \\ 0 & * \end{bmatrix} \\
& \begin{bmatrix} c_7 & -s_7 \\ s_7 & c_7 \end{bmatrix} \begin{bmatrix} x_7^2 & y_7^2 \\ y_7^2 & x_8^1 \end{bmatrix} \begin{bmatrix} c_7 & s_7 \\ -s_7 & c_7 \end{bmatrix} = \begin{bmatrix} \tilde{x}_7 & y_7^3 \\ y_7^3 & x_8^2 \end{bmatrix} = \begin{bmatrix} \tilde{x}_7 & \tilde{y}_7 \\ \tilde{y}_7 & \tilde{x}_8 \end{bmatrix}
\end{aligned}$$

So we introduce one free variable. Same as before we can also start from (1,3), i.e., the first and third rows and columns:

$x_1$	$y_1$	0	$0^1$	0	0	0	0
$y_1$	$x_2$	$y_2$	$0^2$	$0^2$	0	0	0
0	$y_2$	$x_3$	$y_3$	$0^3$	$0^3$	0	0
$0^1$	$0^2$	$y_3$	$x_4$	$y_4$	$0^4$	$0^4$	0
0	$0^2$	$0^3$	$y_4$	$x_5$	$y_5$	$0^5$	$0^5$
0	0	$0^3$	$0^4$	$y_5$	$x_6$	$y_6$	$0^6$
0	0	0	$0^4$	$0^5$	$y_6$	$x_7$	$y_7$
0	0	0	0	$0^5$	$0^6$	$y_7$	$x_8$

We can again reduce to the case (1,2).

Now similar to corresponding part in changing  $y_i, i = 1, \dots, n-1$ , we can introduce  $n-1$  independent free variables, which allows us to make diagonals into alternating pattern. Also from following analysis of characteristic polynomial of  $J$  with alternating diagonals, we can see that we only need to make  $x_i, i = 1, \dots, n-2$  into the alternating patten,  $\alpha, -\alpha, \alpha, -\alpha, \dots$ , and  $x_{n-1}$  and  $x_n$  will be automatically become  $\alpha$  and  $-\alpha$ .

## 7.4 Characteristic polynomial $\det(\lambda I + J)$

Now assume we have diagonals in alternating pattern, we will show that, in the characteristic polynomial  $\det(\lambda I + J)$  we would see even terms only, if  $n$  is an even integer, and odd terms only with no constant term if  $n$  is an odd integer.

Assume  $n = 2m$  is an even number. Since we are looking for the determinant of  $\lambda I + J$ , we can use elementary transformation to act on  $\lambda I + J$ , and we are able to get rid of certain row and column if the diagonal in that row and column is one and is the only nonzero element in that row.

$$\begin{aligned}
 & \begin{bmatrix} \lambda + \alpha & y_1 & & & & \\ & y_1 & \lambda - \alpha & y_2 & & \\ & & y_2 & \lambda + \alpha & y_3 & \\ & & & y_3 & \lambda - \alpha & y_4 \\ & & & & \ddots & \ddots \\ & & & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
 \rightarrow & \begin{bmatrix} \lambda + \alpha & 0 & & & & \\ y_1 & \lambda - \alpha - \frac{y_1^2}{\lambda + \alpha} & y_2 & & & \\ & y_2 & \lambda + \alpha & y_3 & & \\ & & y_3 & \lambda - \alpha & y_4 & \\ & & & \ddots & \ddots & \ddots \\ & & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
 \rightarrow & \begin{bmatrix} 1 & 0 & & & & \\ (\lambda + \alpha)y_1 & \lambda^2 - \alpha^2 - y_1^2 & (\lambda + \alpha)y_2 & & & \\ & y_2 & \lambda + \alpha & y_3 & & \\ & & y_3 & \lambda - \alpha & y_4 & \\ & & & \ddots & \ddots & \ddots \\ & & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & & y_{n-1} & \lambda - \beta \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
& \rightarrow \begin{bmatrix} \lambda^2 - \alpha^2 - y_1^2 & (\lambda + \alpha)y_2 & & & & \\ & y_2 & \lambda + \alpha & y_3 & & \\ & & y_3 & \lambda - \alpha & y_4 & \\ & & & \ddots & \ddots & \ddots \\ & & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
& \rightarrow \begin{bmatrix} \lambda^2 - \alpha^2 - y_1^2 & 0 & & & & \\ & y_2 & \lambda + \alpha - \frac{(\lambda + \alpha)y_2^2}{\lambda^2 - \alpha^2 - y_1^2} & y_3 & & \\ & & y_3 & \lambda - \alpha & y_4 & \\ & & & \ddots & \ddots & \ddots \\ & & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
& \rightarrow \begin{bmatrix} 1 & 0 & & & & \\ (\lambda^2 - \alpha^2 - y_1^2)y_2 & (\lambda^2 - \alpha^2 - y_1^2 - y_2^2)(\lambda + \alpha) & (\lambda^2 - \alpha^2 - y_1^2)y_3 & & & \\ & y_3 & \lambda - \alpha & y_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
& \rightarrow \begin{bmatrix} (\lambda^2 - \alpha^2 - y_1^2 - y_2^2)(\lambda + \alpha) & (\lambda^2 - \alpha^2 - y_1^2)y_3 & & & & \\ & y_3 & \lambda - \alpha & y_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & y_{n-1} & \lambda - \beta \end{bmatrix} \\
& \rightarrow \begin{bmatrix} (\lambda^2 - \alpha^2 - y_1^2 - y_2^2)(\lambda + \alpha) & 0 & & & & \\ & y_3 & \lambda - \alpha - \frac{(\lambda^2 - \alpha^2 - y_1^2)y_3^2}{(\lambda^2 - \alpha^2 - y_1^2 - y_2^2)(\lambda + \alpha)} & y_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & y_{n-2} & \lambda + \beta & y_{n-1} \\ & & & & y_{n-1} & \lambda - \beta \end{bmatrix}
\end{aligned}$$

$$\rightarrow \begin{bmatrix} g_2(\lambda^2) & h_1(\lambda^2)(\lambda + \alpha)y_4 & & \\ & \ddots & \ddots & \ddots \\ & & y_{n-3} & \lambda - \alpha & y_{n-2} \\ & & & (\lambda - \beta)y_{n-2} & \lambda^2 - \beta^2 - y_{n-1}^2 \end{bmatrix}$$

$$\text{with } g_2(\lambda^2) = (\lambda^2 - \alpha^2 - y_1^2 - y_2^2 - y_3^2)(\lambda^2 - \alpha^2) + y_1^2 y_3^2$$

$$\text{and } h_1(\lambda^2) = ((\lambda^2 - \alpha^2 - y_1^2 - y_2^2)$$

$\rightarrow \dots$

$$\begin{aligned} &\rightarrow \begin{bmatrix} g_{m-1}(\lambda^2) & h_{m-2}(\lambda^2)(\lambda + \alpha)y_{n-2} \\ (\lambda - \beta)y_{n-2} & \lambda^2 - \beta^2 - y_{n-1}^2 \end{bmatrix} \\ &= g_{m-1}(\lambda^2)(\lambda^2 - \beta^2 - y_{n-1}^2) - h_{m-2}(\lambda^2)(\lambda + \alpha)(\lambda - \beta) \\ &= g_{m-1}(\lambda^2)(\lambda^2 - \beta^2 - y_{n-1}^2) - h_{m-2}(\lambda^2)[\lambda^2 + (\alpha - \beta)\lambda - \alpha\beta] \\ &= \lambda^{2m} \end{aligned}$$

$\Rightarrow \alpha = \beta$  and we have even terms only in characteristic polynomial.

The reason that the last two diagonals are  $\beta$  and  $-\beta$  is that  $\sum_{i=1}^{n-2} x_i = 0 \Rightarrow x_{n-1} + x_n = 0$ . We get  $x_n = -\beta$  if we assume  $x_{n-1} = \beta$ .

One more thing is that we actually have  $g_{m-1}(\lambda^2) = f_{n-2}(\lambda)$  and  $h_{m-2}(\lambda^2)(\lambda + \alpha) = f_{n-3}(\lambda)$ , where  $f_i(\lambda)$  is the minor of the first  $i$  rows and columns, which is also for the case that  $n$  is odd. We can see that  $\lambda + \alpha$  is a factor of  $f_i(\lambda)$  if  $i$  is an odd integer.  $f_i(\lambda)$  has even terms only if  $i$  is even, i.e., the degrees of  $\lambda$ -terms in  $f_i(\lambda)$  are all even numbers.

The proof for the case that  $n$  is odd is similar. We have in this case  $\alpha = 0$ , so  $f_i(\lambda)$  has even terms only if  $i$  is even, and  $f_i(\lambda) = \lambda * h_i(\lambda^2)$  if  $i$  is odd.  $n - 3$  is even and  $n - 2$  is odd.

$$\begin{aligned} \det(\lambda I + J) &= \det \begin{bmatrix} f_{n-2}(\lambda) & f_{n-3}(\lambda)y_{n-2} \\ (\lambda - \beta)y_{n-2} & \lambda^2 - \beta^2 - y_{n-1}^2 \end{bmatrix} \\ &= f_{n-2}(\lambda) * (\lambda^2 - \beta^2 - y_{n-1}^2) - f_{n-3}(\lambda)(\lambda - \beta)y_{n-2}^2 \\ &= f_{n-2}(\lambda) * (\lambda^2 - \beta^2 - y_{n-1}^2) - f_{n-3}(\lambda)\lambda y_{n-2}^2 + f_{n-3}(\lambda)\beta y_{n-2}^2 \\ &= \lambda^n \\ &\Rightarrow f_{n-3}(\lambda)\beta y_{n-2}^2 \equiv 0 \text{ (the only even terms)} \\ &\Rightarrow \beta = 0 \quad (y_{n-2} \neq 0, f_{n-3}(\lambda) = \lambda^{n-3} + \dots) \end{aligned}$$

The reason that we don't have the constant term, i.e.,  $\det J = 0$  is always true, is that,



using the recursive formula for the determinant of tri-diagonal matrix

$$f_n \lambda = (\lambda - x_n) f_{n-1}(\lambda) - y_{n-1}^2 f_{n-2}(\lambda),$$

we can see that only  $y_i^2$  will appear in the determinant. If  $n = 2m + 1$  is odd,  $\alpha = 0$ , and in the constant term, there is no  $y_i$  involved (or we would have even factors generating the constant term), i.e., the constant term for  $\det(\lambda I + J)$  is always zero, and we get no equation from the constant term. We will have only  $m - 1$  equations if  $n = 2m + 1$  (the trace equation  $\sum_{i=1}^n x_i = 0$  is always true).

*Remark 7.4.1.* We have not given proofs for the alternating diagonals and acting on off-diagonal part yet, which we believe it would be something from symmetric varieties and orthogonal groups. For small  $n$  we can compute all  $x_i$  and  $y_i$  explicitly (see following the case  $n=3$  and 4), but it's hard for us to keep going by calculating directly and would not make much sense even if we could.

In the following two sections we follow the ideas introduced above to work on the small  $n$ :  $n = 3$  and 4.

## 7.5 Case $n = 3$

Let's look at the case of  $n = 3$ . Since we cannot eliminate  $y_1$  using  $x_1$ , from Lemma 7.1.1 we can assume that  $x_1, y_1$  and  $x_2$  satisfy following equality:

$$x_2 - x_1 = \pm 2\sqrt{-1}y_1$$

or it can be written as

$$tI - \begin{bmatrix} \sqrt{-1}y_1 & y_1 \\ y_1 & -\sqrt{-1}y_1 \end{bmatrix}.$$

When we multiply a rotation matrix to the first and second rows and columns, we would get

$$tI - y_1(c + s\sqrt{-1})^2 \begin{bmatrix} \sqrt{-1}y_1 & y_1 \\ y_1 & -\sqrt{-1}y_1 \end{bmatrix}.$$

By taking certain  $c$  we can make the new  $x_1 = 0$ , noticing that following operations would

not affect  $x_1$  at all. We have  $x_2 = x_3 = 0$  immediately from our calculations in illustrating the odd/even pattern of  $\det(\lambda I + J)$ . We will have following matrix to work on

$$J = \begin{bmatrix} 0 & y_1 & 0 \\ y_1 & 0 & y_2 \\ 0 & y_2 & 0 \end{bmatrix}$$

where  $y_1^2 + y_2^2 = 0$  from  $\det(\lambda I + J) = \lambda^n = \lambda^3$ , so we can actually write it as

$$J = t \begin{bmatrix} 0 & \pm\sqrt{-1} & 0 \\ \pm\sqrt{-1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

By multiplying a diagonal matrix  $\begin{bmatrix} \pm I & \\ & 1 \end{bmatrix}$  we can make the  $\pm\sqrt{-1}$  into  $\sqrt{-1}$  or  $-\sqrt{-1}$  only. We take  $\sqrt{-1}$  here. By one permutation matrix and one diagonal matrix with only -1 on its diagonals we could have

$$J = t \begin{bmatrix} 0 & \sqrt{-1} & 0 \\ \sqrt{-1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \rightarrow t \begin{bmatrix} 0 & 0 & \sqrt{-1} \\ 0 & 0 & -1 \\ \sqrt{-1} & -1 & 0 \end{bmatrix}.$$

$t \neq 0$  can be any nonzero number in the field  $\mathbb{F}$  through following calculations:

$$\begin{bmatrix} \frac{1}{2}(t + \frac{1}{t}) & -\frac{1}{2}(t - \frac{1}{t})\sqrt{-1} \\ \frac{1}{2}(t - \frac{1}{t})\sqrt{-1} & \frac{1}{2}(t + \frac{1}{t}) \end{bmatrix} \cdot \begin{bmatrix} \sqrt{-1} \\ -1 \end{bmatrix} = \begin{bmatrix} t\sqrt{-1} \\ -t \end{bmatrix} = t \begin{bmatrix} \sqrt{-1} \\ -1 \end{bmatrix}$$

Before we claim that the above  $J$  with  $t = \sqrt{-1}$  to give us one  $K$ -orbit in  $P = G/K$ , we need to check that there is indeed some matrix  $L \in G = \mathrm{SL}(n, \mathbb{F})$  such that  $LL^T = g = I + J$ . This can be done by direct calculations

$$L = \begin{bmatrix} 1 & 0 & 0 \\ \sqrt{-1} & \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } LL^T = g = \begin{bmatrix} 1 & \sqrt{-1} & 0 \\ \sqrt{-1} & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

So for  $n = 3$ , we will have  $K$ -orbits of unipotent elements in the form of  $I + J$  with  $J$  being

one of following matrices

$J = 0$  (the identity orbit);

$$J = \begin{bmatrix} 1 & \sqrt{-1} & 0 \\ \sqrt{-1} & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{from the case } n = 2);$$

$$J = \begin{bmatrix} 0 & \sqrt{-1} & 0 \\ \sqrt{-1} & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that we can make 1 into -1 or  $\sqrt{-1}$  into  $-\sqrt{-1}$  in the third canonical form, and we can permute -1 and  $\sqrt{-1}$  too.

## 7.6 Case $n = 4$

The case  $n = 4$ .

First given  $\alpha \neq 0$  we can make  $x_1 = \alpha$  and  $x_2 = -\alpha$  using calculations we have done above. We can automatically have  $x_3 = \alpha$  and  $x_4 = -\alpha$ . Once we have alternating diagonals, we can introduce one free variable into  $y_i, i = 1, 2, 3$ . Especially we are able to make  $y_1$  satisfy  $\alpha^2 = y_1^2$ , then

$$\begin{aligned} \det(\lambda I + J) &= \det \begin{bmatrix} \lambda + \alpha & y_1 & & \\ & \lambda - \alpha & y_2 & \\ & y_2 & \lambda + \alpha & y_3 \\ & & y_3 & \lambda - \alpha \end{bmatrix} \\ &= \det \begin{bmatrix} \lambda^2 - \alpha^2 - y_1^2 & (\lambda + \alpha)y_2 \\ (\lambda - \alpha)y_2 & \lambda^2 - \alpha^2 - y_3^2 \end{bmatrix} \\ &= (\lambda^2 - \alpha^2 - y_1^2)(\lambda^2 - \alpha^2 - y_3^2) - (\lambda^2 - \alpha^2)y_2^2 \\ &= \lambda^4 - (2\alpha^2 + y_1^2 + y_2^2 + y_3^2)\lambda^2 + (\alpha^2 + y_3^2)(\alpha^2 + y_1^2) + \alpha^2 y_2^2 \end{aligned}$$

$$\begin{aligned}
&\Rightarrow \begin{cases} 2\alpha^2 + y_1^2 + y_2^2 + y_3^2 = 0 \\ (\alpha^2 + y_3^2)(\alpha^2 + y_1^2) + \alpha^2 y_2^2 = \alpha^4 + \alpha^2(y_1^2 + y_2^2 + y_3^2) + y_1^2 y_3^2 = 0 \end{cases} \\
&\Leftrightarrow \begin{cases} 2\alpha^2 + y_1^2 + y_2^2 + y_3^2 = 0 \\ \alpha^4 = y_1^2 y_3^2 \end{cases} \\
&\Leftrightarrow \begin{cases} y_1^2 = \alpha^2 & (\text{known}) \\ y_2^2 = -4\alpha^2 \\ y_3^2 = \alpha^2 \end{cases}
\end{aligned}$$

Here we take  $\alpha = 1$ , so we can make  $y_1 = y_3 = 1$  and  $y_2 = \pm 2\sqrt{-1}$ . Again there needs to be a matrix  $L \in \text{SL}(n, \mathbb{F})$  such that  $LL^T = g = I + J$ . This can be done by calculations again. For example in the case  $y_2 = 2\sqrt{-1}$ , we have

$$L = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{\sqrt{-1}}{\sqrt{2}} & 0 & 0 \\ 0 & 2\sqrt{2} & -2\sqrt{-1} & 0 \\ 0 & 0 & \frac{-1}{2\sqrt{-1}} & \frac{1}{2} \end{bmatrix} \text{ and } LL^T = g = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & 2\sqrt{-1} & 0 \\ 0 & 2\sqrt{-1} & 2 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Since we are only able to introduce even number of -1 into  $J$  to keep the determinant to be 1, we would have two orbits with  $\text{rank}(J) = 3$ . The  $K$ -orbits of unipotent elements for  $n = 4$  are the two orbits introduced by these two representative matrices of rank 3 and all orbits included in the case  $n = 3$ .

## 7.7 Some remarks

1. If  $n = 2m + 1$  is odd, we should always have only one  $K$ -orbit of unipotent elements with rank  $n - 1$ , but we will have two  $K$ -orbits of unipotent elements since we have  $n - 1$  off-diagonals and are only able to introduce even numbers of  $-1$ 's. All  $K$ -orbits will be the ones of rank  $n - 1$  and the ones included in  $n - 1$  case. If  $n$  is odd, then only one  $K$ -orbit of rank  $n - 2$  from the  $K$ -orbits of  $n - 1$  case is included.

2. The key point here is whether we can make diagonals into alternating pattern, and off-diagonals into the entries we want. We can introduce some free variables. It's pretty clear that, in the way we introduce free variables, say we introduce  $m$  unknowns, the first few diagonals/off-diagonals will have degree  $m$  of freedom, but how can we show it's also true that all diagonals/off-diagonals have degree  $m$  of freedom, which would allow us to move from one solution  $x_i, y_i$  to another one  $\bar{x}_i, \bar{y}_i$ . This is the key part since, if we can do it, combining with condition  $\det(\lambda I + J) = \lambda^n$  we should be able to uniquely determine one or two solutions as our

representatives, which would actually finish our classification over algebraically closed fields. This is clearly seen in the case  $n = 3$  and 4. We will have same observation if we work on the case  $n = 5$  or 6, and so on.

3. Even if we get the desired tri-diagonal matrices satisfying certain conditions, we still need to show that there exists some matrix  $L \in \mathrm{SL}(n, \mathbb{F})$  such that  $LL^T = g = I + J$ . This guarantees that we are studying unipotent elements in the symmetric variety  $P = G/K$ .

4. Most work done in this dissertation is from the viewpoint of matrix theory. The possible solution for the problem we meet is theories from symmetric variety and orthogonal groups. If we still want to follow the matrix approach, we can build up tools parallel to the  $\lambda$ -matrix theory which gives us the JCM, since we see the pattern of  $x_i$  and  $y_i$  in the characteristic polynomial. The question is how we can translate the role of rotations in the characteristic polynomial into a less complicated way: to keep the matrix in the tri-diagonal form, we always need to use a series of rotations, which give us a complicated orthogonal matrix.

## REFERENCES

- [1] Michael Artin. *Algebras*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1991.
- [2] P. Bala and R. W. Carter. Classes of unipotent elements in simple algebraic groups. I. *Math. Proc. Camb. Phil. Soc.*, 79(3):401–425, 1976.
- [3] P. Bala and R. W. Carter. Classes of unipotent elements in simple algebraic groups. II. *Math. Proc. Camb. Phil. Soc.*, 80(1):1–18, 1976.
- [4] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer Verlag, New York, second edition, 1991.
- [5] M. Brion and Aloysius G. Helminck. On orbit closures of symmetric subgroups in flag varieties. *Canad. J. Math.*, 52(2):265–292, 2000.
- [6] F. Bruhat and J. Tits. Groupes réductifs sur un corps local. i. données radicielles valuées. *Inst. Hautes Études Sci. Publ. Math.*, 41:5–251, 1972.
- [7] J. L. Brylinski and P. Delorme. Vecteurs distributions  $h$ -invariant pour les séries principales généralisées d’espaces symétriques réductifs et prolongement méromorphe d’intégrals d’eisenstein. *Invent. Math.*, 109(3):619–664, 1992.
- [8] J. Carmona and P. Delorme. Base méromorphe de vecteurs distributions  $h$ -invariants pour les séries principales généralisées d’espaces symétriques réductifs: equation fonctionnelle. *J. Funct. Anal.*, 122(1):152–221, 1994.
- [9] E. Cartan. Groups simples clos et géométrie riemannienne. *J. Math. Pures Appl.*, 8:1–33, 1929.
- [10] C. De Concini and C. Procesi. Complete symmetric varieties. In *Invariant theory (Montecatini, 1982)*, volume 996 of *Lecture notes in Math*, pages 1–44. Springer-Verlag, Berlin, 1983.
- [11] C. De Concini and C. Procesi. Complete symmetric varieties. II. Intersection theory. In *Algebraic groups and related topics (Kyoto/Nayoya, 1983)*, pages 481–513, Amsterdam, North-Holland, 1985.
- [12] P. Delorme. Formule de plancherel pour les espaces symétriques réductifs. *Anal. Math.*, 147(2):417–452, 1998.
- [13] J. Faraut. Distributions sphérique sur les espaces hyperboliques. *J. Math. Pures Appl.*, 58:369–444, 1979.
- [14] M. Flensted-Jensen. Discrete series for semisimple symmetric spaces. *Annals of Math.*, 111:253–311, 1980.
- [15] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics*. American Mathematical Society, 2001. Reprint with corrections.

- [16] A. G. Helminck and S. P. Wang. On rationality properties of involutions of reductive groups. *Adv. in Math.*, 99:26–96, 1993.
- [17] Aloysius G. Helminck. *Symmetric  $k$ -varieties*, volume 56 of *Proc. Sympos. Pure Math.*, pages 233–279. Amer. Math. Soc., 1994.
- [18] Aloysius G. Helminck. *Combinatorics related to orbit closures of symmetric subgroups in flag varieties*, volume 35 of *CRM Proc. Lecture Notes*, pages 71–90. Amer. Math. Soc., Providence, RI, 2004.
- [19] Aloysius G. Helminck. On orbit decompositions of  $k$ -variety. In *Symmetry and spaces*, volume 278 of *Progress in Mathematics*, pages 83–127. Birkhäuser Boston Inc., 2010.
- [20] Aloysius G. Helminck and Gerald W. Schwarz. Orbits and invariants associated with a pair of commuting involutions. *Duke Math. J.*, 106(2):237–279, 2001.
- [21] Aloysius G. Helminck and Gerald W. Schwarz. Real double cosets spaces and their invariants. *J. Algebra*, 322(1):219–236, 2009.
- [22] Aloysius G. Helminck and Ling Wu. Classification of involutions of  $SL(2, k)$ . *Comm. Algebra*, 30(1):193–203, 2002.
- [23] Aloysius G. Helminck, Ling Wu, and Christopher E. Dometrius. Involutions of  $SL(n, k)$ ,  $n > 2$ . *Acta Appl. Math.*, 90(12):91–119, 2006.
- [24] F. Hirzebruch and P. J. Slodowy. Elliptic genera, involutions, and homogeneous spin manifolds. *Geom. Dedicata*, 35(1-3):309–343, 1990.
- [25] James E. Humphreys. *Linear algebraic groups*. Number 21 in Graduate Texts in Mathematics. SpringerVerlag, New York, 1975.
- [26] G. Lusztig and D. A. Vogan. Singularities of closures of  $K$ -orbits on flag manifolds. *Invent. Math.*, 71:365–379, 1983.
- [27] George Lusztig. On the finiteness of the number of unipotent classes. *Invent. Math.*, 34:201–213, 1976.
- [28] Toshihiko Matsuki. Double Coset Decompositions of Reductive Lie Groups Arising from Two Involutions. *J. Algebra*, 197(1):49–91, 1997.
- [29] T. Oshima and T. Matsuki. A description of discrete series for semisimple symmetric spaces. *Adv. Stud. Pure Math.*, 4:331–390, 1980.
- [30] T. Oshima and J. Sekiguchi. Eigenspaces of invariant differential operators in an affine symmetric space. *Invent. Math.*, 57:1–81, 1980.
- [31] R. W. Richardson. On orbits of algebraic groups and Lie groups. *Bull. Austral. Math. Soc.*, 25(1):1–28, 1982.
- [32] R. W. Richardson. Orbits, invariants and representations associated to involutions of reductive groups. *Invent. math.*, 66:287–312, 1982.

- [33] T. A. Springer. *Some results on algebraic groups with involutions*, volume 6 of *Adv. Stud. In Pure Math.*, pages 525–543. Academic Press, Orlando, FL, 1984.
- [34] T. A. Springer. *Linear algebraic groups*, volume 9 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, second edition, 1998.
- [35] E. P. van den Ban and H. Schlichtkrull. Multiplicities in the plancheral decomposition for a semisimple symmetric space. *Contemp. Math.*, 145:163–180, 1993.
- [36] E. P. van den Ban and H. Schlichtkrull. Fourier transforms on a semisimple symmetric space. *Invent. Math.*, 130:517–574, 1997.
- [37] E. P. van den Ban and H. Schlichtkrull. The most-continuous part of the plancherel decomposition for a reductive symmetric space. *Ann. Math.*, 145:267–364, 1997.
- [38] Varadarajan, editor. *Harish-Chandra: Collected papers*, volume I-IV. Springer-Verlag, New York, 1984.
- [39] J. A. Vust. Finiteness of orbit structure for real flag manifolds. *Geom. Dedicata*, 3:377–384, 1974.