Images of Polynomial Maps with Constants on Algebras

विद्या वाचस्पति की

उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

द्वारा / By

प्राची सैनी / Prachi Saini

पंजीकरण सं. / Registration No.: 20213603

शोध प्रबंध पर्यवेक्षक / Thesis Supervisor:

Prof. Anupam Singh

IISER PUNE

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

2026

To my constants

Daksh and Kanishk

Declaration of Authorship

Name of Student	: Prachi Saini
Reg. No.	: 20213603
Thesis Supervisor	: Prof. Anupam Singh
Department	: Mathematics
Date of joining program	: September 1, 2021
Date of Pre-Synopsis Seminar	: February 3, 2026

Title of Thesis: Images of Polynomial Maps with Constants on Algebras

I declare that this written submission represents my ideas in my own words and that where others' ideas have been included, I have adequately cited and referenced the original sources. I declare that I have acknowledged collaborative work and discussions wherever such work has been included. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented, fabricated, or falsified any idea, data, fact, or source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and may also invoke penal action from the concerned sources.

The work reported in this thesis is the original work done by me under the guidance of Prof. Anupam Singh.

Date: 09.04.2026

Signature: Prachi Saini

Certificate

Certified that the work incorporated in the thesis entitled ***Images of Polynomial Maps with Constants on Algebras***, Submitted by **Prachi Saini**, was carried out by the candidate under my supervision. The work presented here, or any part of it, has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

(Supervisor)

Anupam Singh

Date:

Acknowledgements

I express my deepest gratitude to my supervisor, Prof. Anupam Singh, for his unwavering guidance, patience, and encouragement throughout my PhD journey. Whenever I encountered particularly challenging problems, he offered timely advice while giving me the space and confidence to think deeply and arrive at solutions independently. Beyond providing thoughtful and detailed feedback, he consistently challenged me to ask deeper questions and develop a rigorous approach to problem-solving. His mentorship has been a guiding force in both my academic development and personal growth.

I am sincerely grateful to my research advisory committee members, Prof. Amit Kulshrestha and Prof. Rama Mishra, for their valuable feedback and thoughtful suggestions, which greatly enriched my research. I also thank Prof. Soumen Maity for his continued support throughout my time in the department. I would also like to acknowledge the faculty members from my coursework, Prof. Steven Spallone, Dr. Debargha Banerjee, Dr. Supriya Pisolkar, Dr. Haripada Sahu, and Dr. Chandrasheel Bhagwat, whose teaching and mentorship provided me with a strong foundation. I am especially thankful to Dr. Chandrasheel Bhagwat for his additional guidance throughout my PhD journey.

I also thank Prof. B. Sury, Dr. Maneesh Thakur, Prof. K.N. Raghavan, Prof. Amritanshu Prasad, and Prof. Sudhir Ghorpade for the many valuable discussions we shared during conferences, workshops, and other meetings. Their insights and thoughtful conversations not only helped me explore new perspectives but also significantly shaped my approach to research.

The guidance and support of my master's teachers at IIT Hyderabad, Dr. Pradipto Banerjee, Prof. Narsimha Kumar, and Prof. Venku Naidu, were instrumental in building my foundational knowledge and inspiring me to pursue research with rigor and curiosity.

I am deeply grateful to the administrative staff, especially Suvarna Bhardwaj and Yogesh Kolap, for their support in handling departmental matters. I also thank Sayalee Damle for her help related to my fellowship, which ensured the smooth management of logistics. I am grateful to CSIR for the fellowship, which provided the financial support necessary to pursue this research.

I am deeply thankful to Saikat, for the many thoughtful and insightful discussions we shared throughout my PhD. I am especially grateful for his guidance, and constant support, which helped me navigate both the challenges of research and the pressures of the program. During my visits to Harish-Chandra Research Institute and Indian Statistical Institute, Bangalore, he generously hosted me and engaged in numerous discussions, meetings, and collaborative sessions, all of which had a lasting influence on my research.

I would like to acknowledge my friends Ramandeep, Hitendra, Aniket, Ajinkya, Ronit, Bittu, Tushar, Subhash, and Narayanan at IISER, with special thanks to Ronit for his support during the initial years of my PhD. My thanks are also due to Clifford, whose long discussions over chai, spanning mathematics and life, brought fresh ideas, new perspectives, and laughter, enriching both my research and my life. I would also like to thank Hrithik, Aanshi, Vivek, Himanshu, Pulkit, and Utkarsh, whose company provided balance and joy beyond academic life.

I thank Rishabh for being a constant presence in my life over the past years, standing by me through both highs and lows. During times when circumstances became difficult, his encouragement and steady support helped me regain strength and perspective. His friendship has been an essential part of this journey.

My family has been a source of inspiration, support, and balance throughout this journey. Daksh's hard work, persistence, and dedication have consistently motivated me through personal and academic challenges, while Kanishk's warmth, energy, and constant presence lifted my spirits and helped me maintain perspective during difficult times. My cousins, Bhumika, Nikunj, and Lavish, have been sources of joy and relaxation, often keeping me connected to their lives and reminding me to take breaks and enjoy life beyond work. I would also like to remember Ritu di, who is no longer with us. Her encouragement in mathematics during my school days planted the seed that eventually led me here, and her inspiration continues to guide me.

Finally, I am profoundly grateful to my parents, whose unwavering love, sacrifices, and faith in me have been the foundation of everything I have achieved. Without their support, patience, and guidance, this journey would not have been possible.

To all of you, I offer my heartfelt thanks.

Prachi

List of Publications

Publications included in the thesis

1. Saikat Panja, **Prachi Saini**, and Anupam Singh, *Surjectivity of polynomial maps on matrices*, European Journal of Mathematics **11** (2025), no. 3, Paper No. 62, 22 pp.
2. Saikat Panja, **Prachi Saini**, and Anupam Singh, *Images of polynomial maps with constants*, Mathematika **71** (2025), no. 3, Paper No. e70031, 17 pp.
3. Saikat Panja, **Prachi Saini** and Anupam Singh, *Polynomial maps with constants on split octonion algebras*, Communications in Algebra **54** (2026), no. 2, 579–594.

Abstract

This thesis studies *polynomial maps with constants* evaluated on algebras, with particular emphasis on *surjectivity* and *classification of images.* Given an algebra $\mathscr{A}$ and a noncommutative polynomial ω whose coefficients lie in $\mathscr{A}$, the associated evaluation map

$$\omega\colon \mathscr{A}^m \to \mathscr{A}$$

naturally raises several fundamental questions: when is the map surjective, when is its image a vector space, and how do these properties depend on the choice of coefficients?

The main focus of this thesis is on *diagonal polynomial maps* and *Lv́ov–Kaplansky type maps* on central simple algebras, such as matrix algebras and quaternion algebras, as well as on the octonion algebra, which is non-associative. In particular, we address the following problems.

Surjectivity of diagonal maps. Let m be a positive integer and $n \geq 2$. Given integers $k_1, k_2, \ldots, k_m \geq 1$ and nonzero elements $A_1, \ldots, A_m \in \mathscr{A}$, consider the diagonal map

$$\begin{aligned} \omega\colon M_n(\mathscr{A})^m &\longrightarrow M_n(\mathscr{A}), \\ (x_1, \ldots, x_m) &\mapsto A_1 x_1^{k_1} + \cdots + A_m x_m^{k_m}. \end{aligned}$$

We study the minimum value of m for which ω is surjective. In the case $m = 2$, we

determine conditions on A_1 and A_2 that ensure surjectivity.

Lv́ov–Kaplansky type maps. Let $\mathbb{F}$ be a field and $\mathscr{A} = M_n(\mathbb{F})$. For

$$\omega = A_1(x_1x_2) - A_2(x_2x_1) \in \mathscr{A}\langle x_1, x_2\rangle,$$

with $A_1, A_2 \in \mathscr{A}$, we determine precisely when the image of the associated map is a vector space.

These problems unify themes arising from Waring-type problems, polynomial identities, and orbit classification under automorphism group actions, and contribute to the broader program of understanding images of polynomial maps on algebras.

To address these questions, we employ tools such as canonical forms of matrices, actions of automorphism groups, simultaneous conjugation, and the reduction of solvability over extension fields to solvability over base fields.

We first study diagonal maps on matrix algebras over sufficiently large finite fields, algebraically closed fields, and the real field with scalar coefficients, determining the minimum number of variables required for surjectivity. As a consequence, we also describe the images of such maps on Hamiltonian quaternions and division octonion algebras.

Using the theory of central simple algebras, we then analyze diagonal maps with coefficients from the algebra itself on $M_2(\overline{\mathbb{F}})$. By classifying orbit representatives under the action of the automorphism group, we obtain explicit conditions on the coefficients that guarantee surjectivity.

Next, we extend this approach to the split octonion algebra, obtained via the Cayley–Dickson construction. Using the classification of orbit representatives under the action of the exceptional group G_2, we determine conditions ensuring surjectivity of diagonal maps in two variables over an algebraically closed field.

Finally, we classify Lv́ov–Kaplansky type maps on $M_2(\overline{\mathbb{F}})$, determining exactly when their images form vector spaces. Together, these results contribute to a systematic understanding of polynomial images on matrix algebras and related algebraic structures.

Notation

$\mathbb{N}$	natural numbers
$\mathbb{Z}$	integers
$\mathbb{F}$	field
$\mathbb{F}^{\times}$	$\mathbb{F} \setminus \{0\}$
$\overline{\mathbb{F}}$	algebraic closure of $\mathbb{F}$
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\mathbb{F}_q$	finite field with q elements
$\mathscr{A}$	algebra
$\mathscr{R}$	ring
D	division algebra
$\mathscr{O}$	octonion algebra

$\mathscr{Q}$	quaternion algebra
G	group
e_i	vector with 1 in the i-th place and 0 otherwise
$M_n(\mathbb{F})$	space of all $n \times n$ matrices over $\mathbb{F}$
E_{ij}	matrix with 1 in the ij-th place and 0 otherwise
$\mathrm{GL}_n(\mathbb{F})$	general linear group over $\mathbb{F}$
$\mathrm{PGL}_n(\mathbb{F})$	projective general linear group over $\mathbb{F}$
$\mathbb{F} \cdot I_n$	space of scalar matrices of size n over $\mathbb{F}$
$\mathfrak{sl}_n(\mathbb{F})$	space of trace 0 matrix of size n over $\mathbb{F}$
$\mathrm{Aut}(\mathscr{A})$	automorphism group of an algebra $\mathscr{A}$
G_2	automorphism group of octonion algebra
$\mathscr{Z}(\mathscr{A})$	center of $\mathscr{A}$
$\mathscr{C}_G(H)$	centralizer of H in G
$\cong$	isomorphism
$\oplus$	direct sum
$\otimes$	tensor product

Σ	sum
$\mathscr{N}$	bilinear form
N	norm or quadratic form
$\mathscr{C}$	composition algebra
$\bar{}$	conjugation of an element
Tr	trace
$\star$	free product
J	Jordan form

Contents

Chapter 1

Introduction

1.1 The Classical Waring Problem

In 1770, Edward Waring in his book [63] conjectured that every natural number can be expressed as a sum of finitely many k-th powers. In other words, for every natural number $n \in \mathbb{N}$ and for each integer $k \geq 2$, there exist a number $g(k) \in \mathbb{N}$ such that

$$n = a_1^k + a_2^k + \cdots + a_{g(k)}^k$$

where $a_i \in \mathbb{N} \cup \{0\}$, for each $1 \leq i \leq g(k)$. This came to be known as *classical Waring problem.*

In 1772, Lagrange proved the celebrated *four-square theorem* [36], settling the case for $k = 2$ of the classical Waring problem. The theorem states that for every $n \in \mathbb{N}$, there exist a_1, a_2, a_3 and a_4 in $\mathbb{N} \cup \{0\}$, such that

$$n = a_1^2 + a_2^2 + a_3^2 + a_4^2.$$

The number four is minimal, since not every natural number can be represented as a sum of three squares. In fact, by a theorem of Legendre, a natural number n

can be written as a sum of three squares if and only if it is not of the form

$$n = 4^k(8m+7)$$

for integers $k,m \geq 0$. Following Lagrange's proof, Waring's conjecture came to be viewed as a far-reaching generalisation of the four-square theorem. Hilbert in 1909 [27] proved the existence of $g(k)$ for every $k \geq 2$.

1.2 Waring-Type Problems for Matrices and Algebras

While the original conjecture concerned integers, the classical problem has inspired a rich body of research, including generalisations to algebraic structures beyond integers. I. Connell posed a natural matrix analogue of the classical Waring problem: For a given $A \in M_2(\mathbb{Z})$, what is the minimum $g(2) \in \mathbb{N}$ such that

$$A = A_1^2 + A_2^2 + \cdots + A_{g(2)}^2$$

for some $A_1, A_2, \ldots, A_{g(2)} \in M_2(\mathbb{Z})$?

L. Carlitz subsequently answered this question by proving that $g(2) = 3$, that is, every 2×2 integer matrix can be expressed as a sum of squares of integer matrices, see [13]. This result provides a natural extension of the classical Waring problem to the setting of matrices and lays the foundation for further study of Waring-type problems over algebras.

1.2.1 Diagonal Forms

These problems are closely related to the study of images of polynomial maps evaluated on algebraic structures. One such family of maps studied in this thesis is stated as:

Question 1.2.1. Let $k_1, k_2, \ldots, k_m \geq 1$ be integers, and $\delta_1, \ldots, \delta_m \in \mathbb{F}$ be non-zero scalars. Consider the diagonal polynomial

$$\omega(x_1, \ldots, x_m) = \delta_1 x_1^{k_1} + \delta_2 x_2^{k_2} + \cdots + \delta_m x_m^{k_m} \tag{1.2.1}$$

in m-variables.

For a given $\mathbb{F}$-algebra $\mathscr{A}$, what is the minimum value of m for which the associated polynomial map is surjective when evaluated on $\mathscr{A}$?

When $k_1 = \ldots = k_m = k$ the Equation (1.2.1) is a k-form. Such a k-form is said to be **universal** on $\mathscr{A}$ if the map ω is surjective. The classical Waring problem corresponds to the case $\delta_1 = \delta_2 = \ldots = \delta_m = 1$ and asks for the smallest m such that the form is universal.

The universality of the quadratic form ($k = 2$ case) over fields has been extensively studied, including arithmetic aspects (see, for example, Bhargava [8] and Bhargava et al. [9]). Diagonal forms over function fields were studied by Voloch [61], while further foundational contributions to quadratic and diagonal forms over function fields were made by Colliot-Thélène and Sansuc [16] and by Parimala and Suresh [50]. The image of k-forms over matrix algebras also received considerable attention.

In [25], Griffin and Krusemeyer studied the image of quadratic form, i.e., for $k = 2$, in three variables over matrix algbera $M_n(\mathbb{F})$, where $\mathbb{F}$ is a field of characteristic not equal to two, and showed that $g(2) = 3$. Moreover, they proved that if n is odd and $\alpha \in \mathbb{F}$, then αI_n is a sum of two squares if and only if α is a sum of two squares in $\mathbb{F}$. They conjectured that $g(2) \neq 2$, and that the only matrices not expressible as sums of two squares are scalar matrices of the form αI_n where α is not representable as sum of two squares. This conjecture was settled by Richman [53] using elementary techniques from linear algebra.

More generally, the Waring problem for an algebra $\mathscr{A}$ is to determine $g(k)$. Richman [52] considered $\mathscr{A} = M_n(\mathscr{R})$ where $\mathscr{R}$ is a commutative associative ring with unity and $n \geq k$. He showed that $g(k) \leq 7$, and $g(2) \leq 4$. Combining results of Vaserstein [59] and Richman [52], one obtains $g(2) = 3$ for $M_n(\mathbb{Z})$. Since then, there have been several generalisations of *Waring-type problem* on algebras.

Leep, Shapiro, and Wadsworth [40] studied the sum of squares in central simple algebras and established conditions under which

$$D = \sum D^2$$

for a finite-dimensional division algebra D of characteristic not equal to two. Since the quaternion algebra serves as the foundational example of a non-commutative 4-dimensional division algebra, and octonion algebras are obtained from quaternion algebras via the Cayley-Dickson Doubling process, it is natural to extend results from the quaternion setting to the octonion case. In this direction, Pümplün [51] generalised the corresponding division algebra results to octonion algebras, which are non-associative.

1.3 Polynomial Maps with Constants

Motivated by the developments for diagonal forms, the study of Waring-type problems in both associative and non-associative algebras has emerged as a natural extension of the classical problem. The generalisation of this problem can be stated as follows:

Let $\mathscr{F}_m := \mathbb{F}\langle x_1, x_2, \ldots, x_m\rangle$ be a free algebra of rank m over the field $\mathbb{F}$. Let $\mathscr{A}$ be an $\mathbb{F}$-algebra. Consider the free algebra over $\mathscr{A}$, given by $\mathscr{A}_m := \mathscr{A} * \mathscr{F}_m$. An element $\omega(x_1,\ldots,x_m) \in \mathscr{A}_m$ is a finite sum of the form

$$\sum (a_{j_1}) x_{i_1}^{k_1} (a_{j_2}) x_{i_2}^{k_2} \ldots (a_{j_r}) x_{i_r}^{k_r}$$

where $a_{j_s} \in \mathscr{A}$ and $i_s, k_s \in \mathbb{Z}_{\geq 0}$ for all $1 \leq s \leq r$. Such an element defines an evaluation map over the algebra $\mathscr{A}$ by

$$\omega : \mathscr{A}^m \longrightarrow \mathscr{A}$$
$$(a_1, a_2, \ldots, a_m) \mapsto \omega(a_1, a_2, \ldots, a_m).$$

These maps are referred to as *polynomial maps with constants*. These maps allow for a broader class of expressions and reveal structural properties of the underlying algebra, including the identities satisfied by its elements.

1.3.1 The Lv́ov-Kaplansky Conjecture and Its Generalisations

One of the central conjectures concerning images of polynomial maps is the Lv́ov-Kaplansky Conjecture, see [1].

Conjecture 1 (Lv́ov-Kaplansky)**.** *Let $\mathbb{F}$ be an infinite field, and let f be a multilinear polynomial. Then the image of f, evaluated on the matrix algebra of $n \times n$ matrices over $\mathbb{F}$, forms a vector subspace. Moreover, the image is one of the following:*

$$\{0\}, \quad \mathbb{F} \cdot I_n, \quad \mathfrak{sl}_n(\mathbb{F}), \quad \textit{or} \quad M_n(\mathbb{F}).$$

Although the conjecture remains open, several partial results are known. Kanel-Belov, Malev, and Rowen [30] proved that the conjecture holds for $n = 2$ over quadratically closed fields. For arbitrary fields, Malev [44] showed that the image of a multilinear polynomial evaluated on $M_2(\mathbb{F})$ is of the form: $\{0\}$, or $\mathbb{F} \cdot I_2$, or a subset of $\mathfrak{sl}_2(\mathbb{F})$, or the whole algebra $M_2(\mathbb{F})$. He further proved that for $\mathbb{F} = \mathbb{R}$ and $n = 2$, the conjecture holds. Subsequently, Belov, Malev and Rowen [29] classified the possible images of multilinear polynomial evaluations for $n = 3$. Further contributions in this direction include works by Fagundes [17], Wang-Zhou-Luo [62], and Gargate–Mello [18]. The images of multilinear polynomials have also been studied over composition algebras, notably the quaternion algebra [45] and the octonion algebra [31].

The classical motivation for this conjecture comes from early works on polynomial images and trace structures in matrix algebras. Amitsur and Levitzki [3] studied identities for matrix algebras, later extending the results to semisimple algebras. Albert and Muckenhoupt [2] and K. Shoda [57] showed that the matrices with trace zero lie in the image of a commutator map. Moreover, the theory on the image of the commutator map on central simple algebras and division rings has been studied by Amitsur and Rowen, see [4]. They show that for a central division algebra D over the field $\mathbb{F}$, any non-central element of $M_n(D)$, with $n \geq 2$, having reduced trace 0 lies in the image of the commutator map.

Beyond predicting a precise classification of possible images, a fundamental aspect of the conjecture is the assertion that the image of a multilinear polynomial evaluated on a matrix algebra is always a vector subspace. Thus, the conjecture addresses both the algebraic structure and the linear nature of polynomial images.

The framework of polynomial maps with constants naturally leads to a broader class of image problems. In this setting, it is natural to ask whether analogues of the Lv́ov-Kaplansky conjecture continue to hold, and in particular, whether the images of such maps remain vector subspaces. This perspective motivates the following generalised Lv́ov-Kaplansky problem, which concerns polynomial maps with fixed coefficients from the underlying algebra.

1.3.1.1 The generalized Lv́ov-Kaplansky problem

Let $\mathbb{F}$ be a field and let $\mathscr{A} = M_n(\mathbb{F})$. The classical Lv́ov-Kaplansky conjecture predicts that the image of a multilinear polynomial without constants evaluated on $M_n(\mathbb{F})$ is a linear subspace of $M_n(\mathbb{F})$. In particular, the conjecture asserts that the image is a vector space. Motivated by this conjecture, one is naturally led to consider polynomial maps in which coefficients are allowed to lie in the algebra itself. In particular, let,

$$f(x_1,\ldots,x_m) = \sum_{\sigma \in S_m} A_{\sigma,0}\, x_{\sigma(1)} A_{\sigma,1}\, x_{\sigma(2)} \cdots x_{\sigma(m)} A_{\sigma,m},$$

where each $A_{\sigma,i} \in M_n(\mathbb{F})$ is a fixed. The variables $x_1,\ldots,x_m$ are non-commuting, and the polynomial f is multilinear in the sense that each variable appears exactly once in every monomial. Such a polynomial is called a *multilinear polynomial with constants.*

The following problem may be viewed as a natural extension of the Lv́ov-Kaplansky conjecture.

Conjecture 2 (Generalized Lv́ov-Kaplansky problem)**.** *Let f be a multilinear polynomial with coefficients in $M_n(\mathbb{F})$. Is the image*

$$f\big(M_n(\mathbb{F}),\ldots,M_n(\mathbb{F})\big)$$

necessarily a linear subspace of $M_n(\mathbb{F})$?

Motivated by the results on the commutator map, one such class of maps studied in this thesis consists of generalisations of the commutator map, which arise naturally in the investigation of images of polynomial maps on matrix algebras and are closely related to questions motivated by the Lv́ov–Kaplansky conjecture.

Question 1.3.1 (Generalized Commutator Map)**.**

$$\omega(x_1,x_2) = A_1(x_1x_2) - A_2(x_2x_1) \in M_n(\mathbb{F}) \star \langle x_1, x_2 \rangle.$$

Determine for which pair (A_1, A_2), the image of the map ω is a vector subspace of $M_n(\mathbb{F})$.

Results describing the images of specific polynomial maps with constants provide supporting evidence for this generalized problem. In particular, our results on commutator-type polynomial maps with constants show that their images form linear subspaces, thereby supporting the general philosophy underlying the generalized Lv́ov-Kaplansky problem.

The map belongs to the class of *polynomial maps with constants*. Parallel developments have also influenced the study of images of polynomial maps in the theory of word maps with constants in group theory; see, for instance, [28]. In this direction, Gordeev, Kunyavskii, and Plotkin initiated a systematic study of word maps with constants and established several foundational results for simple algebraic groups; see [23, 24]. Following their work, a growing body of literature has emerged investigating the structural and image-theoretic properties of word maps with constants in various group settings; see, for example, [21, 22, 56].

1.3.2 Generalized Waring Problem

In the context of group theory, the Waring problem in two variables for finite simple groups was studied by Larsen, Shalev, and Tiep [38], who showed that for groups of sufficiently large order, every element can be expressed as a product of two values of a fixed power word. This result highlights that it suffices to work with only two variables to obtain surjectivity.

A parallel phenomenon appears in the context of matrix algebras. Following a question of Larsen, Kishore and Singh [35] studied the Waring problem for matrices over finite fields. Their result shows that, over a finite field of sufficiently large cardinality, every matrix can be expressed as a sum of two k-th powers. In recent work [49], it was shown that for Waring-type problems on matrix algebras over algebraically closed fields, it is sufficient to consider diagonal polynomial maps in two variables, that is, $m = 2$ in (1.2.1). This observation provides a strong analogy with the group-theoretic setting and motivates a focused study of two-variable polynomial maps with constants. Guided by these developments, this thesis investigates the following problem, which is stated in a general setting:

Question 1.3.2. Let $k_1, k_2 \geq 1$ be integers and $A_1, A_2 \in \mathscr{A}$ be non-zero elements. Consider

$$\omega(x_1, x_2) = A_1(x_1)^{k_1} + A_2(x_2)^{k_2} \in \mathscr{A} * \mathbb{F}_2.$$

For what $(A_1, A_2) \in \mathscr{A}^2$, is the map ω surjective on $\mathscr{A}$?

1.4 Scope and Contributions of the Thesis

Motivated by the classical Waring problem and its algebraic analogues, this thesis studies the surjectivity and image structure of polynomial maps on associative and non-associative algebras. In the Introduction, we formulated three guiding questions (Questions 1.2.1, 1.3.2 and 1.3.1), which address diagonal maps, polynomial maps with coefficients, and commutator-type maps, respectively. The principal contributions of this thesis provide answers to these questions over matrix algebras, which is the starting point of studying such maps on central simple algebras and composition algebras, including the split octonion algebra.

1.4.1 Diagonal maps on Central Simple Algebras

Question 1.2.1 asks when diagonal maps in equation 1.2.1 are surjective when evaluated on matrix algebras over various fields, and, in the surjective case, what is the minimal number of variables m required.

This question is motivated by the classical Waring problem and its extensions to non-commutative settings, particularly matrix algebras. Waring-type problems for matrix algebras have been extensively studied; see, for example, the work of Katre and collaborators [32, 33], Garge [5, 19], and Brešar–Semrl [11, 12, 10]. More generally, problems concerning the image structure of polynomial maps on algebras have been investigated in, for instance, [39, 46].

The following theorems, proved in Chapter 7, provide a complete answer to Question 1.2.1 in several important cases.

Theorem A. *Let m be a positive integer and $n \geq 2$. Given integers $k_1, k_2, \ldots, k_m \geq 1$, and $\delta_1, \ldots, \delta_m \in \mathbb{R}$ all non-zero, consider the diagonal map $\omega\colon M_n(\mathbb{R})^m \to M_n(\mathbb{R})$ for $n \geq 2$ given by*

$$\omega(x_1, \ldots, x_m) = \delta_1 x_1^{k_1} + \delta_2 x_2^{k_2} + \cdots + \delta_m x_m^{k_m}.$$

Then we have, for $m = 2$, the map ω is surjective except when n is odd, $\delta_1\delta_2 > 0$ and k_1, k_2 are both even (in that case the image misses negative scalar matrices). It is surjective for $m \geq 3$.

This result is proved via two intermediate theorems, appearing as Theorem 7.1.2 and Theorem 7.1.3.

Theorem B. *Let m be a positive integer and $n \geq 2$. Given integers $k_1, k_2, \ldots, k_m \geq 1$, and $\delta_1, \ldots, \delta_m \in \mathbb{H}$ all non-zero, consider the diagonal map $\omega\colon M_n(\mathbb{H})^m \to M_n(\mathbb{H})$*

for $n \geq 2$ given by

$$\omega(x_1, \ldots, x_m) = \delta_1 x_1^{k_1} + \delta_2 x_2^{k_2} + \cdots + \delta_m x_m^{k_m}.$$

Then the map ω is surjective for all $m \geq 2$.

This theorem appears as Theorem 7.1.4. As a consequence, we also determine the image of the diagonal map on the division octonion algebra $\mathscr{O}$ over $\mathbb{R}$, which is stated in Corollary 7.1.5.

Theorem C. *Let $k_1, k_2 \geq 1$ and $n \geq 2$ be integers and β be a non-zero element in the finite field $\mathbb{F}_q$. Consider the map $\omega \colon M_n(\mathbb{F}_q) \times M_n(\mathbb{F}_q) \to M_n(\mathbb{F}_q)$ given by $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$. Then, there exists a constant $\mathscr{K}(k_1, k_2)$ (which depends only on k_1 and k_2) such that for all $q > \mathscr{K}(k_1, k_2)$, the map ω is surjective.*

This result appears as Theorem 7.1.6.

1.4.2 Commutator map on Matrix Algebra

The question 1.3.1 focuses on generalised commutator-type polynomial maps, which are closely related to classical commutator maps studied by Shoda, Herstein and Amitsur-Rowen. The following result addresses this question in the case of 2×2 matrices and is answered in Chapter 6 (see Theorem 6.1.2).

Theorem D. *Let $\mathbb{F}$ be an algebraically closed field. Consider the polynomial map ω given by $A_1(x_1x_2) - A_2(x_2x_1)$ on $M_2(\mathbb{F})$ where $A_1, A_2 \in M_2(\mathbb{F})$ both non-zero. Then, the image is a vector subspace of $M_2(\mathbb{F})$.*

1.4.3 Polynomial Maps with Constants on Composition Algebras

Question 1.3.2 concerns polynomial maps with coefficients from the underlying algebra, which naturally generalises diagonal maps and arises in connection with word maps with constants in group theory.

In this part, we investigate the image structure and surjectivity of such maps on composition algebras, with particular emphasis on matrix algebras of size 2 and the split octonion algebra. The results addressing this question are established in Chapter 6 (see Theorem 6.1.1) and Chapter 8 (see Theorem 8.1.1).

Theorem E. *Let $\mathbb{F}$ be an algebraically closed field. Consider the polynomial map ω given by $A_1x_1^{k_1} + A_2x_2^{k_2}$ on $M_2(\mathbb{F})$ where $A_1, A_2 \in M_2(\mathbb{F})$ both non-zero. Then, the image of ω is a vector subspace of $M_2(\mathbb{F})$. Further, ω is surjective if and only if A_1 and A_2 can be simultaneously conjugated to a pair of matrices such that both the matrices do not have the same zero rows.*

Theorem F. *Let $\mathbb{F}$ be an algebraically closed field, and $\mathscr{O}(\mathbb{F})$ be the split octonion algebra over $\mathbb{F}$. Then, the map induced by $A_1(x_1^{k_1}) + A_2(x_2)^{k_2}$ on $\mathscr{O}(\mathbb{F})$, where $A_1, A_2 \in \mathscr{O}(\mathbb{F})\setminus\{0\}$, is surjective if and only if the pair (A_1, A_2) under $G_2(\mathbb{F})$-action does not represent one of the following pairs:*

1. $\left(\begin{pmatrix}\alpha_1 & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}, \begin{pmatrix}\beta_1 & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}\right)$
2. $\left(\begin{pmatrix}0 & \mathbf{0}\\ \mathbf{0} & \alpha_8\end{pmatrix}, \begin{pmatrix}0 & \mathbf{0}\\ \mathbf{0} & \beta_8\end{pmatrix}\right)$
3. $\left(\begin{pmatrix}\alpha_1 & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}, \begin{pmatrix}\beta_1 & (1,0,0)\\ (0,\beta_6,0) & 0\end{pmatrix}\right)$
4. $\left(\begin{pmatrix}0 & \mathbf{0}\\ \mathbf{0} & \alpha_8\end{pmatrix}, \begin{pmatrix}0 & (1,0,0)\\ (0,\beta_6,0) & \beta_8\end{pmatrix}\right)$

5. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}\right)$

6. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} \beta_1 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix}\right)$

7. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$

8. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & \mathbf{0} \\ (0,1,0) & 0 \end{pmatrix}\right)$

where $\alpha_1, \alpha_8, \beta_1, \beta_6, \beta_8 \in \mathscr{F}$.

Organisation of the Thesis

This thesis is organised as follows.

Chapter 1 serves as an introduction to the themes and motivations of the thesis. We place our work in the context of Waring-type problems and their algebraic analogues, and we introduce polynomial maps on algebras as a unifying framework for studying image and surjectivity problems.

Chapter 2 is devoted to the basic theory of central simple algebras. We recall fundamental definitions and structural results, including matrix algebras, division algebras, and the Artin–Wedderburn theorem. This chapter establishes the algebraic background required for the study of polynomial maps in later chapters.

Chapter 3 focuses on composition algebras. We review their defining properties, including the norm form and the composition property, and discuss key examples and structural results that will be used throughout the thesis.

Chapter 4 is devoted to quaternion and octonion algebras. We study their algebraic structure and properties in detail, highlighting both associative and non-associative aspects. This chapter prepares the groundwork for analysing polynomial maps on these algebras.

Chapter 5 introduces polynomial maps with constants. We discuss their definition, basic properties, and motivation, and explain how they generalise constant-free polynomial maps and diagonal maps. This chapter provides the conceptual framework for the results developed in subsequent chapters.

Chapter 6 contains our results on polynomial maps with constants on matrix algebras. We analyse the structure of their images, establish surjectivity results, and examine how the presence of constants affects the behaviour of polynomial maps in the matrix setting.

Chapter 7 is devoted to our results on diagonal polynomial maps on central simple algebras. Beginning with matrix algebras over finite fields, the complex field, and the real field, we study surjectivity and image structure. We then extend these results to quaternion and octonion algebras, obtaining a unified treatment of diagonal maps over all finite-dimensional real division algebras.

Chapter 8 studies polynomial maps with constants on the split octonion algebra. Using the Zorn vector matrix realisation, we analyse the image structure and surjectivity of such maps and identify phenomena that arise specifically from non-associativity and the isotropic nature of the split octonion algebra.

Chapter 9 concludes the thesis with a discussion of future directions and open problems. We outline possible extensions of the results obtained in this work and suggest avenues for further research related to polynomial maps on algebras and Waring-type problems.

Chapter 2

Central Simple Algebras

In this chapter, we introduce the notion of a central simple algebra along with the essential algebraic tools used to address our main problem. In the context of this thesis, central simple algebras provide the natural algebraic framework for studying polynomial maps and Waring-type problems on noncommutative algebras, particularly matrix algebras, quaternion algebras, and their generalizations. We begin with basic definitions and examples, and then proceed to structural results such as the Skolem-Noether theorem, the Artin-Wedderburn theorem, and the splitting fields of central simple algebras. The material presented here is based on [7] and [20]. Throughout this chapter, we assume $\mathbb{F}$ is a field of arbitrary characteristic unless stated otherwise.

2.1 Definitions and Examples

Definition 2.1.1. An **associative algebra** over a field $\mathbb{F}$ is a vector space $\mathscr{A}$ over $\mathbb{F}$ equipped with a bilinear map

$$\mu : \mathscr{A} \times \mathscr{A} \longrightarrow \mathscr{A}, \quad (x,y) \mapsto xy$$

called the multiplication, such that the following conditions hold:

1. **Associativity:**

$$(xy)z = x(yz) \quad \text{for all} \quad x, y, z \in \mathscr{A}$$

2. **Compatibility:**

$$(\lambda x)y = x(\lambda y) = \lambda(xy) \quad \text{for all} \quad \lambda \in \mathbb{F}, x, y \in \mathscr{A}.$$

Definition 2.1.2. A **non-associative algebra** over a field $\mathbb{F}$ is a vector space $\mathscr{A}$ over $\mathbb{F}$ equipped with a bilinear map

$$\mu : \mathscr{A} \times \mathscr{A} \longrightarrow \mathscr{A}, \quad (x, y) \mapsto xy,$$

called the multiplication, such that the compatibility condition

$$(\lambda x)y = x(\lambda y) = \lambda(xy) \quad \text{for all } \lambda \in \mathbb{F},\ x, y \in \mathscr{A}$$

holds. No associativity condition is imposed; that is, in general,

$$(xy)z \neq x(yz) \quad \text{for } x, y, z \in \mathscr{A}$$

might not hold.

Remark 2.1.3. If there exists an element $e \in \mathscr{A}$ such that

$$ex = xe = x \quad \text{for all} \quad x \in \mathscr{A},$$

then $\mathscr{A}$ is called a **unital algebra** and e is called the identity element.

Definition 2.1.4. An algebra $\mathscr{A}$ is **simple** if it has no non-trivial two-sided ideals, i.e., the only ideals are $\{0\}$ and $\mathscr{A}$ itself.

Definition 2.1.5. The center of an algebra $\mathscr{A}$ is defined as

$$\mathscr{Z}(\mathscr{A}) = \{z \in \mathscr{A} : za = az \quad \text{for all} \quad a \in \mathscr{A}\}.$$

Definition 2.1.6. An $\mathbb{F}$-algebra $\mathscr{A}$ is called **central simple** if it satisfies the following conditions:

1. $\mathscr{A}$ is associative and finite-dimensional over $\mathbb{F}$,
2. $\mathscr{A}$ is simple,
3. $\mathscr{A}$ is central, that is, $\mathscr{Z}(\mathscr{A}) = \mathbb{F}$.

Example 2.1.7. The field $\mathbb{F}$, viewed as an $\mathbb{F}$-algebra, is a central simple algebra.

Example 2.1.8 (Quaternion Algebra)**.** For $a, b \in \mathbb{F}^\times$, let

$$i = \begin{pmatrix} 0 & a & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & a \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad j = \begin{pmatrix} 0 & 0 & b & 0 \\ 0 & 0 & 0 & -b \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

be elements of $M_4(\mathbb{F})$. These matrices satisfy the relations

$$i^2 = aI_4, \quad j^2 = bI_4, \quad \text{and} \quad ij = -ji.$$

Then the subalgebra generated by i and j is a central simple algebra of dimension 4 over $\mathbb{F}$, where the characteristic of $\mathbb{F} \neq 2$. The algebra generated by i and j, denoted by $(a,b)_{\mathbb{F}}$, is called the **quaternion algebra**.

Example 2.1.9 (Symbol Algebra)**.** Let $\mathbb{F}$ be a field containing a primitive n-th root of unity ζ_n, and let $a,b \in \mathbb{F}^{\times}$. The *symbol algebra* $(a,b)_{\zeta_n,\mathbb{F}}$ is the associative $\mathbb{F}$-algebra generated by two elements x and y subject to the relations

$$x^n = a, \qquad y^n = b, \qquad yx = \zeta_n xy.$$

As a vector space over $\mathbb{F}$, the algebra $(a,b)_{\zeta_n,\mathbb{F}}$ has basis

$$\{x^i y^j \mid 0 \leq i, j \leq n-1\},$$

and hence

$$\dim_{\mathbb{F}}(a,b)_{\zeta_n,\mathbb{F}} = n^2.$$

The center of $(a,b)_{\zeta_n,\mathbb{F}}$ is $\mathbb{F}$, and the algebra is central simple over $\mathbb{F}$.

Remark 2.1.10. When $n = 2$ and $\zeta_2 = -1$, the symbol algebra $(a,b)_{2,\mathbb{F}}$ is generated by elements i and j satisfying

$$i^2 = a, \qquad j^2 = b, \qquad ij = -ji.$$

In this case, $(a,b)_{\zeta,2}$ coincides with the quaternion algebra $(a,b)_{\mathbb{F}}$. Thus quaternion algebras are precisely the degree 2 symbol algebras.

Example 2.1.11. Let D be a finite-dimensional division algebra with center $\mathscr{Z}(D)$. Then the matrix algebra $M_n(D)$ is a central simple algebra over $\mathscr{Z}(D)$.

Proof. Since D is finite-dimensional over $\mathscr{Z}(D)$, we have

$$\dim_{\mathscr{Z}(D)} M_n(D) = n^2 \dim_{\mathscr{Z}(D)} D < \infty.$$

Thus, $M_n(D)$ is finite-dimensional as a $\mathscr{Z}(D)$-algebra.

Let I be a two-sided nonzero ideal of $M_n(D)$. Choose a nonzero matrix

$$X = (x_{ij}) \in I.$$

Then there exist indices k,l such that $x_{kl} \neq 0$. Since D is a division algebra, x_{kl} is invertible. For arbitrary i, j, we compute

$$E_{ik} X E_{lj} = x_{kl} E_{ij} \in I.$$

Multiplying by x_{kl}^{-1}, we obtain $E_{ij} \in I$ for all i, j. Hence I contains all matrix units and therefore

$$I = M_n(D).$$

Thus $M_n(D)$ is simple.

Let $A = (a_{ij}) \in \mathscr{Z}(M_n(D))$. Commuting with all diagonal matrix units forces

$$a_{ij} = 0 \quad \text{for } i \neq j,$$

and commuting with off-diagonal matrix units implies

$$a_{11} = a_{22} = \cdots = a_{nn} = d$$

for some $d \in D$. Thus $A = dI_n$.

Finally, commutativity with all matrices in $M_n(D)$ implies that d lies in the center of D, hence $d \in \mathscr{Z}(D)$. Therefore,

$$\mathscr{Z}(M_n(D)) = \{dI_n \mid d \in Z(D)\} \cong \mathscr{Z}(D).$$

□

The matrix algebra $M_n(\mathbb{F})$ is a central simple $\mathbb{F}$-algebra. The proof follows from above by $D = \mathbb{F}$.

Definition 2.1.12. A central simple algebra is called **split** if it is isomorphic to a matrix algebra.

One of the principal tools employed in this thesis to study the image of a polynomial map is the automorphism group of the algebra under consideration. In the following section, we establish the Skolem–Noether theorem. The proof closely follows the approach in Chapter 3 of [7].

2.2 Skolem-Noether Theorem

Theorem 2.2.1 (Skolem-Noether)**.** *Let $\mathscr{A}$ be a finite-dimensional central simple algebra over a field $\mathbb{F}$. Let $B \subseteq \mathscr{A}$ be a simple $\mathbb{F}$-subalgebra, and let*

$$\varphi : B \to \mathscr{A}$$

be an $\mathbb{F}$-algebra homomorphism. Then there exists an invertible element $u \in A^\times$ such that

$$\varphi(b) = ubu^{-1} \quad \text{for all } b \in B.$$

Proof. Since $\mathscr{A}$ is a finite-dimensional central simple algebra over $\mathbb{F}$, it is simple as a ring. Let $B \subseteq A$ be a simple $\mathbb{F}$-subalgebra, and let $\varphi : B \to A$ be an $\mathbb{F}$-algebra homomorphism.

We consider $\mathscr{A}$ as a left B-module in two different ways. First, $\mathscr{A}$ is a left B-module via left multiplication. Second, we define a left B-module structure on $\mathscr{A}$ by letting

$$b \cdot a = \varphi(b)a \quad \text{for } b \in B, \ a \in A.$$

We denote this twisted module by ${}_{\varphi}\mathscr{A}$.

Since B is a finite-dimensional simple $\mathbb{F}$-algebra, it is semisimple, and hence both $\mathscr{A}$ and ${}_{\varphi}\mathscr{A}$ are semisimple left B-modules. Moreover, they have the same finite dimension over $\mathbb{F}$, so they are isomorphic as left B-modules. Let

$$\psi : \mathscr{A} \to \mathscr{A}$$

be an isomorphism of left B-modules from A to ${}_{\varphi}A$.

The B-linearity of ψ implies that for all $b \in B$ and $a \in \mathscr{A}$,

$$\psi(ba) = \varphi(b)\psi(a).$$

Set

$$u = \psi(1) \in \mathscr{A}.$$

Then for any $b \in B$, we have

$$\begin{aligned} \varphi(b)u &= \varphi(b)\psi(1) \\ &= \psi(b \cdot 1) \\ &= \psi(b) \end{aligned}$$

and also

$$\begin{aligned} \psi(b) &= \psi(1 \cdot b) \\ &= \psi(1)b \\ &= ub. \end{aligned}$$

Therefore,

$$\varphi(b)u = ub \quad \text{for all } b \in B.$$

Since ψ is an isomorphism, $u \neq 0$. Moreover, the identity $\varphi(b)u = ub$ implies that u implements an isomorphism of B-modules, which forces u to be invertible. Hence $u \in \mathscr{A}^{\times}$. Multiplying the above identity on the right by u^{-1} yields

$$\varphi(b) = ubu^{-1} \quad \text{for all } b \in B.$$

This proves the theorem. □

Corollary 2.2.2. Let $\mathscr{A}$ be a finite-dimensional central simple algebra over a field $\mathbb{F}$. Then every $\mathbb{F}$-algebra automorphism of $\mathscr{A}$ is inner. That is, for every

$$\varphi \in \mathrm{Aut}_{\mathbb{F}}(\mathscr{A}),$$

there exists $u \in \mathscr{A}^{\times}$ such that

$$\varphi(a) = uau^{-1} \quad \text{for all } a \in \mathscr{A}.$$

Proof. Apply the Skolem–Noether theorem with $B = \mathscr{A}$. Since $\mathscr{A}$ is central and simple, all the hypotheses are satisfied, and the result follows immediately. □

Remark 2.2.3. The Skolem–Noether theorem shows that all $\mathbb{F}$-algebra automorphisms of a central simple algebra are inner. This result is fundamental in the structure theory of central simple algebras and will be used throughout this thesis to study properties invariant under algebra automorphisms.

The objective of this thesis is to investigate the image structure of polynomial maps evaluated on central simple algebras and on certain non-associative algebras. Since central simple algebras are naturally studied up to algebra isomorphism, results obtained for a representative algebra extend to all algebras within the same isomorphism class. This perspective motivates the use of the Artin–Wedderburn theorem, which classifies central simple algebras up to isomorphism.

2.3 The Artin-Wedderburn Theorem

The Artin–Wedderburn theorem describes the structure of finite-dimensional semisimple algebras over a field and plays a fundamental role in the theory of central simple algebras. A general reference for the material in this section is [7]. The statement of the theorem below follows [15, Chapter 5].

Theorem 2.3.1 (Artin-Wedderburn)**.** *Let $\mathscr{A}$ be a finite-dimensional semisimple algebra over a field $\mathbb{F}$. Then there exist division algebras $D_1,\ldots,D_r$, each finite-dimensional over $\mathbb{F}$, and positive integers $n_1,\ldots,n_r$ such that*

$$\mathscr{A} \cong \bigoplus_{i=1}^{r} M_{n_i}(D_i)$$

as $\mathbb{F}$-algebras.

In particular, if $\mathscr{A}$ is a finite-dimensional simple algebra over $\mathbb{F}$, then there exist a division algebra D finite-dimensional over $\mathbb{F}$ and an integer $n \geq 1$ such that

$$\mathscr{A} \cong M_n(D).$$

Moreover, if $\mathscr{A}$ is a central simple algebra over $\mathbb{F}$, then the division algebra D is central over $\mathbb{F}$, that is,

$$\mathscr{Z}(D) = \mathbb{F}.$$

Since a substantial part of this thesis focuses on central simple algebras over an algebraically closed field, this structural simplification plays a central role in our study. As a consequence of the Artin-Wedderburn theorem, every central simple algebra over an algebraically closed field is isomorphic to a full matrix algebra. The precise statement is given below.

Corollary 2.3.2. Let $\mathbb{F}$ be an algebraically closed field, and let $\mathscr{A}$ be a finite-dimensional central simple algebra over $\mathbb{F}$. Then there exists an integer $n \geq 1$ such that

$$\mathscr{A} \cong M_n(\mathbb{F}).$$

Proof. By the Artin-Wedderburn theorem, there exist a finite-dimensional division algebra D over $\mathbb{F}$ and an integer $n \geq 1$ such that

$$\mathscr{A} \cong M_n(D).$$

Since $\mathscr{A}$ is central over $\mathbb{F}$, we have $\mathscr{Z}(D) = \mathbb{F}$. As $\mathbb{F}$ is algebraically closed, the only finite-dimensional division algebra over $\mathbb{F}$ is $\mathbb{F}$ itself. Hence $D = \mathbb{F}$, and the result follows. □

In the next section, we discuss the behavior of central simple algebras under scalar extension and use it to define the degree and dimension of such algebras.

2.4 Degree and Dimension

Let $\mathscr{A}$ be a finite-dimensional $\mathbb{F}$-algebra, and let $L/\mathbb{F}$ be a field extension. The scalar extension of $\mathscr{A}$ is defined by

$$\mathscr{A}_L := \mathscr{A} \otimes_{\mathbb{F}} L,$$

which is an L-algebra. The algebra $\mathscr{A}_L = \mathscr{A} \otimes_{\mathbb{F}} L$ consists of finite sums

$$x = \sum_{i=1}^{m} a_i \otimes \lambda_i, \qquad a_i \in \mathscr{A},\ \lambda_i \in L.$$

Addition in $\mathscr{A}_L = \mathscr{A} \otimes_{\mathbb{F}} L$ is defined by linearity. For simple tensors, it is given by

$$(a \otimes \lambda) + (a' \otimes \lambda') = a \otimes \lambda + a' \otimes \lambda',$$

and this rule extends to arbitrary elements. The multiplication on $\mathscr{A}_L$ is determined by the rule

$$(a_1 \otimes \lambda_1)(a_2 \otimes \lambda_2) = (a_1 a_2) \otimes (\lambda_1 \lambda_2),$$

and is extended bilinearly to all elements of $\mathscr{A}_L$. Consequently, the product of two general elements

$$\sum_i a_i \otimes \lambda_i.$$

With this multiplication, $\mathscr{A}_L$ becomes an L-algebra, and $\mathscr{A}$ embeds naturally into $\mathscr{A}_L$ via the map $a \mapsto a \otimes 1$.

Before moving to the result, we state the following proposition. The proposition describes the center of a tensor product of finite-dimensional simple algebras. This result will be used to deduce the centrality of scalar extensions of central simple algebras. The reference for the proposition is [20, Chapter IV]

Proposition 2.4.1. *Let $\mathscr{A}$ and $\mathscr{B}$ be finite-dimensional simple algebras over a field $\mathbb{F}$. Then*

$$\mathscr{Z}_{\mathscr{A}\otimes\mathscr{B}}(\mathscr{A} \otimes_{\mathbb{F}} \mathscr{B}) = \mathscr{Z}_{\mathscr{A}}(\mathscr{A}) \otimes_{\mathbb{F}} \mathscr{Z}_{\mathscr{B}}(\mathscr{B}).$$

Proposition 2.4.2. *Let $\mathscr{A}$ be a finite-dimensional central simple algebra over $\mathbb{F}$, and let $L/\mathbb{F}$ be any field extension. Then $\mathscr{A} \otimes_{\mathbb{F}} L$ is a finite-dimensional central simple algebra over L.*

Proof. Since $\mathscr{A}$ is finite-dimensional over $\mathbb{F}$, the algebra $\mathscr{A} \otimes_{\mathbb{F}} L$ is finite-dimensional over L.

Simplicity. Let I be a nonzero two-sided ideal of $\mathscr{A} \otimes_{\mathbb{F}} L$. Choose a nonzero element

$$x = \sum_{i=1}^{m} a_i \otimes \lambda_i \in I$$

such that m is minimal among all such expressions, and such that $\lambda_1, \ldots, \lambda_m$ are linearly independent over $\mathbb{F}$. In particular, we may assume $a_1 \neq 0$.

Since $\mathscr{A}$ is simple, the two-sided ideal generated by a_1 is $\mathscr{A}$. Hence there exist elements $u_j, v_j \in \mathscr{A}$ such that

$$\sum_j u_j a_1 v_j = 1.$$

Multiplying x on the left and right by elements of $\mathscr{A} \otimes 1$, we obtain an element

$$y = 1 \otimes \lambda_1 + \sum_{i=2}^{m} b_i \otimes \lambda_i \in I,$$

for suitable $b_i \in \mathscr{A}$.

Suppose that $b_k \neq 0$ for some $k \geq 2$. Since $\mathscr{A}$ is simple, the two-sided ideal of $\mathscr{A}$ generated by b_k is equal to $\mathscr{A}$. Hence there exist elements $u_j, v_j \in \mathscr{A}$ such that

$$\sum_j u_j b_k v_j = 1.$$

Consider the element

$$z := \sum_j (u_j \otimes 1)\, y\, (v_j \otimes 1) \in I.$$

Then z can be written in the form

$$z = c \otimes \lambda_1 + 1 \otimes \lambda_k + \sum_{i \neq 1,k} c_i \otimes \lambda_i,$$

for some $c, c_i \in \mathscr{A}$.

Since $\lambda_1, \ldots, \lambda_m$ are linearly independent over $\mathbb{F}$, we may subtract a suitable element of the form $(a \otimes 1)x(b \otimes 1)$ from z so as to eliminate the term $1 \otimes \lambda_k$ without affecting the remaining coefficients. The resulting element is a nonzero element of I that can be expressed as a sum of strictly fewer than m simple tensors, which contradicts the minimality of m.

Hence $b_i = 0$ for all $i \geq 2$, and therefore

$$y = 1 \otimes \lambda$$

for some $\lambda \in L^\times$. Multiplying by $1 \otimes \lambda^{-1}$, we conclude that $1 \otimes 1 \in I$, and consequently

$$I = \mathscr{A} \otimes_{\mathbb{F}} L.$$

Central. By Proposition 2.4.1,

$$\begin{aligned} \mathscr{Z}_{\mathscr{A} \otimes L}(\mathscr{A} \otimes_{\mathbb{F}} L) &= \mathscr{Z}_{\mathscr{A}}(\mathscr{A}) \otimes_{\mathbb{F}} \mathscr{Z}_L(L) \\ &= \mathscr{Z}_{\mathscr{A}}(\mathscr{A}) \otimes_{\mathbb{F}} L \end{aligned}$$

Since $\mathscr{A}$ is a central simple algebra over $\mathbb{F}$, we have $\mathscr{Z}(\mathscr{A}) = \mathbb{F}$, and since L is a field, $\mathscr{Z}(L) = L$. Therefore,

$$\mathscr{Z}(\mathscr{A} \otimes_{\mathbb{F}} L) = \mathbb{F} \otimes_{\mathbb{F}} L \cong L.$$

Hence $\mathscr{A} \otimes_{\mathbb{F}} L$ is central over L. □

Theorem 2.4.1 (Descent of Central Simplicity)**.** *Let $\mathscr{A}$ be a finite-dimensional $\mathbb{F}$-algebra and let $L/\mathbb{F}$ be a field extension. If*

$$\mathscr{A} \otimes_{\mathbb{F}} L$$

is a central simple algebra over L, then $\mathscr{A}$ is a central simple algebra over $\mathbb{F}$.

Proof. **Simplicity.** Let I be a nonzero two-sided ideal of $\mathscr{A}$. Then

$$I \otimes_{\mathbb{F}} L$$

is a nonzero two-sided ideal of $\mathscr{A} \otimes_{\mathbb{F}} L$, contradicting the simplicity of $\mathscr{A} \otimes_{\mathbb{F}} L$. Hence $\mathscr{A}$ is simple.

Centrality. By Proposition 2.4.1, we have

$$\mathscr{Z}(\mathscr{A} \otimes_{\mathbb{F}} L) = \mathscr{Z}(\mathscr{A}) \otimes_{\mathbb{F}} \mathscr{Z}(L).$$

Since $\mathscr{A} \otimes_{\mathbb{F}} L$ is central over L, we have

$$\mathscr{Z}(\mathscr{A} \otimes_{\mathbb{F}} L) = L.$$

Moreover, $\mathscr{Z}(L) = L$, and therefore

$$\mathscr{Z}(\mathscr{A}) \otimes_{\mathbb{F}} L \cong L.$$

This is possible only if

$$\mathscr{Z}(\mathscr{A}) = \mathbb{F}.$$

Hence $\mathscr{A}$ is central over $\mathbb{F}$.

Therefore, $\mathscr{A}$ is a central simple algebra over $\mathbb{F}$. □

Corollary 2.4.2. A $\mathbb{F}$-algebra $\mathscr{A}$ is central simple over $\mathbb{F}$ if and only if

$$\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}}$$

is isomorphic to a full matrix algebra over the algebraic closure $\overline{\mathbb{F}}$.

Proof. If $\mathscr{A}$ is central simple over $\mathbb{F}$, then $\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}}$ is central simple over $\overline{\mathbb{F}}$. Since $\overline{\mathbb{F}}$ is algebraically closed, the Artin–Wedderburn theorem implies

$$\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}} \cong M_n(\overline{\mathbb{F}})$$

for some n.

Conversely, if $\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}}$ is a matrix algebra, then it is central simple over $\overline{\mathbb{F}}$, and by descent, $\mathscr{A}$ is central simple over $\mathbb{F}$. □

Let $\mathscr{A}$ be a central simple algebra over a field $\mathbb{F}$. Then by above corollary, for some $n \in \mathbb{N}$,

$$\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}} \cong M_n(\overline{\mathbb{F}})$$

Thus,

$$\dim_{\mathbb{F}}(\mathscr{A}) = \dim_{\overline{\mathbb{F}}}(\mathscr{A} \otimes_{\mathbb{F}} \overline{\mathbb{F}}) = \dim_{\overline{\mathbb{F}}}(M_n(\overline{\mathbb{F}}) = n^2$$

We conclude that the dimension of a central simple algebra is always a square of some positive integer. Now, we define the degree associated to a central simple algebra.

Definition 2.4.3. Let $\mathscr{A}$ be a finite-dimensional central simple algebra over a field $\mathbb{F}$. The *degree* of $\mathscr{A}$, denoted by $\deg(\mathscr{A})$, is defined as

$$\deg(\mathscr{A}) := \sqrt{\dim_{\mathbb{F}} \mathscr{A}}.$$

2.5 Splitting Fields and Maximal Subfields

Let $\mathscr{A}$ be a finite-dimensional central simple algebra over a field $\mathbb{F}$. A central theme in the structure theory of such algebras is that, after extending scalars to a suitable field extension, $\mathscr{A}$ becomes a matrix algebra. This motivates the notion of splitting fields.

Definition 2.5.1. A field extension $L/\mathbb{F}$ is called a *splitting field* of a central simple algebra $\mathscr{A}$ if

$$\mathscr{A} \otimes_{\mathbb{F}} L \cong M_n(L)$$

for some integer $n \geq 1$.

By the Artin–Wedderburn theorem, every finite-dimensional central simple algebra admits a splitting field. Moreover, the integer n is uniquely determined and is called the *degree* of $\mathscr{A}$. Thus, splitting fields allows one to reduce questions about $\mathscr{A}$ to questions about matrix algebras over fields.

Definition 2.5.2. Let $\mathscr{A}$ be a finite-dimensional central simple algebra over $\mathbb{F}$. A subfield $K \subseteq \mathscr{A}$ is called a *maximal subfield* of $\mathscr{A}$ if K is a field extension of $\mathbb{F}$ contained in $\mathscr{A}$ and

$$[K:\mathbb{F}] = \deg(\mathscr{A}).$$

2.5.1 Maximal Subfields as Splitting Fields

The relation between maximal subfields and splitting fields is fundamental.

Proposition 2.5.1. *Let $\mathscr{A}$ be a finite-dimensional central simple algebra over $\mathbb{F}$, and let $K \subseteq \mathscr{A}$ be a maximal subfield. Then $K/\mathbb{F}$ is a splitting field of $\mathscr{A}$, that is,*

$$\mathscr{A} \otimes_{\mathbb{F}} K \cong M_n(K),$$

where $n = \deg(\mathscr{A})$.

Remark 2.5.3. The converse of the above proposition is not true in general. Indeed, any algebraically closed field containing $\mathbb{F}$ is a splitting field for $\mathscr{A}$, but such a field need not embed into $\mathscr{A}$ and hence need not be a maximal subfield.

Since every finite-dimensional central simple algebra over $\mathbb{F}$ is isomorphic to a matrix algebra over a central division algebra, it suffices to study maximal subfields of division algebras. In particular, if $\mathscr{D}$ is a central division algebra over $\mathbb{F}$, then any maximal subfield of $\mathscr{D}$ is a splitting field of $\mathscr{D}$.

As an immediate consequence, every finite-dimensional central simple algebra over $\mathbb{F}$ admits a maximal subfield.

Proposition 2.5.2. *Let $\mathscr{D}$ be a finite-dimensional central division algebra over $\mathbb{F}$, and let $K \subseteq \mathscr{D}$ be a field containing $\mathbb{F}$. Then the following statements are equivalent:*

1. *K is a maximal subfield of $\mathscr{D}$.*
2. *$C_{\mathscr{D}}(K) = K$, where $C_{\mathscr{D}}(K)$ denotes the centralizer of K in $\mathscr{D}$.*
3. *$[K : \mathbb{F}] = \deg(\mathscr{D})$.*
4. *K is a splitting field of $\mathscr{D}$, that is,*

$$\mathscr{D} \otimes_{\mathbb{F}} K \cong M_{\deg(\mathscr{D})}(K).$$

Thus, for central division algebras, the notions of maximal subfield and splitting field coincide. This fact plays a crucial role in understanding the structure and invariants of division algebras, such as degree and index, and allows one to study them via suitable field extensions.

Corollary 2.5.4. Every finite-dimensional central division algebra over $\mathbb{F}$ admits a maximal subfield.

Corollary 2.5.5. Let $\mathscr{A}$ be a finite-dimensional central simple algebra over $\mathbb{F}$. Then $\mathscr{A}$ admits a maximal subfield.

Proof. By the Artin-Wedderburn theorem, there exists a finite-dimensional central division algebra $\mathscr{D}$ over $\mathbb{F}$ and an integer $n \geq 1$ such that

$$\mathscr{A} \cong M_n(\mathscr{D}).$$

Let $K \subseteq \mathscr{D}$ be a maximal subfield of $\mathscr{D}$, which exists by the theory of central division algebras. Identifying $\mathscr{D}$ with the subalgebra of scalar matrices inside $M_n(\mathscr{D})$, we may view K as a subfield of $\mathscr{A}$.

Suppose $K \subsetneq L \subseteq \mathscr{A}$ is a field. Then $L \cap \mathscr{D}$ is a field strictly containing K, contradicting the maximality of K in $\mathscr{D}$. Hence, K is a maximal subfield of $\mathscr{A}$. □

2.5.2 Reduced Trace and Reduced Norm

Let $\mathscr{A}$ be a finite-dimensional central simple algebra over a field $\mathbb{F}$ of degree n, and let $K \subseteq \mathscr{A}$ be a maximal subfield (so that $[K:\mathbb{F}] = n$ and K is a splitting field). Then the scalar extension

$$\mathscr{A} \otimes_{\mathbb{F}} K \cong M_n(K)$$

allows us to define the *reduced trace* and *reduced norm* of elements of $\mathscr{A}$.

Definition 2.5.6 (Reduced Trace and Norm)**.** Let $a \in \mathscr{A}$, and consider its image in $M_n(K)$ under the isomorphism $\mathscr{A} \otimes_{\mathbb{F}} K \cong M_n(K)$. Then we define:

$$\begin{aligned} \mathrm{Trd}_{\mathscr{A}/\mathbb{F}}(a) &:= \text{trace of } a \text{ in } M_n(K), \\ \mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}(a) &:= \text{determinant of } a \text{ in } M_n(K). \end{aligned}$$

These maps satisfy

1. $\mathrm{Trd}_{\mathscr{A}/\mathbb{F}}(a) \in \mathbb{F}$ and $\mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}(a) \in \mathbb{F}$.
2. $\mathrm{Trd}_{\mathscr{A}/\mathbb{F}}$ is linear over $\mathbb{F}$, and $\mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}$ is multiplicative:

$$\mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}(ab) = \mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}(a)\,\mathrm{Nrd}_{\mathscr{A}/\mathbb{F}}(b).$$

Remark 2.5.7. For a central division algebra $\mathscr{D}$, the reduced norm detects invertibility, that is, $a \in \mathscr{D}^\times$ is invertible if and only if $\mathrm{Nrd}_{\mathscr{D}/\mathbb{F}}(a) \neq 0$. Similarly, the reduced trace provides a natural linear invariant for studying polynomial images, e.g., for trace identities in the sense of Lv́ov-Kaplansky conjecture.

Example 2.5.8 (Quaternion Algebra)**.** Let $(a,b)_{\mathbb{F}}$ be a quaternion algebra over $\mathbb{F}$. Then for $x = x_0 + x_1 i + x_2 j + x_3 ij$, the reduced trace and norm are

$$\begin{aligned} \mathrm{Trd}(x) &= 2x_0, \\ \mathrm{Nrd}(x) &= x_0^2 - ax_1^2 - bx_2^2 + abx_3^2. \end{aligned}$$

These coincide with the standard trace and norm of quaternions described in Chapter 4 and are special cases of the general definitions above.

One of the principal reasons for introducing splitting fields and maximal subfields is that they allow a uniform way to embed a central simple algebra into a matrix algebra over a field extension. If $\mathscr{A}$ is a finite-dimensional central simple algebra over $\mathbb{F}$ and $L/\mathbb{F}$ is a splitting field, then

$$\mathscr{A} \otimes_{\mathbb{F}} L \cong M_n(L)$$

for a uniquely determined integer n. This integer is called the *degree* of $\mathscr{A}$. The degree serves as a measure of the intrinsic size of $\mathscr{A}$ and controls the dimensions of its maximal subfields, all of which have degree n over $\mathbb{F}$. Moreover, embeddings of $\mathscr{A}$ into matrix algebras over splitting fields preserve algebraic properties such as simplicity, centrality, and allow one to study structural questions and invariants, such as traces, norms, and images of polynomial maps, in the well-understood setting of matrices over a field. The degree, therefore, plays a central role in determining both the ambient matrix size and the possible image structure of such maps.

Chapter 3

Composition Algebras

This chapter introduces the theory of composition algebras and establishes the foundational definitions that will serve as the basis for the structural results presented later in the thesis. The exposition largely follows Springer and Veldkamp [58], with standard definitions taken from Grove [26] and Lam [37].

3.1 Bilinear and Quadratic forms

Throughout, let V be a vector space of dimension n over the field $\mathbb{F}$ with $\mathsf{char} \neq 2$.

Definition 3.1.1. A **bilinear form** on V is a function $\mathscr{N} : V \times V \longrightarrow \mathbb{F}$ satisfying the following properties for all $a, a', b, b' \in V$ and $\lambda \in \mathbb{F}$:

1. $\mathscr{N}(\lambda a + a', b) = \lambda \mathscr{N}(a, b) + \mathscr{N}(a', b)$,
2. $\mathscr{N}(a, \lambda b + b') = \mathscr{N}(a, b) + \lambda \mathscr{N}(a, b')$.

A bilinear form $\mathscr{N}$ is **symmetric** if $\mathscr{N}(a, b) = \mathscr{N}(b, a)$ for all $a, b \in V$. A symmetric bilinear form $\mathscr{N}$ on V naturally defines a quadratic form on V

Definition 3.1.2. A **quadratic form** on V is a function $N : V \longrightarrow \mathbb{F}$ satisfying the following properties for all $a, b \in V$:

1. $N(\lambda a) := \lambda^2 N(a)$,

2. the map $\mathscr{N} : V \times V \longrightarrow \mathbb{F}$ given by $\mathscr{N}(a,b) = \frac{1}{2}\big(N(a+b) - N(a) - N(b)\big)$ defines a symmetric bilinear form on V.

Given a symmetric bilinear form $\mathscr{N}$, the associated quadratic form N is given by $N(a) := \mathscr{N}(a,a)$. Since $\mathsf{char}(\mathbb{F}) \neq 2$, there is a one-to-one correspondence between the symmetric bilinear forms and the quadratic forms on V.

A quadratic form (V,N) is **isotropic** if for some non-zero vector $a \in V$, $N(a) = 0$. If no such vector exists, then it is **anisotropic**.

Given a bilinear form $(V,\mathscr{N})$ and a vector $a \in V$, define the linear maps

$$\begin{aligned} l_a : V &\longrightarrow \mathbb{F} & r_a : V &\longrightarrow \mathbb{F} \\ b &\longmapsto \mathscr{N}(a,b) & b &\longmapsto \mathscr{N}(b,a) \end{aligned}$$

The bilinear form $\mathscr{N}$ on V is said to be **non-degenerate** if the only vector $a \in V$ satisfying $l_a = 0$ (equivalently, $r_a = 0$) is $a = 0$. Otherwise, $\mathscr{N}$ is called **degenerate**.

For a linear subspace $\mathscr{W}$ of V, the *orthogonal complement* is defined by

$$\mathscr{W}^{\perp} := \{\, a \in V \;\big|\; \mathscr{N}(a,b) = 0 \text{ for all } b \in \mathscr{W} \,\}.$$

The subspace $\mathscr{W}$ is called **non-singular** if the restriction of $\mathscr{N}$ on $\mathscr{W}$ is non-degenerate. If the subspace $\mathscr{W}$ of V is non-singular, then $V = \mathscr{W} \oplus \mathscr{W}^{\perp}$.

3.2 Composition Algebra

Throughout this section, let $\mathscr{C}$ be an algebra, not necessarily associative, over a field $\mathbb{F}$. We assume that $\mathscr{C}$ is unital, meaning that it contains a multiplicative identity element, which we denote by e.

Definition 3.2.1. A **composition algebra** is a pair $(\mathscr{C},N)$ consisting of an algebra $\mathscr{C}$ together with a non-degenerate quadratic form $N : \mathscr{C} \to \mathbb{F}$ satisfying the *composition property*

$$N(ab) = N(a)N(b) \qquad \text{for all } a,b \in \mathscr{C}.$$

The quadratic form N is called the *norm* of the algebra $\mathscr{C}$. Its associated symmetric bilinear form, obtained by polarization and denoted by $\mathscr{N}$, will be referred to as the *inner product* on $\mathscr{C}$. A non-singular subspace of $\mathscr{C}'$ of $\mathscr{C}$ is called **subalgebra** of a composition algebra if it is closed under multiplication and contains e.

Definition 3.2.2. The **conjugation** in a composition algebra $\mathscr{C}$ is the map

$$^{-} : \mathscr{C} \longrightarrow \mathscr{C}, \qquad a \longmapsto \overline{a} = \mathscr{N}(a,e)\,e - a,$$

where e is the identity element and $\mathscr{N}$ denotes the associated inner product.

Using the definitions introduced above, we recall some fundamental properties of composition algebras, as established in [58, Chapter 1].

Lemma 3.2.1. *Let $a,b \in \mathscr{C}$. The following properties hold:*

1. *Every element a satisfies*

$$a^2 - \mathscr{N}(a,e)\,a + N(a)\,e = 0.$$

2. $a\overline{a} = \overline{a}a = N(a)\,e.$
3. $\overline{\overline{a}} = a.$
4. $a\,(\overline{a}b) = N(a)\,b.$
5. $(a\overline{b})\,b = N(b)\,a.$
6. *The element a is invertible if and only if $N(a) \neq 0$. Moreover, in this case,*

$$a^{-1} = N(a)^{-1}\,\overline{a}.$$

7. $(ab)a = a(ba)$
8. $a(ab) = a^2 b$ *and* $(ab)b = ab^2$

Properties (7) *and* (8) *together are referred to as* alternative laws. *These identities imply power associativity, so that expressions of the form* a^n *are unambiguous.*

3.3 Classification of Composition Algebra

In this section, we recall the doubling process, also known as the Cayley-Dickson process, and then classify composition algebras by dimension.

To understand the structure of composition algebras, it is helpful to view them as obtained by successive doublings. The Cayley–Dickson construction provides a canonical way to enlarge a given composition algebra, and, remarkably, every composition algebra arises from a suitable subalgebra by this process. We first recall a proposition that formalises how to obtain the composition algebra from its composition proper subalgebra; see [58, Proposition 1.5.1].

Proposition 3.3.1. *Let* $(\mathscr{C}, N)$ *be a composition algebra over a field* $\mathbb{F}$*. Suppose* $\mathscr{C}_0 \subseteq \mathscr{C}$ *is a composition proper subalgebra and let* $u \in \mathscr{C}_0^{\perp}$ *be an element such that* $N(u) \neq 0$*. Then the space*

$$\mathscr{C}_1 = \mathscr{C}_0 \oplus \mathscr{C}_0 u$$

forms a composition subalgebra of $\mathscr{C}$ *with* $\dim(\mathscr{C}_1) = 2 \cdot \dim(\mathscr{C}_0)$*. Moreover, for* $a, a', b, b' \in \mathscr{C}_0$*, the operations on* $\mathscr{C}_1$ *are given by:*

$$\begin{aligned} \textit{Multiplication:} \quad & (a+bu)(a'+b'u) = \left(aa' + \alpha \overline{b'} b\right) + \left(b'a + b\overline{a}\right)u, \\ \textit{Norm:} \quad & N(a+bu) = N(a) - \alpha N(b), \\ \textit{Conjugation:} \quad & \overline{a+bu} = \overline{a} - bu, \end{aligned}$$

where $\bar{\cdot}$ *denotes the canonical involution on* $\mathscr{C}_0$ *and* $\alpha = -N(u)$*.*

Given a composition algebra, one can construct a new composition algebra via the doubling process. Below is the proposition from [58, Proposition 1.5.2].

Proposition 3.3.2. *Let $\mathscr{C}_0$ be an associative composition algebra and $\alpha \in \mathbb{F}^\star$. Consider the space $\mathscr{C} = \mathscr{C}_0 \oplus \mathscr{C}_0$. For $a, a', b, b' \in \mathscr{C}_0$, define the product*

$$(a,b)(a',b') = (aa' + \alpha \overline{b'} b, b'a + b\overline{a'})$$

and norm as

$$N((a,b)) = N(a) - \alpha N(b).$$

Then $\mathscr{C}$ is a composition algebra equipped with the quadratic form N. Moreover, $\mathscr{C}$ is associative if and only if $\mathscr{C}$ is commutative and associative.

As a result, starting from $\mathbb{F}$ as a 1-dimensional composition algebra, successive applications of the doubling construction generate the classical finite-dimensional composition algebras. Doubling the 1-dimensional algebra produces a 2-dimensional quadratic algebra, which is commutative and associative. Doubling a quadratic algebra gives a 4-dimensional quaternion algebra, which is non-commutative but associative. A further doubling results in an 8-dimensional octonion algebra, which is neither commutative nor associative, but satisfies the alternative laws. Consequently, the only finite-dimensional composition algebras are of dimension $1,2,4$ and 8.

This classification can be found in [58, Theorem 1.6.2].

Chapter 4

Quaternion and Octonion Algebras

In this chapter, we recall the basic structure and properties of quaternion and octonion algebras over a field $\mathbb{F}$ of characteristic not equal to 2. The quaternion algebra provides one of the earliest non-trivial examples of a central simple algebra. It also serves as a fundamental example of a composition algebra, being non-commutative yet associative. Another important example of a composition algebra, obtained via the Cayley–Dickson doubling process, is the octonion algebra, which is non-commutative and non-associative. We discuss their main properties, as well as the distinction between split and division cases. The exposition in this chapter is largely based on [7] and [58].

4.1 Quaternion Algebra

We start by recalling the definition of quaternion algebra from 2.1.8 introduced in 2. A **quaternion algebra** over $\mathbb{F}$ is a four-dimensional $\mathbb{F}$-algebra of the form

$$\mathscr{Q} = (a,b)_{\mathbb{F}},$$

where $a, b \in \mathbb{F}^{\times}$, generated by elements i, j subject to the relations

$$i^2 = a, \qquad j^2 = b, \qquad ij = -ji.$$

Setting $k = ij$, every element of $\mathscr{Q}$ can be written uniquely as

$$x = x_0 + x_1 i + x_2 j + x_3 k, \qquad x_0, x_1, x_2, x_3 \in \mathbb{F}.$$

Thus, $\dim_{\mathbb{F}} \mathscr{Q} = 4$ and hence the degree of $\mathscr{Q}$ is 2.

4.1.1 Conjugation, Trace, and Norm

For $x \in \mathscr{Q}$, the **conjugate** of x is defined to be

$$\bar{x} = x_0 - x_1 i - x_2 j - x_3 k.$$

Using conjugation, one defines the **reduced trace** and **reduced norm** of an element $x \in \mathscr{Q}$ by

$$Tr(x) = x + \bar{x} = 2x_0, \qquad N(x) = x\bar{x}.$$

Explicitly,

$$N(x) = x_0^2 - a x_1^2 - b x_2^2 + ab x_3^2 \in \mathbb{F}.$$

In the context of composition algebra, $(\mathscr{Q}, N)$ forms a composition algebra of dimension 4 with the non-degenerate form N.

4.1.2 Split and Division Quaternion Algebras

Based on the norm of $\mathscr{Q}$, one can classify the quaternion algebra as either split or division.

Definition 4.1.1. A quaternion algebra $\mathscr{Q}$ over $\mathbb{F}$ is called *split* if $\mathscr{Q}$ if there exist non-zero element $x \in \mathscr{Q}$ such that $N(x) = 0$. Otherwise, it is a division algebra.

In other words, $\mathscr{Q}$ is a division algebra if and only if $N(x) = 0$ implies $x = 0$ for $x \in \mathscr{Q}$. We now show that every split quaternion algebra is isomorphic to a full matrix algebra.

Theorem 4.1.2. *Let $(a,b)_{\mathbb{F}}$ be a split quaternion algebra over $\mathbb{F}$. Then*

$$(a,b)_{\mathbb{F}} \cong M_2(\mathbb{F}).$$

Proof. Since $(a,b)_{\mathbb{F}}$ is split, it is not a division algebra. However, $(a,b)_{\mathbb{F}}$ is a finite-dimensional central simple algebra over $\mathbb{F}$ of dimension 4.

By the Artin-Wedderburn theorem, there exist a division algebra $\mathscr{D}$ over $\mathbb{F}$ and an integer $n \geq 1$ such that

$$(a,b)_{\mathbb{F}} \cong M_n(\mathscr{D}).$$

Taking dimensions over $\mathbb{F}$, we obtain, $4 = n^2 \dim_{\mathbb{F}}(\mathscr{D})$. The only possibility is $n = 2$ and $\dim_{\mathbb{F}}(\mathscr{D}) = 1$, which implies $\mathscr{D} = \mathbb{F}$. It follows that

$$(a,b)_{\mathbb{F}} \cong M_2(\mathbb{F}).$$

□

4.1.3 Splitting Field of Quaternion Algebra

A field extension $L/\mathbb{F}$ is called a *splitting field* of a quaternion algebra $\mathscr{Q}$ if

$$\mathscr{Q} \otimes_{\mathbb{F}} L \cong M_2(L).$$

Let $\mathscr{Q} = (a,b)_{\mathbb{F}}$ be a division algebra. Then the subfields $\mathbb{F}(i)$ and $\mathbb{F}(j)$ are quadratic field extensions of $\mathbb{F}$, since $i^2 = a$ and $j^2 = b$. In particular,

$$\mathbb{F}(i) \cong \mathbb{F}(\sqrt{a}), \qquad \mathbb{F}(j) \cong \mathbb{F}(\sqrt{b}).$$

By the general theory of central division algebras (see Proposition 2.5.2), any maximal subfield of $\mathcal{Q}$ is a splitting field. Hence both $\mathbb{F}(i)$ and $\mathbb{F}(j)$ are splitting fields of $\mathcal{Q}$, and therefore

$$\mathcal{Q} \otimes_{\mathbb{F}} \mathbb{F}(i) \cong M_2(\mathbb{F}(i)), \qquad \mathcal{Q} \otimes_{\mathbb{F}} \mathbb{F}(j) \cong M_2(\mathbb{F}(j)).$$

Consequently, every division quaternion algebra admits a splitting field of degree 2 over its center, and these splitting fields are precisely its maximal subfields.

This observation is particularly useful when studying polynomial maps on $\mathcal{Q}$. By extending scalars to a maximal (equivalently, splitting) field L, one may replace $\mathcal{Q}$ by the matrix algebra $M_2(L)$ and analyze the induced polynomial map in the matrix setting. Properties such as dominance, surjectivity, or restrictions to trace-zero or norm-one elements can then be investigated using tools from matrix theory and subsequently interpreted back in $\mathcal{Q}$.

4.2 Octonion Algebra

A composition algebra of dimension 8, obtained from a quaternion algebra via the Cayley-Dickson doubling process, is called an **octonion algebra**.

The norm on the octonion algebra is induced from the norm on the underlying quaternion composition algebra, as described in Proposition 3.3.2. In particular, if the quaternion algebra is equipped with a multiplicative quadratic norm N, then the doubled algebra inherits a quadratic form, again denoted by N, which satisfies the composition property

$$N(xy) = N(x)N(y), \qquad \text{for all } x, y \in \mathcal{O}.$$

Thus, the algebra $\mathscr{O}$, equipped with the norm N, is an octonion algebra.
As discussed for quaternion algebras, an octonion algebra is either split or a division algebra, depending on the behavior of the norm N.

Proposition 4.2.1. *Let $\mathscr{O}$ be an octonion algebra with norm N. If there exists a nonzero element $x \in \mathscr{O}$ such that $N(x) = 0$, then $\mathscr{O}$ is split. Otherwise, $\mathscr{O}$ is a division algebra.*

In this thesis, we are primarily interested in split octonion algebras; hence, in the remaining chapter, we study the structure of split octonions.

4.2.1 Split Octonion

Instead of describing split octonions via the Cayley–Dickson doubling process, we present their structure using vector matrices over the field $\mathbb{F}$. A description of this can be found in [58, Chapter 1].

Let $\mathbb{F}$ be a field with $\mathsf{char}(\mathbb{F}) \neq 2$. The split octonion algebra $\mathscr{O}$ over $\mathbb{F}$ can be realized as the set of all matrices of the form

$$\mathscr{O}(\mathbb{F}) = \left\{ \begin{pmatrix} \eta & \mathbf{x} \\ \mathbf{y} & \zeta \end{pmatrix} \middle| \eta, \zeta \in \mathbb{F}, \mathbf{x}, \mathbf{y} \in \mathbb{F}^3 \right\}.$$

The addition is defined entry-wise, whereas the multiplication is defined using the following rule

$$\begin{pmatrix} \eta & \mathbf{x} \\ \mathbf{y} & \zeta \end{pmatrix} \begin{pmatrix} \eta' & \mathbf{x}' \\ \mathbf{y}' & \zeta' \end{pmatrix} = \begin{pmatrix} \eta\eta' + \mathbf{x}\cdot\mathbf{y}' & \eta\mathbf{x}' + \zeta'\mathbf{x} + \mathbf{y}\times\mathbf{y}' \\ \zeta\mathbf{y}' + \eta'\mathbf{y} + \mathbf{x}\times\mathbf{x}' & \zeta\zeta' + \mathbf{y}\cdot\mathbf{x}' \end{pmatrix}.$$

Here $\mathbf{x}\cdot\mathbf{y} = \sum_{i=1}^{3} x_i y_i$, where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$, and the exterior power is given by

$$\mathbf{x} \times \mathbf{y} \cdot \mathbf{z} = \det(\mathbf{x}, \mathbf{y}, \mathbf{z}).$$

For an element

$$A = \begin{pmatrix} \eta & \mathbf{x} \\ \mathbf{y} & \zeta \end{pmatrix} \in \mathscr{O}(\mathbb{F}),$$

the *conjugation* is defined by

$$\overline{A} = \begin{pmatrix} \zeta & -\mathbf{x} \\ -\mathbf{y} & \eta \end{pmatrix}.$$

The *norm* of A is given by

$$N(A) = \eta\zeta - \mathbf{x} \cdot \mathbf{y},$$

and the *trace* of A is

$$Tr(A) = \eta + \zeta.$$

With these definitions, one has

$$A\overline{A} = \overline{A}A = N(A)\,I.$$

It is well known that over a field $\mathbb{F}$, there exists, up to isomorphism, a unique split octonion algebra, which is realized by the above construction.

4.2.1.1 Properties

We recall some standard properties of the octonion algebra $\mathscr{O}(\mathbb{F})$ that will be used in later parts of the thesis. Let $A, B \in \mathscr{O}(\mathbb{F})$. Then the following holds.

1. **Power associativity.** For all integers $n \geq 0$, one has

$$A(A^n) = (A^n)A = A^{n+1}.$$

2. **Flexible law.** The octonion algebra satisfies the flexible identity

$$A(BA) = (AB)A.$$

3. **Cancellation by invertible elements.** If $N(A) \neq 0$, then A is invertible and

$$A^{-1}(AB) = B.$$

Chapter 5

Polynomial maps with constants

This chapter introduces the concept of *polynomial maps with constants* and presents the methodology used throughout the thesis to study the questions about surjectivity and image sets obtained through the evaluation of polynomial maps with constants over algebras.

5.1 Polynomial maps

Polynomial maps play a central role in the study of algebraic structures, particularly in understanding how noncommutative polynomials act on algebras. Given an $\mathbb{F}$-algebra $\mathscr{A}$, a polynomial $f(x_1,\dots,x_m)$ in non-commuting variables with coefficients in $\mathbb{F}$ defines a map

$$f\colon \mathscr{A}^m \longrightarrow \mathscr{A}$$
$$(a_1,a_2,\dots,a_m) \mapsto f(a_1,a_2,\dots,a_m)$$

via evaluation. Classical problems in this area concern the description of the image of such maps and criteria for their surjectivity, especially when $\mathscr{A}$ is a matrix algebra or a central simple algebra. Notable examples include multilinear polynomials, whose images are closely related to structural properties of the underlying algebra and appear prominently in conjectures such as the Lv́ov-Kaplansky conjecture.

In many natural situations, however, one is led to consider polynomial expressions that involve fixed elements of the algebra itself as coefficients. This motivates the study of polynomial maps *with constants*, which form the main object of study in this chapter.

5.2 Polynomial maps with constants

For a field $\mathbb{F}$, let $\mathscr{F}_m$ denote the free algebra of rank m over $\mathbb{F}$. Let $\mathscr{A}$ be an $\mathbb{F}$-algebra. Consider the free product $\mathscr{A} \star \mathscr{F}_m$, which forms a free algebra over $\mathbb{F}$. We denote this by $\mathscr{A}_m$. This construction provides a convenient framework for studying polynomial maps over $\mathscr{A}$ in m variables.

Any element $\omega \in \mathscr{A}_m$ can be expressed in the form

$$\sum (a_{j_1}) x_{i_1}^{k_1} (a_{j_2}) x_{i_2}^{k_2} \dots (a_{j_r}) x_{i_r}^{k_r}$$

where $a_{j_s} \in \mathscr{A}$ and $i_s, k_s \in \mathbb{Z}_{\geq 0}$ for all $1 \leq s \leq r$. Each such element ω induces an evaluation map over the algebra $\mathscr{A}$, denoted by $\widetilde{\omega}$, defined as

$$\widetilde{\omega} : \mathscr{A}^n \longrightarrow \mathscr{A}$$
$$(a_1, a_2, \dots, a_n) \mapsto \omega(a_1, a_2, \dots, a_n).$$

The concept serves as an analogue of word maps with constants in group theory, extending the idea to the setting of algebras. In this thesis, we are particularly interested in studying the surjectivity and image sets of such maps.

5.3 Questions

This section formulates the main questions of the thesis, which are addressed in the following chapters for various algebras $\mathscr{A}$. The first question concerns the generalisation of diagonal polynomial maps and is stated as follows:

Question 5.3.1. Let m be a positive integer. Given integers $k_1, k_2, \dots, k_m \geq 1$, and $A_1, \dots, A_m \in \mathscr{A}$ all non-zero, consider the diagonal map $\omega \colon \mathscr{A}^m \to \mathscr{A}$ given by

$$\omega(x_1, \dots, x_m) = A_1(x_1)^{k_1} + A_2(x_2)^{k_2} + \cdots + A_m(x_m)^{k_m}.$$

What is the minimum m for which the given map ω is surjective? If $m = 2$, what are the conditions on $A_1, A_2 \in \mathscr{A}$ for ω to be surjective?

It is well known that any homogeneous polynomial of degree k can be obtained from a multilinear polynomial in m variables via the polarization process. Consequently, understanding the surjectivity of the above diagonal map also provides information in the direction of the Lv́ov–Kaplansky conjecture.

The second question studied in this thesis is stated as:

Question 5.3.2. Let $\mathbb{F}$ be a field and $\mathscr{A} = M_n(\mathbb{F})$ be the matrix algebra. Consider $\omega = A_1x_1x_2 - A_2x_2x_1 \in \mathscr{A}\langle x_1, x_2 \rangle$ for $A_1, A_2 \in M_n(\mathbb{F})$. For what $A_1, A_2 \in \mathscr{A}$, is the image of the map a vector space?

Understanding this problem directly contributes to the study of the Lv́ov-Kaplansky conjecture for matrix algebras of size n, since the polynomial $A_1x_1x_2 - A_2x_2x_1$ is a symmetric multilinear polynomial up to coefficients from $\mathscr{A}$. Determining when the map is surjective provides insight into the general classification of images of multilinear polynomials on matrix algebras.

Throughout the thesis, we will primarily focus on the case $m = 2$, except for certain results where higher numbers of variables are required to study the surjectivity. The next section outlines the general methodology used to study such questions.

5.4 Methodology

To address the surjectivity and image questions for polynomial maps with constants, we adopt a unified approach that applies across the different types of algebras considered in this thesis. We exploit invariants and decomposition techniques to reduce the problem to a more tractable form. By presenting the method here, the reader gains an overview of the tools and strategies that will be elaborated upon in the subsequent chapters.

For non-zero elements $A_1, A_2, \ldots, A_m \in \mathscr{A}$ and $k_1, k_2, \ldots, k_m \geq 1$, consider the polynomial

$$\omega(x_1, \ldots, x_m) = A_1(x_1)^{k_1} + A_2(x_2)^{k_2} + \cdots + A_m(x_m)^{k_m}.$$

With each such map, one can associate the coefficient tuple $(A_1, A_2, \ldots, A_m) \in \mathscr{A}^m$ and consider the diagonal action of the automorphism group $\mathrm{Aut}(\mathscr{A})$ on $\mathscr{A}^m$, given by,

$$\begin{aligned} \mathrm{Aut}(\mathscr{A}) \times \mathscr{A}^m &\longrightarrow \mathscr{A}^m \\ (g, (A_1, A_2, \ldots, A_m)) &\mapsto (g \cdot A_1, g \cdot A_2, \ldots, g \cdot A_m) \end{aligned}$$

If the map ω when evaluated on $\mathscr{A}$ is surjective for a given tuple $(A_1, A_2, \ldots, A_m)$, then the map $g \cdot \omega$ for $g \in \mathrm{Aut}(\mathscr{A})$, with the coefficient tuple $(g \cdot A_1, g \cdot A_2, \ldots, g \cdot A_m)$, is also surjective and vise-versa. Consequently, it suffices to study the surjectivity of the map associated with a representative of the $\mathrm{Aut}(\mathscr{A})$-orbit of the tuple $(A_1, A_2, \ldots, A_m)$.

5.5 Classification problem

The problem of classifying tuples of elements in an algebra up to the natural action of its automorphism group is, in general, a wild problem. In particular, even for well-structured classes of algebras, such classification problems are typically too complicated to admit a complete or manageable description.

In [6], Belitskii and Sergeichuk proved that the problem of classifying pairs of matrices up to simultaneous conjugation contains the problem of classifying m-tuples of matrices under simultaneous conjugation, for arbitrary m. In the next chapter, we present explicit orbit representatives for the action of simultaneous conjugation on pairs of matrices for size 2.

An analogue of this problem, namely, the classification of pairs of elements over the octonions, was studied by Lopatin and Zubkov in [43].

In the later chapters, we make use of such classifications for matrix algebras and split octonion algebras in order to study the images of polynomial maps with constants on these algebras.

Chapter 6

Polynomial maps with Constants: Matrix Algebra

In this chapter, we study the images of polynomial maps with constants, viewed through the broader framework of central simple algebras. Among these, matrix algebras play a fundamental role. Every finite-dimensional central simple algebra over a field is isomorphic to a matrix algebra over a division algebra, and thus $M_n(\mathbb{F})$ serves as the prime and most representative example of a central simple algebra.

An important related structure is the split quaternion algebra, which is isomorphic to the matrix algebra $M_2(\mathbb{F})$. This connection highlights how results obtained for matrix algebras might naturally extend to other split central simple algebras and provides additional intuition about how constants influence the behaviour of polynomial maps in different algebraic settings. To understand the behaviour of polynomial maps with constants evaluated over a central simple algebra, we start with the matrix algebra over an algebraically closed field. Throughout this chapter, we assume $\mathbb{F}$ is an algebraically closed field.

The discussion presented in this chapter is based on our article [47], with necessary modifications and contextual adaptations to align with the broader theme of this thesis. Our focus here is to explore how the introduction of constants affects the structure of the image of a polynomial map.

We investigate the dependence of the image on the choice of constants, analyze whether such images retain invariance under algebra automorphisms, and compare the results with those obtained for constant-free polynomial maps in a later chapter of this thesis.

We prove the following:

Theorem 6.1.1. *Let $\mathbb{F}$ be an algebraically closed field and k_1, k_2 be positive integers. Consider the polynomial map ω given by $A_1x^{k_1} + A_2y^{k_2}$ on $M_2(\mathbb{F})$ where $A_1, A_2 \in M_2(\mathbb{F})$ both non-zero. Then, the image of ω is a vector subspace of $M_2(\mathbb{F})$. Furthermore, ω is surjective if and only if A_1 and A_2 can be simultaneously conjugated to a pair of matrices such that both the matrices do not have the same zero rows.*

Theorem 6.1.2. *Let $\mathbb{F}$ be an algebraically closed field. Consider the polynomial map ω given by $A_1(x_1x_2) - A_2(x_2x_1)$ on $M_2(\mathbb{F})$ where $A_1, A_2 \in M_2(\mathbb{F})$ both non-zero. Then, the image is a vector subspace of $M_2(\mathbb{F})$.*

As discussed in the previous chapter, we deploy the method of simultaneous conjugation to simplify the study of polynomial maps with constants. In general, finding the class representatives for a pair of matrices (A_1, A_2) under simultaneous conjugation is a wild problem, meaning no complete classification exists for arbitrary n. However, for $n = 2$, we give an explicit description of representatives, which allows us to reduce the analysis of the polynomial maps to a finite set of canonical cases. With the help of this, we compute the images for each of those cases and prove Theorem 6.1.1 and Theorem 6.1.2.

6.2 Classification under simultaneous conjugation

Consider the tuple $(A_1,A_2) \in M_2(\mathbb{F})^2$. This section aims to find the representatives of (A_1,A_2) under the conjugation action of $\mathrm{GL}_2(\mathbb{F})$, acting diagonally. We proceed in two steps: first, we find the representative of A_1 under the action of $\mathrm{GL}_2(\mathbb{F})$, say $\widetilde{A}_1$, and then restrict the action of $\mathrm{GL}_2(\mathbb{F})$ to the centralizer of $\widetilde{A}_1$, denoted by $\mathscr{C}_{\mathrm{GL}_2(\mathbb{F})}(\widetilde{A}_1)$, on A_2 to get it's representative.

For each $A_1 \in M_2(\mathbb{F})$, there exists $P_{A_1} \in \mathrm{GL}_2(\mathbb{F})$ such that $P_{A_1}A_1P_{A_1}^{-1} = J_{A_1}$ where J_{A_1} denotes the Jordan canonical form of A_1. We want to comprehend which matrices in $M_2(\mathbb{F})$ lie in the image of map ω when evaluated over $M_2(\mathbb{F})$ i.e. for which C there exists matrices X_1 and X_2 in $M_2(\mathbb{F})$ such that

$$C = \omega(X_1,X_2). \tag{6.2.1}$$

Let $T \in \mathscr{C}_{\mathrm{GL}_2}(J_{A_1})$. Then $TP_{A_1} \in \mathrm{GL}_2(\mathbb{F})$. Conjugating Equation 6.2.1 by TP_1, we get,

$$TP_{A_1}(C)P_{A_1}^{-1}T^{-1} = TP_{A_1}\omega(X_1,X_2)P_{A_1}^{-1}T^{-1}.$$

For both maps considered in this chapter, we obtain

$$\begin{aligned} TP_{A_1}(C)P_{A_1}^{-1}T^{-1} &= TP_{A_1}(A_1)P_{A_1}^{-1}T^{-1}(X_1^{k_1}) + TP_{A_1}(A_2)P_{A_1}^{-1}T^{-1}(X_2^{k_2}) \\ &= J_{A_1}X_1^{k_1} + (T\widetilde{A}_2T^{-1})X_2^{k_2} \end{aligned}$$

and

$$\begin{aligned} TP_{A_1}(C)P_{A_1}^{-1}T^{-1} &= TP_{A_1}(A_1)P_{A_1}^{-1}T^{-1}(X_1X_2) - TP_{A_1}(A_2)P_{A_1}^{-1}T^{-1}(X_2X_1) \\ &= J_{A_1}(X_1X_2) - (T\widetilde{A}_2T^{-1})(X_2X_1) \end{aligned}$$

where $\widetilde{A}_2 = P_{A_1} A_2 P_{A_1}^{-1}$.

Thus, to obtain representatives, it is enough to consider the action of the centralizer $\mathscr{C}_{GL_2}(J_{A_1})$ on $\widetilde{A}_2$ and determine the orbit representatives for each type of Jordan form of A_1. The Jordan form of $A_1 \in M_2(\mathbb{F})$ is one of the following:

$$\begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}_{\lambda \in k^\times} \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}_{\lambda \neq \mu} \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}_{\lambda \in k}.$$

The Table 6.2.1 describes the centralizer corresponding to each Jordan form.

J_{A_1}	$\mathscr{C}_{\mathrm{GL}_2}(J_{A_1})$
$\begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}_{\lambda \neq 0}$	$\mathrm{GL}_2(\mathbb{F})$
$\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}$	$\left\{ \begin{pmatrix} d_1 & \\ & d_2 \end{pmatrix} \middle\vert d_1 d_2 \neq 0 \right\}$
$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}$	$\left\{ \begin{pmatrix} a_1 & b_1 \\ & a_1 \end{pmatrix} \middle\vert a_1 \neq 0 \right\}$

Table 6.2.1: Description of the centralizers

Let $\widetilde{A}_2 = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in M_2(\mathbb{F})$. The representatives of $\widetilde{A}_2$ under the conjugation action are given as below.

(A) For $\mathscr{C}_{\mathrm{GL}_2} = \mathrm{GL}_2$, the representatives are

$$\begin{pmatrix} \mu_1 & \\ & \mu_1 \end{pmatrix}_{\mu_1 \in \mathbb{F}^\times}, \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2}, \begin{pmatrix} \mu_1 & 1 \\ & \mu_1 \end{pmatrix}_{\mu_1 \in \mathbb{F}}.$$

(B) For $\mathscr{C}_{\mathrm{GL}_2} = \left\{ \begin{pmatrix} d_1 & \\ & d_2 \end{pmatrix} \middle| d_1 d_2 \neq 0 \right\}$ and $T \in \mathscr{C}_{\mathrm{GL}_2}$, we have

$$T\widetilde{A}_2 T^{-1} = \begin{pmatrix} a' & d_1 d_2^{-1} b' \\ d_1^{-1} d_2 c' & d' \end{pmatrix}.$$

For $c' \neq 0$, let $T = \begin{pmatrix} c' & \\ & 1 \end{pmatrix}$. If $c' = 0$, then $T = \begin{pmatrix} 1 & \\ & b' \end{pmatrix}$. Thus, the representatives obtained are

$$\begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_1 \end{pmatrix}_{\mu_1 \in \mathbb{F}^\times}, \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2}, \begin{pmatrix} \mu_1 & 1 \\ 0 & \mu_1 \end{pmatrix}_{\mu_1 \in \mathbb{F}}, \begin{pmatrix} \mu_1 & 1 \\ 0 & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2}$$

$$\begin{pmatrix} \mu_1 & 0 \\ 1 & \mu_1 \end{pmatrix}_{\mu_1 \in \mathbb{F}}, \quad \begin{pmatrix} \mu_1 & 0 \\ 1 & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2}, \quad \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix}_{z_i \neq 0}, \quad \begin{pmatrix} 0 & z_2 \\ 1 & z_3 \end{pmatrix}_{z_3 \neq 0}$$

$$\begin{pmatrix} z_1 & z_2 \\ 1 & 0 \end{pmatrix}_{z_1 \neq 0}, \quad \begin{pmatrix} 0 & z_2 \\ 1 & 0 \end{pmatrix}_{z_2 \neq 0}.$$

(C) For $\mathscr{C}_{\mathrm{GL}_2} = \left\{ \begin{pmatrix} a_1 & b_1 \\ 0 & a_1 \end{pmatrix} \middle| a_1 \neq 0 \right\}$. Then TB_AT^{-1} is

$$\begin{pmatrix} a' + c'a_1^{-1}b_1 & (d'-a')\,a_1^{-1}b_1 + b' - c'(a_1^{-1}b_1)^2 \\ c' & d' - c'a_1^{-1}b_1 \end{pmatrix}$$

where $T \in \mathscr{C}_{GL_2}$. Let $a_1^{-1}b_1 = \beta$. Then

$$T\widetilde{A}_2T^{-1} = \begin{pmatrix} a' + c'x & (d'-a')x + b' - c'x^2 \\ c' & d' - c'x \end{pmatrix}.$$

If $c' \neq 0$, then, since $\mathbb{F}$ being an algebraically closed field there exists β such that

$$(d'-a')\beta + b' - c'\beta^2 = 0.$$

If $c' = 0$ and $d' \neq a'$, then there exist β such that $(d'-a')x + b' = 0$. Thus, considering these cases, the representatives are given by

$$\begin{pmatrix} \mu_1 & \\ z & \mu_1 \end{pmatrix}_{\mu_1, z \in \mathbb{F}^\times}, \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2, z \in \mathbb{F}^\times}, \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}_{\mu_1 \neq \mu_2 \in \mathbb{F}}, \begin{pmatrix} \mu_1 & z \\ & \mu_1 \end{pmatrix}_{\mu_1 z \in \mathbb{F}^\times}.$$

Remark 6.2.1. From the above discussion, we see that the surjectivity of polynomial maps $\omega \in \mathscr{A}\langle x_1, x_2 \rangle$ of the form $A_1 x_1^{k_1} + A_2 x_2^{k_2}$ and $A_1 x_1 x_2 - A_2 x_2 x_1$ is preserved under simultaneous conjugation. Moreover, if the image of ω corresponding to the tuple (A_1', A_2') obtained under simultaneous conjugation is a vector subspace, then so is the image of ω corresponding to the original tuple (A_1, A_2). Consequently, it suffices to study the maps for the reduced forms

$$J_{A_1} x_1^{k_1} + A_2' x_2^{k_2} \quad \text{and} \quad J_{A_1} x_1 x_2 - A_2' x_2 x_1,$$

where A_2' is a representative of A_2 obtained under the conjugation action induced by an automorphism sending A_1 to its Jordan canonical form J_{A_1}, followed by the conjugation action of the centralizer $\mathscr{C}_{\mathrm{GL}_2(\mathbb{F})}(J_{A_1})$.

Remark 6.2.2. Note that the argument for surjectivity is symmetric with respect to the representatives of (A_1, A_2). If the polynomial associated to (J_{A_1}, A_2') is surjective then so is for (A_1', J_{A_2}) and vice-versa.

By a slight abuse of notation, we write the pair (A_1, A_2) instead of (J_{A_1}, A_2') to study the images of maps.

6.3 Proof of Theorem 6.1.1

In this section, we study the image of the polynomial map

$$\omega(x_1, x_2) = A_1 x_1^{k_1} + A_2 x_2^{k_2}.$$

Proposition 6.3.1. *Let* $\mathbb{F}$ *be an algebraically closed field of characteristic 0, and let* $A_1, A_2 \in M_n(\mathbb{F})$. *For* $C \in M_n(\mathbb{F})$ *such that either* $\det(CA_1)$ *or* $\det(CA_2)$ *is non-zero, then* C *lies in the image of the evaluation map* $A_1x_1^{k_1} + A_2x_2^{k_2}$ *over* $M_n(\mathbb{F})$.

Proof. Assume $CA_1 \in \mathrm{GL}_n(\mathbb{F})$, so does $A_1^{-1}C$. Let $P_C \in \mathrm{GL}_n(\mathbb{F})$ such that $C' := P_C(A_1^{-1}C)P_C^{-1}$ is the Jordan canonical form of $A_1^{-1}C$. Clearly, $C' \in \mathrm{GL}_n(\mathbb{F})$. Since $\mathbb{F}$ is an algebraically closed field, every invertible matrix admits a k_1-th root in $\mathrm{GL}_n(\mathbb{F})$. Therefore, there exist $X_1' \in \mathrm{GL}_n(\mathbb{F})$ such that $C' = (X_1')^{k_1}$. For $X_1 = P_C^{-1}X_1'P_C$, we have $C = A_1X_1^{k_1}$ and hence,

$$C = A_1X_1^{k_1} + A_2X_2^{k_2}$$

where $X_2 = 0$. A similar argument works if $\det(CA_2)$ is non-zero. □

Following the same argument, we state the next proposition for an algebraically closed field having an arbitrary characteristic:

Proposition 6.3.2. *Let* $C \in M_2(\mathbb{F})$ *and* $\mathrm{char}(\mathbb{F}) \nmid k_1, k_2$. *If* $\det(CA_1)$ *or* $\det(CA_2)$ *is non-zero, then* C *lies in the image of the polynomial map* $A_1x_1^{k_1} + A_2x_2^{k_2}$.

Proof. Since $A_1^{-1}C \in \mathrm{GL}_n(\mathbb{F})$. The Jordan canonical form for $A_1^{-1}C$ possible are:

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \quad or \quad \begin{pmatrix} \lambda_1 & 1 \\ & \lambda_1 \end{pmatrix}$$

with $\lambda_1\lambda_2 \in \mathbb{F}^\times$. The diagonal matrix is k_1-th power of the diagonal matrix having entries $\sqrt[k_1]{\lambda_1}$ and $\sqrt[k_1]{\lambda_2}$. Moreover, the matrix $\begin{pmatrix} \sqrt[k_1]{\lambda_1} & 1 \\ & \sqrt[k_1]{\lambda_1} \end{pmatrix}^{k_1}$ is conjuagte to

$\begin{pmatrix} \lambda_1 & 1 \\ & \lambda_1 \end{pmatrix}^{k_1}$ as $\mathsf{char}(\mathbb{F}) \nmid k_1$. Therefore, in either case, take $X_2 = 0$ and $C = A_1 X_1^{k_1}$ has solution over $M_2(\mathbb{F})$. The similar argument follows for $\det(CA_2) \neq 0$. □

Proposition 6.3.3. *Let $\mathbb{F}$ be an algebraically closed field such that either $\mathsf{char}(\mathbb{F}) \nmid k_1$ or $\mathsf{char}(\mathbb{F}) \nmid k_2$ and $A_1, A_2 \in \mathrm{GL}_n(\mathbb{F})$. The map $\omega = A_1 x_1^{k_1} + A_2 x_2^{k_2}$ is surjective on $M_n(\mathbb{F})$.*

Proof. Assume $\mathsf{char}(\mathbb{F}) \nmid k_2$. The map ω is surjective on $M_n(\mathbb{F})$ if and only if the map $\widetilde{\omega} = x_1^{k_1} + A_1^{-1} A_2 x_2^{k_2}$ is surjective on $M_n(\mathbb{F})$. Choose a scalar $\alpha \in \mathbb{F}$ such that $C - (\alpha I_n)^{k_1}$ is invertible. This is always possible since C has only finitely many eigenvalues. By Proposition 6.3.1, there exist X_2 such that

$$A_1 A_2^{-1}(C - (\alpha I_2)^{k_1}) = X_2^{k_2}.$$

Setting $X_1 := \alpha I_n$, gives

$$C = X_1^{k_1} + A_1^{-1} A_2 X_2^{k_2}$$

and hence ω is surjective on $M_n(\mathbb{F})$. □

Thus, we focus on the cases where either A_1 or A_2 is singular with C being singular, and the case where A_1 and A_2 both are singular, and C is arbitrary. In the case $C = 0$, a solution always exists, given by $x_1 = x_2 = 0$. Thus, we will always consider C to be a non-zero matrix. To handle these systematically, we divide the proof into three propositions. The first of these is stated as follows.

Proposition 6.3.4. *Let $\omega = A_1 x_1^{k_1} + A_2 x_2^{k_2} \in M_2(\mathbb{F})\langle x_1, x_2 \rangle$, with A_1, A_2 nonzero matrices. If A_1 is a scalar matrix, then the map $\widetilde{\omega}$ is surjective.*

Proof. Let $A_1 = \lambda I_2$ for $\lambda \in \mathbb{F}^\times$. For $C \in M_2(\mathbb{F})$, we want the existence of solution of equation

$$\lambda x_1^{k_1} + A_2 x_2^{k_2} - C = 0.$$

Consider C to be a singular matrix. Let P_C be the matrix that conjugates C to its Jordan form and is one of the following:

$$\begin{pmatrix} \mu & 0 \\ 0 & 0 \end{pmatrix}_{\mu \neq 0} = \mu E_{11} \ or \ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = E_{12}.$$

Conjugate the above equation by P_C and consider C upto its Jordan form.

For the first case, since $\mu - \lambda x^{k_1} = 0$ has solution, so does $\mu E_{11} - \lambda x_1^{k_1} = 0$. For the second case, if there exists an α such that $E_{12} - \alpha A_2$ is non singular, then consider $y = \sqrt[k_2]{\alpha} I_2$. By Proposition 6.3.2, there exist X_1 such that C lies in the image of ω.

Suppose there doesn't exist any $\alpha \in \mathbb{F}$ such that $E_{12} - \alpha A_2$ is non-singular. Since A_2 is singular, it must be of the following form:

$$\begin{pmatrix} 0 & b_1 \\ 0 & b_2 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b_1 & b_2 \\ 0 & 0 \end{pmatrix}.$$

For the first case, let $X_2 \in M_2(\mathbb{F})$ such that $X_2^{k_2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then $C - A_2 X_2^{k_2} = \begin{pmatrix} b_1 & 1 \\ b_2 & 0 \end{pmatrix}$. If $b_2 = 0$, take $X_1 = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & 0 \end{pmatrix}$. Then $\lambda X_1^{k_1}$ is $\begin{pmatrix} \alpha_1^{k_1} & \alpha_1^{k_1-1}\alpha_2 \\ 0 & 0 \end{pmatrix}$. Since $b_1 \neq 0$, α_1 and α_2 can be chosen in such a way that $b_1 = \alpha_1^{k_1}$ and $1 = \alpha_1^{k_1-1}\alpha_2$. If

$b_2 \neq 0$, in this case $\begin{pmatrix} b_1 & 1 \\ b_2 & 0 \end{pmatrix}$ is a non-singular matrix and hence there exist X_1 such that $\lambda X_1^{k_1} = C - A_2 X_2^{k_2}$. The other case when A_2 is given by $\begin{pmatrix} b_1 & b_2 \\ 0 & 0 \end{pmatrix}$ follows the similar proof. Therefore, the given map is surjective. □

Corollary 6.3.1. If A_1 or A_2 is non singular, then the map ω defined by $\omega = A_1 x_1^{k_1} + A_2 x_2^{k_2}$ is surjective.

Proof. Let us assume A_1 is a non-singular matrix. The map ω is surjective if and only if

$$\widetilde{\omega} : x_1^{k_1} + A_1^{-1} A_2 x_2^{k_2}$$

is surjective. The map $\widetilde{\omega}$ is surjective, as follows from Proposition 6.3.4. □

By the preceding lemma, it suffices to consider the case in which both A_1 and A_2 are singular. We now formulate the following proposition, which characterises surjectivity when A_1 is diagonal up to conjugation.

Proposition 6.3.5. *Let* $\omega = A_1 x_1^{k_1} + A_2 x_2^{k_2} \in M_2(\mathbb{F})\langle x_1, x_2 \rangle$*, with* A_1*,* A_2 *nonzero matrices. Suppose up to conjugation,* A_1 *is a diagonal matrix* $\operatorname{diag}(\lambda, \mu)$*, then the associated map* $\widetilde{\omega}$ *is surjective if and only if one of the following conditions holds:*

1. A_1 *is invertible;*
2. $\lambda = 0$ *and the first row of an orbit representative of* A_2 *is nonzero;*
3. $\mu = 0$ *and the second row of an orbit representative of* A_2 *is nonzero.*

6.3.1 Proof of Proposition 6.3.5

In this subsection, we analyze the case in which the matrix A_1 has distinct eigenvalues. More precisely, we assume that A_1 is conjugate to the diagonal matrix

$\operatorname{diag}(\lambda,\mu)$ with $\lambda \neq \mu$.

Under this assumption, the equation under consideration takes the form

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} x_1^{k_1} + A_2 x_2^{k_2}.$$

As established earlier, it suffices to restrict our attention to the case in which A_1 is singular, that is, when either $\lambda = 0$ or $\mu = 0$.

To complete the proof, we examine representatives of the orbits arising from the action of the centralizer of the diagonal Jordan form $\operatorname{diag}(\lambda,\mu)$. Since A_2 is assumed to be singular, it leads to a finite collection of singular cases, summarised in Section 6.2, each of which is considered below separately. We begin by stating the two auxiliary lemmas that will be repeatedly used in the discussion of the cases of each representative.

Lemma 6.3.1. *Let* $\omega_1 : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ *given by* $\omega_1(x) = \begin{pmatrix} \delta_1 & \\ & 0 \end{pmatrix} x^k$ *for some* $\delta_1 \in \mathbb{F}^\times$. *The image of* ω_1 *is* $\left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a \neq 0 \text{ or } b \neq 0 \right\}$.

Proof. For $a \neq 0$, the equation $\delta_1 x^k - a = 0$ always has a non-zero solution, say α. Then $X = \begin{pmatrix} \alpha & \left(\alpha^{k-1}\delta_1\right)^{-1} b \\ & 0 \end{pmatrix}$ satisfies $\omega_1(X) = \begin{pmatrix} a & b \\ & 0 \end{pmatrix}$. For $a = 0$, the matrix $X = \begin{pmatrix} 0 & \delta_1^{-1} b \\ & 1 \end{pmatrix}$ satisfies $\omega_1(X) = \begin{pmatrix} 0 & b \\ & 0 \end{pmatrix}$, completing the proof. □

Lemma 6.3.2. *Let* $\omega_2 : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ *given by* $\omega_2(x) = \begin{pmatrix} 0 & \\ & \delta_2 \end{pmatrix} x^k$ *for some* $\delta_2 \in \mathbb{F}^{\times}$*. The image of* ω_2 *is* $\left\{ \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} : a \neq 0 \text{ or } b \neq 0 \right\}$.

Proof. Since $\begin{pmatrix} \delta_2 & 0 \\ 0 & 0 \end{pmatrix}$ is conjugate to $\begin{pmatrix} 0 & 0 \\ 0 & \delta_2 \end{pmatrix}$ by the matrix $P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in $\mathrm{GL}_2(\mathbb{F})$. Thus, the $\mathrm{Image}(\omega_2) = P\,(\mathrm{Image}(\omega_1))\,P^{-1} = \left\{ \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} : a \neq 0 \text{ or } b \neq 0 \right\}$. ∎

We now proceed to consider the cases corresponding to each representative.

6.3.1.1 Representative has distinct eigenvalues

Consider the map

$$\omega(x_1, x_2) := \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix} x_2^{k_2}.$$

Since both A_1 and A_2 are singular, we have $\lambda\mu = 0$ and $\mu_1\mu_2 = 0$. If $\lambda = \mu_1 = 0$ (respectively, $\mu = \mu_2 = 0$), then the map ω fails to contain all the matrices which have non-zero first row (respectively, non-zero last row). Consequently, the original map ω, before the action of the automorphism group, also fails to be surjective as the image does not contain all the conjugates of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with a or b being non-zero (respectively, with c or d being non-zero).

We therefore restrict our attention to $\lambda\mu_2 \in \mathbb{F}^\times$ and $\mu\mu_1 \in \mathbb{F}^\times$. In either case, the map ω can be decomposed as

$$\omega(x_1, x_2) = \omega_1(x_1) + \omega_2(x_2).$$

where the image of ω_1 consists precisely of matrices whose second row is zero (see Lemma 6.3.1), and the image of ω_2 is precisely all matrices having first row zero (see Lemma 6.3.2). It follows ω is surjective on $M_2(\mathbb{F})$ if and only if $\lambda\mu_2 \in \mathbb{F}^\times$ or $\mu\mu_1 \in \mathbb{F}^\times$, when both A_1 and A_2 are singular.

6.3.1.2 Representative is an upper triangular matrix

Consider the map

$$\omega(x_1, x_2) = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & 1 \\ & \mu_2 \end{pmatrix} x_2^{k_2}.$$

The representative of A_2 is the upper triangular matrix $\begin{pmatrix} \mu_1 & 1 \\ & \mu_2 \end{pmatrix}$. As stated before, A_1 and A_2 are both singular. Therefore, $\lambda\mu = 0$ and $\mu_1\mu_2 = 0$. We state the auxiliary lemmas that will be used to discuss this case.

Lemma 6.3.3. *For $\mu_1 \in \mathbb{F}$, define $\omega_3 : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ such that* $\omega_3(x) = \begin{pmatrix} \mu_1 & 1 \\ & 0 \end{pmatrix} x^k$ *for a positive integer k. Then the image of ω_3 is exactly the space of matrices spanned by E_{11} and E_{12}.*

Proof. The matrix $\begin{pmatrix} \mu_1 & 1 \\ & 0 \end{pmatrix}$ has second row zero. Multiplying by any matrix on the right will result in the second row being zero. Thus, we have to show that the image of ω_3 is exactly all the matrices having the second row equal to zero. Let C be a matrix of form $\begin{pmatrix} a & b \\ & 0 \end{pmatrix}$. For $b \in \mathbb{F}^{\times}$, there exist nonzero α_1 such that $\alpha_1^k = b$. Let $X = \begin{pmatrix} & 0 \\ \alpha_1^{-(k-1)}a & \alpha_1 \end{pmatrix}$. Then $X^k = \begin{pmatrix} & 0 \\ \alpha_1^{k-1}\alpha_1^{-(k-1)}a & \alpha_1^k \end{pmatrix} = \begin{pmatrix} 0 & \\ a & b \end{pmatrix}$. Therefore, $\omega_3(X) = C$ for $b \neq 0$. For $b = 0$, $\omega_3(X) = C$ for $X = \begin{pmatrix} 1 & \\ a & 0 \end{pmatrix}$, which completes the proof. □

Lemma 6.3.4. *Let k be a positive integer and $\mu_2 \in \mathbb{F}$. Define $\omega_4 : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ by* $\omega_4(x) = \begin{pmatrix} 0 & 1 \\ & \mu_2 \end{pmatrix} x^k$. *Then the image of ω_4 is precisely the space* $\mathscr{W} = \left\{ \begin{pmatrix} a & b \\ \mu_2 a & \mu_2 b \end{pmatrix} : a, b \in \mathbb{F} \right\}$.

Proof. Suppose $b \in \mathbb{F}^{\times}$, let $X = \begin{pmatrix} & 0 & \\ & & \\ \alpha_1^{-(k-1)}a & & \alpha_1 \end{pmatrix}$ where $\alpha_1^k = b$. For $b = 0$, let $X = \begin{pmatrix} 1 & \\ & \\ a & 0 \end{pmatrix}$. Then, in either case, the image of ω_4 contains the subspace $\mathscr{W}$.

The right multiplication of any matrix in $M_2(\mathbb{F})$ by $\begin{pmatrix} 0 & 1 \\ & \\ & \mu_2 \end{pmatrix}$ gives exactly the elements of the form $\mathscr{W}$ as $\begin{pmatrix} & 1 \\ & \\ 0 & \mu_2 \end{pmatrix} \begin{pmatrix} a & b \\ & \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ & \\ \mu_2 c & \mu_2 d \end{pmatrix}$. We first consider $\mu = 0$ and, using that, we give the solution for the case when $\lambda = 0$. □

We now state the main proposition associated with the given representative.

Proposition 6.3.6. *For $\mu_1, \mu_2 \in \mathbb{F}$. Define the map $\omega : M_2(\mathbb{F}) \times M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ given by*

$$(X_1, X_2) \mapsto \begin{pmatrix} \lambda & \\ & \\ & \mu \end{pmatrix} X_1^{k_1} + \begin{pmatrix} \mu_1 & 1 \\ & \\ & \mu_2 \end{pmatrix} X_2^{k_2}$$

The map ω is surjective if and only if $\lambda\mu_2 \in \mathbb{F}^{\times}$ or $\mu \in \mathbb{F}^{\times}$.

Proof. The proof is divided into two cases, based on $\lambda = 0$ and $\mu = 0$.

(I) Let $\mu = 0 = \mu_2$ and $\mu_1 \in \mathbb{F}$. The map is given by

$$\omega(x_1, x_2) = \begin{pmatrix} \lambda & \\ & \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & 1 \\ & \\ & 0 \end{pmatrix} x_2^{k_2}.$$

Clearly, $\omega(x_1,x_2) = \omega_1(x_1) + \omega_3(x_2)$. The image of ω_1 and ω_3 on $M_2(\mathbb{F})$ are same, see Lemma 6.3.1 and Lemma 6.3.3. Thus, the image of ω on $M_2(\mathbb{F})$ is the image of ω_1, which is exactly the matrices having the second row equal to zero.

(II) Let $\mu = 0 = \mu_1$ and $\mu_2 \in \mathbb{F}$. The map is

$$\omega(x_1,x_2) = \begin{pmatrix} \lambda & \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} 0 & 1 \\ & \mu_2 \end{pmatrix} x_2^{k_2}.$$

The map ω can be decomposed as $\omega_1 + \omega_4$. If $\mu_2 = 0$, then on evaluation on $M_2(\mathbb{F})$, the image of ω_4 is same as image of ω_1. Thus, for $\mu_2 = 0$,

$$\text{Image}(\omega) = \text{Image}(\omega_1) = \left\{ \begin{pmatrix} a & b \\ & 0 \end{pmatrix} : a,b \in \mathbb{F} \right\}.$$

Assume $\mu_2 \in \mathbb{F}^\times$. For $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, by Lemma 6.3.4, there exist $X_2 \in M_2(\mathbb{F})$ such that $\omega_4(X_2) = \begin{pmatrix} \mu_2^{-1}c & \mu_2^{-1}d \\ c & d \end{pmatrix}$. Then,

$$C - \omega_4(X_2) = \begin{pmatrix} a - \mu_2^{-1}c & b - \mu_2^{-1}d \\ 0 & 0 \end{pmatrix},$$

which lies in the image of ω_1 by Lemma 6.3.1. Thus, $\text{Image}(\omega) = M_2(\mathbb{F})$ for $\mu_2 \in \mathbb{F}^\times$.

(III) Let $\lambda = 0$ and $\mu_1, \mu_2 \in \mathbb{F}$. The map is

$$\omega(x_1,x_2) = \begin{pmatrix} 0 & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & 1 \\ & \mu_2 \end{pmatrix} x_2^{k_2}.$$

Decomposing the map

$$\omega(x_1,x_2) = \begin{cases} \omega_2(x_1) + \omega_3(x_2), & \text{if} \quad \mu_2 = 0, \\ \omega_2(x_1) + \omega_4(x_2), & \text{if} \quad \mu_1 = 0. \end{cases}$$

For $\mu_2 = 0$, by Lemma 6.3.3, there exist $X_2 \in M_2(\mathbb{F})$ such that for any $C \in M_2(\mathbb{F})$, $C - \omega_3(X_2) = \begin{pmatrix} & \\ * & * \end{pmatrix}$. By Lemma 6.3.2, there exist $X_1 \in M_2(\mathbb{F})$ such that $C - \omega_3(X_2) = \omega_2(X_1)$. For $\mu_1 = 0$, using Lemma 6.3.4, there exist $X_2 \in M_2(\mathbb{F})$ such that $\omega_4(X_2) = \begin{pmatrix} a & b \\ \mu_2 a & \mu_2 b \end{pmatrix}$. Thus for an arbitary $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{F})$, using Lemma 6.3.2, we have

$$C - \omega_4(X_2) = \begin{pmatrix} 0 & 0 \\ c - \mu_2 a & d - \mu_2 b \end{pmatrix} \in \text{Image}(\omega_2).$$

That is, there exist $X_1 \in M_2(\mathbb{F})$ such that $C = \omega_2(X_1) + \omega_4(X_2)$.

Thus, for $\lambda\mu_2 \in \mathbb{F}^\times$ by **(I)** and **(II)** and for $\mu \in \mathbb{F}^\times$ by **(III)** the map ω is surjective on $M_2(\mathbb{F})$. Suppose $\lambda\mu_2 = 0$ and $\mu = 0$. Since A_1 is a non-zero matrix, we must have $\lambda \neq 0$. Thus $\mu_2 = \mu = 0$ and the map is given by

$$\begin{aligned}\omega(x_1,x_2) &= \begin{pmatrix} \lambda & \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & 1 \\ & 0 \end{pmatrix} x_2^{k_2} \\ &= \omega_1(x_1) + \omega_3(x_2).\end{aligned}$$

By Lemma 6.3.1 and Lemma 6.3.3, the

$$\text{Image}(\omega) = \left\{ \begin{pmatrix} a & b \\ & \end{pmatrix} : a,b \in \mathbb{F} \right\}.$$

□

6.3.1.3 Representative is a lower triangular matrix

In this part, the representative for A_2 is $\begin{pmatrix} \mu_1 & 0 \\ 1 & \mu_2 \end{pmatrix}$.

Proposition 6.3.7. *Let $\lambda, \mu, \mu_1, \mu_2 \in \mathbb{F}$ such that $\lambda\mu = 0$ and $\mu_1\mu_2 = 0$. Define the map*

$$\omega, \widetilde{\omega} : M_2(\mathbb{F}) \times M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$$

given by

$$\omega(x_1,x_2) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & 0 \\ 1 & \mu_2 \end{pmatrix} x_2^{k_2}$$

$$\widetilde{\omega}(x_1,x_2) = \begin{pmatrix} \mu & 0 \\ 0 & \lambda \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_2 & 1 \\ & \mu_1 \end{pmatrix} x_2^{k_2}.$$

The map ω is surjective on $M_2(\mathbb{F})$ if and only if the map $\widetilde{\omega}$ is surjective on $M_2(\mathbb{F})$. Moreover, the map ω is surjective if and only if $\lambda \in \mathbb{F}^\times$ or $\mu\mu_1 \in \mathbb{F}^\times$.

Proof. Suppose the map ω is surjective on $M_2(\mathbb{F})$. That is, for every $C \in M_2(\mathbb{F})$, there exist $X_1, X_2 \in M_2(\mathbb{F})$ such that $C = \omega(X_1, X_2)$. Let $P = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \in \mathrm{PGL}_2(\mathbb{F})$. Conjugating $C = \omega(X_1, X_2)$ by P, we get,

$$PCP^{-1} = \begin{pmatrix} \mu & 0 \\ 0 & \lambda \end{pmatrix} (PX_1P^{-1})^{k_1} + \begin{pmatrix} \mu_2 & 1 \\ & \mu_1 \end{pmatrix} (PX_2P^{-1})^{k_2}.$$

Thus, $PCP^{-1} \in \mathrm{Image}(\widetilde{\omega})$ for every $C \in M_2(\mathbb{F})$. The conjugation action by P on $M_2(F)$ defines an automorphism of $M_2(\mathbb{F})$. Therefore, $\mathrm{Image}(\widetilde{\omega}) = M_2(\mathbb{F})$. Thus, $\widetilde{\omega}$ is surjective on $M_2(\mathbb{F})$. Conversely, conjugating the image of $\widetilde{\omega}$ by the same P gives the image of ω on $M_2(\mathbb{F})$. By Proposition 6.3.6, it follows that the map ω is surjective if and only if $\lambda \in \mathbb{F}^\times$ or $\mu\mu_1 \in \mathbb{F}^\times$. □

6.3.1.4 Representative having second row non-zero

We consider the map whose second representative has all non-zero entries and is given by

$$\omega(x_1,x_2) = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} x_2^{k_2}.$$

Lemma 6.3.5. *For $z_i \in \mathbb{F}$ and $\lambda \in \mathbb{F}^\times$, the map*

$$\omega(x_1,x_2) = \begin{pmatrix} \lambda & \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} x_2^{k_2}$$

is always surjective on $M_2(\mathbb{F})$.

Proof. Let $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{F})$. First, assume that $c \neq 0$. There exists $b_0 \in \mathbb{F}^\times$ such that $b_0^{k_2} = c$. Define $X_2 = \begin{pmatrix} b_0 & b_1 \\ 0 & \end{pmatrix}$, where $b_1 \in \mathbb{F}$ is chosen so that $b_0^{k_2-1} b_1 = d$. If $c = 0$, define $X_2 = \begin{pmatrix} 0 & b_0 \\ & 1 \end{pmatrix}$, where $b_0 = d - z_3 \in \mathbb{F}$. In either case, the matrix $C - \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} X_2^{k_2}$ is of the form $\begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix}$. By Lemma 6.3.1, there exists $X_1 \in M_2(\mathbb{F})$ such that $C = \omega(X_1,X_2)$. This proves that ω is surjective. □

Lemma 6.3.6. *For $\mu \in \mathbb{F}^\times$ and z_1 and z_2 both being non-zero, every element of $M_2(\mathbb{F})$ lies in the image of the map*

$$\omega(x_1,x_2) = \begin{pmatrix} 0 & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} x_2^{k_2}.$$

Proof. Let C be an arbitrary element of $M_2(\mathbb{F})$, given by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If $a \in \mathbb{F}^{\times}$, then let $X_2 = \begin{pmatrix} a_0 & a_1 \\ & 0 \end{pmatrix}$ where $a = z_1 a_0^{k_2}$ and $b = z_1 a_0^{k_2-1} a_1$. Otherwise let $X_2 = \begin{pmatrix} 0 & z_1^{-1}(b - z_2) \\ & 1 \end{pmatrix}$. In either case, we get,

$$C - \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} X_2^{k_2} = \begin{pmatrix} & \\ * & * \end{pmatrix} \in M_2(\mathbb{F}).$$

Using Lemma 6.3.2, any element of this form lies in the image of ω_2 on $M_2(\mathbb{F})$. Thus the map ω is surjective on $M_2(\mathbb{F})$. □

Remark 6.3.2. We note that in the above lemma, if $z_1 = 0$ then it implies z_2 is also zero as the matrix $\begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix}$ is singular. In this case, the map

$$\omega(x_1, x_2) = \begin{pmatrix} 0 & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} 0 & \\ 1 & z_3 \end{pmatrix} x_2^{k_2}$$

has the same image as ω_2 on $M_2(\mathbb{F})$, see Lemma 6.3.2.

We now state the main result for the given representative.

Proposition 6.3.8. *Let z_1, z_2, z_3 be elements in $\mathbb{F}$. The map*

$$\omega(x_1,x_2) = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} x_1^{k_1} + \begin{pmatrix} z_1 & z_2 \\ 1 & z_3 \end{pmatrix} x_2^{k_2}$$

is surjective on $M_2(\mathbb{F})$ if and only if $\lambda \in \mathbb{F}^\times$ or $\mu z_1 \in \mathbb{F}^\times$.

Proof. Suppose ω is surjective. Assume, for contradiction, $\lambda = \mu z_1 = 0$. Then either $\mu = 0$ or $z_1 = 0$. Since the representatives are non-zero and $\lambda = 0$, we must have $\mu \neq 0$ and $z_1 = 0$. By the above remark, $\text{Image}(\omega) = \text{Image}(\omega_2)$ which is a proper subspace of $M_2(\mathbb{F})$, which contradicts the surjectivity. Therefore, we must have $\lambda \in \mathbb{F}^\times$ or $\mu z_1 \in \mathbb{F}^\times$. Conversely, let $\lambda \in \mathbb{F}^\times$ or $\mu z_1 \in \mathbb{F}^\times$. Since both representatives are singular, we distinguish two cases. If $\lambda \in \mathbb{F}^\times$ and $\mu = 0$, the result follows from Lemma 6.3.5. If $\lambda = 0$ and $\mu \in \mathbb{F}^\times$, the conclusion follows from Lemma 6.3.6. This completes the proof. □

Proposition 6.3.9. *Let $\omega = Ax^{k_1} + By^{k_2} \in M_2(\mathbb{F})\langle x,y\rangle$, with A, B nonzero matrices. If up to conjugation A is a unipotent matrix $\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}$, then the map $\widetilde{\omega}$ is surjective if and only if one of the following cases occurs;*

1. *A is invertible*
2. *if $\lambda = 0$, the second row of an orbit representative of B has a nonzero second row.*

6.3.2 Proof of Proposition 6.3.9

In this subsection, we consider the map corresponding to the pair whose first representative is a λ-unipotent matrix. For $\lambda \in \mathbb{F}$, define the map $\omega : M_2(\mathbb{F}) \times M_2(\mathbb{F}) \longrightarrow$

$M_2(\mathbb{F})$ given by $\omega(X_1, X_2) = \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix} X_1^{k_1} + A_2 X_2^{k_2}$, where A_2 is the representative obtained under the action of centralizer of $\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}$. If $\lambda \in \mathbb{F}^\times$, the map ω is surjective, follows from Theorem 6.3.1. Thus, A_1 and A_2 are both singular matrices. Therefore, we now consider the cases when $\lambda = 0$.

6.3.2.1 Representative is a lower triangular matrix

The representative of A_2 is a lower triangular matrix given by $\begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix}$ where $z \in \mathbb{F}^\times$. With the choice of representative, the map takes the form

$$\omega(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix} x_2^{k_2}.$$

Proposition 6.3.10. *For $z \in \mathbb{F}^\times$ and $\mu_1, \mu_2 \in \mathbb{F}$, define the map $\omega, \widetilde{\omega} : M_2(\mathbb{F}) \times M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ as*

$$\omega(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix} x_2^{k_2}$$

$$\widetilde{\omega}(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ 1 & \mu_2 \end{pmatrix} x_2^{k_2}.$$

Then ω is surjective if and only if $\widetilde{\omega}$ is surjective.

Proof. Suppose ω is surjective. Then for any $C \in M_2(\mathbb{F})$, there exists $X_1, X_2 \in M_2(\mathbb{F})$ such that $C = \omega(X_1, X_2)$.

Let $P = \begin{pmatrix} 1 & \\ & z^{-1} \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F})$. Conjugating $C = \omega(X_1, X_2)$ by P we get,

$$\begin{aligned} PCP^{-1} = P\omega(X_1,X_2)P^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix} (PX_1P^{-1})^{k_1} \\ &+ \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} \mu_1 & 0 \\ z & \mu_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix} (PX_2P^{-1})^{k_2} \\ &= \begin{pmatrix} 0 & z \\ & 0 \end{pmatrix} (PX_1P^{-1})^{k_1} + \begin{pmatrix} \mu_1 & \\ 1 & \mu_2 \end{pmatrix} (PX_2P^{-1})^{k_2} \\ &= \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} (zI_2)(PX_1P^{-1})^{k_1} + \begin{pmatrix} \mu_1 & \\ 1 & \mu_2 \end{pmatrix} (PX_2P^{-1})^{k_2}. \end{aligned}$$

Since $\mathbb{F}$ is an algebraically closed field, for the given $z \in \mathbb{F}$, there exists z_0 in $\mathbb{F}$ such that $z_0^{k_1} = z$ and $z_0 I_2 \in \mathscr{Z}(M_2(\mathbb{F}))$. Thus, $z_0 I_2 P X_1 P^{-1} = P(z_0 I_2 X_1)P^{-1}$. Let $\widetilde{X}_1 = P(z_0 I_2 X_1)P^{-1}$ and $\widetilde{X}_2 = PX_2P^{-1}$. We get,

$$PCP^{-1} = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} \widetilde{X}_1^{k_1} + \begin{pmatrix} \mu_1 & \\ 1 & \mu_2 \end{pmatrix} \widetilde{X}_2^{k_2}$$

$$= \widetilde{\omega}(\widetilde{X}_1, \widetilde{X}_2).$$

Conjugation action defines an automorphism on $M_2(\mathbb{F})$. Therefore, for every $PCP^{-1} \in M_2(\mathbb{F})$, there exist $\widetilde{X}_1, \widetilde{X}_2 \in \mathbb{F}$ such that $C = \widetilde{\omega}(\widetilde{X}_1, \widetilde{X}_2)$.

Conversely, suppose for $C \in M_2(\mathbb{F})$, there exist $X_1, X_2 \in M_2(\mathbb{F})$ such that $C = \widetilde{\omega}(X_1, X_2)$. Then conjugating it by $P = \begin{pmatrix} 1 & \\ & z \end{pmatrix}$ gives the surjectivity of the map ω. $\square$

Proposition 6.3.11. *The map $\widetilde{\omega}$ is surjective on $M_2(\mathbb{F})$.*

Proof. Since A_2 is singular, we have $\mu_1\mu_2 = 0$ and $P = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \in \mathrm{PGL}_2(\mathbb{F})$. Suppose $\mu_1 = 0$ and $\mu_2 \in \mathbb{F}$. Then

$$\widetilde{\omega}(x_1, x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} 0 & \\ 1 & \mu_2 \end{pmatrix} x_2^{k_2} = \omega_3(x_1) + \omega_3^P(x_2)$$

where $\omega_3^P : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ is defined as $\omega_3^P(x_2) = P \begin{pmatrix} \mu_2 & 1 \\ & 0 \end{pmatrix} P^{-1} x_2^{k_2}$. By Lemma 6.3.3, the $\mathrm{Image}(\omega_3) = \left\{ \begin{pmatrix} a & b \\ & \end{pmatrix} : a, b \in \mathbb{F} \right\}$ and $\mathrm{Image}(\omega_3^P) = P\mathrm{Image}(\omega_3)P^{-1} = \left\{ \begin{pmatrix} & \\ c & d \end{pmatrix} : c, d \in \mathbb{F} \right\}$. Thus $\mathrm{Image}(\widetilde{\omega}) = M_2(\mathbb{F})$ for $\mu_1 = 0$.

Suppose $\mu_2 = 0$ and $\mu_1 \in \mathbb{F}$. Then

$$\widetilde{\omega}(x_1,x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} 0 & \\ 1 & \mu_2 \end{pmatrix} x_2^{k_2} = \omega_4(x_1) + \omega_4^P(x_2)$$

where $\omega_4^P(x_2) := P \begin{pmatrix} 0 & 1 \\ & \mu_1 \end{pmatrix} P^{-1} x_2^{k_2}$ on $M_2(\mathbb{F})$. Then by Lemma 6.3.2, The $\text{Image}(\widetilde{\omega})$ is $M_2(\mathbb{F})$ for $\mu_2 = 0$. Both cases combined prove the surjectivity of $\widetilde{\omega}$ on $M_2(\mathbb{F})$. □

Corollary 6.3.3. For $z \in \mathbb{F}^\times$, the map $\omega(x_1,x_2) := \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix} x_2^{k_2}$ is surjective on $M_2(\mathbb{F})$.

Proof. Since ω is surjective on $M_2(\mathbb{F})$ if and only if $\widetilde{\omega}$ is surjective on $M_2(\mathbb{F})$. By the above proposition, it follows that ω is surjective on $M_2(\mathbb{F})$. □

6.3.2.2 Representative is diagonal with distinct entries

We consider the map where the second representative is a diagonal matrix with distinct diagonal values. For $\mu_1, \mu_2 \in \mathbb{F}$, consider the map

$$\omega(x_1,x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix} x_2^{k_2}.$$

Proposition 6.3.12. *The map ω is surjective if and only if $\mu_2 \in \mathbb{F}^\times$. Moreover, if $\mu_2 = 0$, the image of ω is the space spanned by E_{11} and E_{12} in $M_2(\mathbb{F})$.*

Proof. The map ω is surjective if and only if $\mu_2 \in \mathbb{F}^\times$ follows from Proposition 6.3.6. If $\mu_2 = 0$ and $\mu_1 \in \mathbb{F}^\times$, then by Lemma 6.3.3, the $\text{Image}(\omega) = \text{Image}(\omega_3)$ which is exactly the space spanned by E_{11} and E_{12}. $\square$

6.3.2.3 Representative is an upper triangular matrix

For $\mu_1 \in \mathbb{F}$, the map ω is defined by

$$\omega(x_1,x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} \mu_1 & z \\ & \mu_1 \end{pmatrix} x_2^{k_2}$$

where the representative of A_2 is $\begin{pmatrix} \mu_1 & z \\ & \mu_1 \end{pmatrix}$ where $\mu_1 \in \mathbb{F}$. Since both representatives are singular, so $\mu_1 = 0$. Thus, the map gets reduced to

$$\omega(x_1,x_2) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x_1^{k_1} + \begin{pmatrix} 0 & z \\ & 0 \end{pmatrix} x_2^{k_2}.$$

Proposition 6.3.13. *The map ω is not surjective on $M_2(\mathbb{F})$.*

Proof. The map $\omega(x_1,x_2) = \omega_3(x_1) + zI_2\omega_3(x_2)$. The image of ω_3 on $M_2(\mathbb{F})$ is $\left\{ \begin{pmatrix} a & b \\ & \end{pmatrix} : a,b \in \mathbb{F} \right\}$ by Lemma 6.3.3. Thus, the image of ω is the same as the image of ω_3, which is a proper subspace of $M_2(\mathbb{F})$. Hence, the map is not surjective. $\square$

The exact images of the map $\omega(x_1,x_2) = A_1x_1^{k_1} + A_2x_2^{k_2}$ obtained under the representative of (A_1,A_2) under the simultaneous conjugation is described in 9.1.5.

6.4 Proof of Theorem 6.1.2

In this section, we study the matrix-valued noncommutative polynomials in the context of Lv́ov–Kaplansky conjecture. Let $\omega \in M_2(\mathbb{F})\langle x_1, x_2\rangle$ defined by

$$A_1(x_1x_2) - A_2(x_2x_1),$$

where A_1, A_2 are nonzero matrices. Shoda [57] proved that the image of the commutator polynomial $x_1x_2 - x_2x_1$ on $M_n(\mathbb{F})$, where $\mathbb{F}$ is a field of characteristic 0, is the Lie algebra $\mathfrak{sl}_n(\mathbb{F})$. This result was later extended by Albert and Muckenhoupt [2] for arbitrary characteristic. These results provided the evidence for Lv́ov–Kaplansky Conjecture. We begin by stating a lemma that simplifies the problem into a few cases.

Lemma 6.4.1. *Let $\omega(x_1,x_2) = A_1(x_1x_2) - A_2(x_2x_1) \in M_2(\mathbb{F})\langle x_1, x_2\rangle$ where A_1, A_2 are non-zero matrices. Then*

1. *If $A_1 = A_2$, then the image of ω on $M_2(\mathbb{F})$ is the vector subspace.*
2. *If $A_1 - A_2 \in \mathrm{GL}_2(\mathbb{F})$, then the image of ω on $M_2(\mathbb{F})$ is exactly $M_2(\mathbb{F})$.*

Proof. When $A_1 = A_2$, the image of the polynomial ω is given by $A_1\mathfrak{sl}_2(\mathbb{F})$, and hence is a vector subspace of $M_2(\mathbb{F})$.

Assume $A_1 - A_2 \in \mathrm{GL}_2(\mathbb{F})$. For any $C \in M_2(\mathbb{F})$, take $x_1 = (A_1 - A_2)^{-1}C$ and $x_2 = I_2$, the identity matrix. Then

$$\omega(x_1,x_2) = A_1x_1 - A_2x_1 = (A_1 - A_2)x_1 = C.$$

□

The lemma shows that the map ω is surjective whenever $A_1 = A_2$ or $A_1 - A_2$ is invertible. Therefore, we consider A_1, A_2 such that $A_1 - A_2$ is singular. As discussed in Section 6.2, we use the simultaneous conjugacy classes and solve the problem on the reduced equation. By abuse of notation, we continue to denote these representatives by A_1, A_2. We divide the cases into subsections depending on the representative of A_1 as classified in Section 6.2.

6.4.1 Representative is a scalar matrix

We begin with the case where A_1 is a scalar matrix. Owing to the large centraliser of A_1, this case admits a straightforward reduction and serves as a useful base case. The following proposition shows that the image of ω is always a vector subspace in this setting.

Proposition 6.4.1. *For $A_1 = \mu_1 I_2 \in \mathrm{GL}_2(\mathbb{F})$ and $A_2 \in M_2(\mathbb{F})$, the image of the map*

$$\omega(x_1,x_2) = A_1(x_1x_2) - A_2(x_2x_1)$$

on $M_2(\mathbb{F})$ is always a vector subspace of $M_2(\mathbb{F})$.

Proof. The representatives of A_2 under the action of the centraliser of the scalar matrix are

$$\begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mu_1 & 1 \\ & \mu_1 \end{pmatrix}.$$

1. For $A_2 = \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}$. Let $x_1 = \begin{pmatrix} a_1 & \\ a_3 & 1 \end{pmatrix}$ and $x_2 \begin{pmatrix} & 1 \\ b_3 & b_4 \end{pmatrix}$. Then

$$\begin{pmatrix} \mu_1 & \\ & \mu_1 \end{pmatrix}\begin{pmatrix} a_1 & \\ a_3 & 1 \end{pmatrix}\begin{pmatrix} & 1 \\ b_3 & b_4 \end{pmatrix} - \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}\begin{pmatrix} & 1 \\ b_3 & b_4 \end{pmatrix}\begin{pmatrix} a_1 & \\ a_3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -\mu_1 a_3 & \mu_1(a_1-1) \\ \mu_1 b_3 - \mu_2 a_1 b_3 - \mu_2 a_3 b_4 & \mu_1 a_3 + (\mu_1-\mu_2) b_4 \end{pmatrix}.$$

Since $\mu_1 - \mu_2 \neq 0$ and $\mu_1 \neq 0$, if we choose a_3, a_1, b_4 and b_3 in the given order, it will cover all matrices of $M_2(\mathbb{F})$. In this case, the map is surjective and hence a vector space.

2. For $A_2 = \begin{pmatrix} \mu_1 & 1 \\ & \mu_1 \end{pmatrix}$. Let $x_1 = \begin{pmatrix} a_1 & a_2 \\ & a_4 \end{pmatrix}$ and $x_2 = \begin{pmatrix} b_1 & b_2 \\ 1 & \end{pmatrix}$. Then

$$\begin{pmatrix} \mu_1 & \\ & \mu_1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ & a_4 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ 1 & \end{pmatrix} - \begin{pmatrix} \mu_1 & 1 \\ & \mu_1 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ 1 & \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ & a_4 \end{pmatrix}$$
$$= \begin{pmatrix} \mu_1 a_2 - a_1 & \mu_1 a_1 b_2 - \mu_1 a_2 b_1 - a_2 - \mu_1 a_4 b_2 \\ \mu_1 a_4 - \mu_1 a_1 & -\mu_1 a_2 \end{pmatrix}.$$

Since $\mu_1 \neq 0$, choosing the elements in the order a_2, a_1, a_4 and b_2 or b_1 (when at least one of a_1, a_2, a_4, $a_1 - a_4$ is nonzero), we get all matrices but $\begin{pmatrix} 0 & \alpha \\ & 0 \end{pmatrix}$. But this matrix can be obtained by setting x_1 to be the identity matrix and $x_2 = \text{diag}(\alpha, \alpha)$. In this case, the map gives the image to be $M_2(\mathbb{F})$ and hence is a vector space.

□

Corollary 6.4.1. Let $A_1, A_2 \in M_2(\mathbb{F})$ be two non-zero elements. If either of A_1 or A_2 is non-singular, then the image of the map $\omega(x_1, x_2) = A_1(x_1 x_2) - A_2(x_2 x_1)$ evaluated over $M_2(\mathbb{F})$, is a vector subspace of $M_2(\mathbb{F})$.

Proof. Let us assume that A_1 is a non-singular matrix. Define the map

$$\begin{aligned}\widetilde{\omega} : M_2(\mathbb{F}) \times M_2(\mathbb{F}) &\longrightarrow M_2(\mathbb{F})\\ (x_1, x_2) &\mapsto (x_1 x_2) - A_1^{-1} A_2 (x_2 x_1).\end{aligned}$$

The map $\widetilde{\omega}$ corresponds to the tuple $(I_2, A_1^{-1}A_2)$. By Proposition 6.4.1, the image of such a map is always a vector subspace of $M_2(\mathbb{F})$. Define $\tau : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ by $x \mapsto A_1(x)$. The map τ is a linear transformation and hence maps a vector subspace to a vector subspace. Consequently, $\text{Image}(\omega) = A_1 \text{Image}(\widetilde{\omega})$ is a vector subspace of $M_2(\mathbb{F})$. Moreover, if $A_1 \neq A_2$, the image of ω is the full space $M_2(\mathbb{F})$. When A_2 is non-singular, the proof follows analogously. □

6.4.2 Representative is a diagonal matrix

We consider the case where the representative of A_1 is a diagonal matrix $\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix}$ with $\lambda \neq \mu$. The representatives of A_2 under the action of the centralizer of A_1 are described in Section 6.2. In this section, we first establish a lemma, which will be used to prove the main result in Proposition 6.4.2.

Lemma 6.4.2. *Let $A_1 \in M_2(\mathbb{F})$ be a singular diagonal matrix with distinct diagonal entries, and let $A_2 \in M_2(\mathbb{F})$ be singular. Then the image of the map $\omega(x_1, x_2) = A_1(x_1x_2) - A_2(x_2x_1)$ is a vector subspace of $M_2(\mathbb{F})$.*

Proof. Write $A_2 = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$. If $A_1 - A_2$ is non-singular, the result follows from Lemma 6.4.1. Otherwise, assume $A_1 - A_2$ is singular. Since A_1 is singular, we first assume $\mu = 0$ and hence $A_1 - A_2 = \begin{pmatrix} \lambda - b_1 & -b_2 \\ -b_3 & -b_4 \end{pmatrix}$. Since the determinant of $A_1 - A_2$ is zero, we get $b_4 = 0$. Moreover, $\det(A_2) = 0$ implies either $b_2 = 0$ or $b_3 = 0$. Therefore, A_2 takes one of the two forms $\begin{pmatrix} b_1 & b_2 \\ & 0 \end{pmatrix}$ or $\begin{pmatrix} b_1 & \\ b_3 & 0 \end{pmatrix}$. This leads to two cases for the map ω:

1. $\omega_1(x_1, x_2) = \begin{pmatrix} \lambda & \\ & 0 \end{pmatrix}(x_1x_2) - \begin{pmatrix} b_1 & b_2 \\ & 0 \end{pmatrix}(x_2x_1)$ and
2. $\omega_2(x_1, x_2) = \begin{pmatrix} \lambda & \\ & 0 \end{pmatrix}(x_1x_2) - \begin{pmatrix} b_1 & \\ b_3 & 0 \end{pmatrix}(x_2x_1)$.

The image of map ω_1 will always be a subset of the space $\left\{ \begin{pmatrix} a & b \\ & \end{pmatrix} : a, b \in \mathbb{F} \right\}$ as the second row of matrices lie in the kernel space of the transformation obtained by left multiplication of matrices A_1 and A_2 on $M_2(\mathbb{F})$. Let $x_1 = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$ and $x_2 = \begin{pmatrix} 0 & \\ g & h \end{pmatrix}$. Then $\omega(x_1, x_2) = \begin{pmatrix} \lambda g & \lambda h - b_2 g \\ & 0 \end{pmatrix}$. Since $\lambda \neq 0$, choosing g and h gives the matrices of the form $\begin{pmatrix} a & b \\ & 0 \end{pmatrix}$. Thus, $\text{Image}(\omega_1)$ is a proper vector subspace of $M_2(\mathbb{F})$.

If b_3 is zero in ω_2, then the image is a subspace which follows from the map ω_1

above, with $b_2 = 0$. If b_3 is non-zero, let $x_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ and $x_2 = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$. Then

$$\omega_2(x_1,x_2) = \begin{pmatrix} \lambda g - b_1 f & \lambda h - b_1 e \\ -b_3 f & -b_3 e \end{pmatrix}.$$

Since λ and b_3 are non-zero, choose f,e and then g,h. This would give the entire space $M_2(\mathbb{F})$.

For $\lambda = 0$, conjugating the map

$$\widetilde{\omega}(x_1,x_2) = \begin{pmatrix} \mu & \\ & 0 \end{pmatrix}(x_1x_2) - A_2(x_2x_1)$$

by the element $P = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \in \mathrm{GL}_2(\mathbb{F})$ shows that the image of ω is a vector subspace of $M_2(\mathbb{F})$. □

We state the main result of this section as follows:

Proposition 6.4.2. *Let $A_1, A_2 \in M_2(\mathbb{F})$ be non-zero matrices, and A_1 is a diagonal matrix. Then the image of the map $\omega(x_1,x_2) = A_1(x_1x_2) - A_2(x_2x_1)$ in $M_2(\mathbb{F})$ is a vector subspace.*

Proof. If A_1 or A_2 is non-singular, then the image of ω is a subspace of $M_2(\mathbb{F})$ by Theorem 6.4.1. If $A_1 - A_2$ is a non-singular matrix, then the image of ω is a subspace of $M_2(\mathbb{F})$ by Lemma 6.4.1. Thus, we assume A_1, A_2 and $A_1 - A_2$ are singular matrices. The image of the map ω is a vector subspace of $M_2(\mathbb{F})$ by Lemma 6.4.2. □

6.4.3 Representative is a unipotent matrix

The representative of A_1 is $\begin{pmatrix} \mu_1 & 1 \\ & \mu_1 \end{pmatrix}$. The representatives of A_2 obtained under the conjugation action of the centralizer of A_1 are described in Section 6.2. We first discuss the image of the map $\omega(x_1,x_2) = A_1(x_1x_2) - A_2(x_2x_1)$ on the representatives of (A_1,A_2) and summarize the result in Proposition 6.4.3.

6.4.3.1 Lower Triangular with Distinct Diagonal Entries

The representative of A_2 under the conjugation action of the centraliser of A_1 is a lower triangular matrix given by $\begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix}$ where $z \in \mathbb{F}^\times$. Since $\det(A_1 - A_2) = 0$ and $z \neq 0$, it follows that λ is distinct from both μ_1 and μ_2. Let $x_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $x_2 = \begin{pmatrix} 1 & \\ 1 & h \end{pmatrix}$. Evaluating $A_1(x_1x_2) - A_2(x_2x_1)$ gives

$$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & \\ 1 & h \end{pmatrix} - \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix} \begin{pmatrix} 1 & \\ 1 & h \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$= \begin{pmatrix} (\lambda-\mu_1)a+\lambda b+c+d & (\lambda h-\mu_1)b+d \\ -(z+\mu_2)a+(\lambda-\mu_2 h)c+\lambda d & -(z+\mu_2)b+h(\lambda_1-\mu_2)d \end{pmatrix}.$$

Since λ and μ_2 can not be simultaneously zero, and the determinant of the coefficient matrix of a,b,c,d is $h[(\lambda-\mu_2)\lambda(h-1)+\mu_2][(\lambda-\mu_1)\mu_2(1-h)+\mu_2 g]$, we can choose h such a way that the determinant does not vanish. Hence, the image is the full matrix algebra $M_2(\mathbb{F})$, which is a vector space.

6.4.3.2 Lower Triangular Representative with Constant Diagonal

The representative of A_2 under the conjugation action of the centralizer of A_1 is a lower triangular matrix which has the same eigenvalue μ given by $\begin{pmatrix} \mu & \\ z & \mu \end{pmatrix}$ with $z \in \mathbb{F}^{\times}$. Since $A_1 - A_2$ has zero determinant, it follows that $\lambda \neq \mu$. The surjectivity of ω in this case follows from 6.4.3.1 by using the condition that $\lambda \neq \mu$.

6.4.3.3 Diagonal Matrix with Distinct Diagonal Entries

The representative of A_2 is given by a diagonal matrix having distinct diagonal entries. The map is defined by

$$\begin{aligned} \omega(x_1,x_2) &= \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}(x_1x_2) - \begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}(x_2x_1) \\ &= -\left(\begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}(x_2x_1) - \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}(x_1x_2)\right). \end{aligned}$$

As discussed in 6.4.2, the map

$$\widetilde{\omega}(x_1,x_2) = \left(\begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix} (x_2x_1) - \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix} (x_1x_2) \right)$$

is surjective on $M_2(\mathbb{F})$ and $\mathbf{Image}(\omega) = \mathbf{Image}(\widetilde{\omega})$. Therefore, the map ω on evaluation on $M_2(\mathbb{F})$ is surjective.

6.4.3.4 Upper Triangular with Constant Diagonal Entries

The representative of A_2 under the conjugation action is an upper triangular matrix with the same diagonal entries given by $\begin{pmatrix} \mu & z \\ & \mu \end{pmatrix}$ where $z \in \mathbb{F}^{\times}$. Since $A_1 - A_2$ has zero determinant, we have $\lambda = \mu$.

Let $x_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $x_2 = \begin{pmatrix} & f \\ 1 & \end{pmatrix}$. On computing $A_1(x_1x_2) - A_2(x_2x_1)$, we get,

$$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} & f \\ 1 & \end{pmatrix} - \begin{pmatrix} \lambda & z \\ & \lambda \end{pmatrix} \begin{pmatrix} & f \\ 1 & \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$= \begin{pmatrix} \lambda b + d - \lambda fc - za & \lambda af + cf - \lambda fd - zb \\ \lambda d - \lambda a & \lambda cf - \lambda b \end{pmatrix}.$$

Then f can be chosen such that the coefficient matrix of the system of equations in q,b,c,d has a non-zero determinant. This shows that the image in this case is

$M_2(\mathbb{F})$. If λ is zero, then the map is given by

$$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}(x_1x_2) - \begin{pmatrix} 0 & z \\ & 0 \end{pmatrix}(x_2x_1) = \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}((x_1x_2) - (zI_2)(x_2x_1)).$$

If $z = 1$, then the image of $x_1x_2 - (zI_2)(x_2x_1)$ is $\mathfrak{sl}_2(\mathbb{F})$. If $z \neq 1$, then the image of $x_1x_2 - (zI_2)(x_2x_1)$ is $M_2(\mathbb{F})$. Hence, the image of ω is the image of the map $\tau : M_2(\mathbb{F}) \longrightarrow M_2(\mathbb{F})$ given by $x \mapsto \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} x$ when restricted to $\mathfrak{sl}_2(\mathbb{F})$ or $M_2(\mathbb{F})$ depending on $z = 1$ or $z \neq 0, 1$. Any linear transformation maps a subspace to a subspace and hence in either case, the image is a subspace of $M_2(\mathbb{F})$ given by the set $\left\{ \begin{pmatrix} a & b \\ & \end{pmatrix} : a, b \in \mathbb{F} \right\}$.

Thus, we summarise the result of this section in the following proposition.

Proposition 6.4.3. *Let A_1 be a λ-potent matrix in $M_2(\mathbb{F})$. Then the image of the map $\omega(x_1, x_2) = A_1(x_1x_2) - A_2(x_2x_1) \in M_2(\mathbb{F})\langle x_1, x_2 \rangle$ is a vector subspace of $M_2(\mathbb{F})$.*

Proof. The centralizer of A_1 acts by conjugation on A_2, producing the representatives discussed in 6.2. If $A_1 - A_2$ has a non-zero determinant, by Lemma 6.4.1, the image of the map ω evaluated on $M_2(\mathbb{F})$ is always a full space $M_2(\mathbb{F})$. Thus, we discuss the cases when $A_1 - A_2$ is a singular matrix. For $A_2 = \begin{pmatrix} \mu_1 & \\ z & \mu_2 \end{pmatrix}$ with $z \in \mathbb{F}^{\times}$. For $\mu_1 \neq \mu_2$, the image is $M_2(\mathbb{F})$ follows from 6.4.3.1. For $\mu_1 = \mu_2$, the image is again $M_2(\mathbb{F})$ followed by 6.4.3.2. When the representative of A_2 is $\begin{pmatrix} \mu_1 & \\ & \mu_2 \end{pmatrix}$, the image is a subspace and is in fact the full space, by 6.4.3.3. For the representative, $\begin{pmatrix} \mu & z \\ & \mu \end{pmatrix}$, the image is a subspace as seen in 6.4.3.4. Moreover, if A_1 is nilpotent, then the image is a proper subspace. Otherwise, the map $\boldsymbol{\omega}$ is sujrective on $M_2(\mathbb{F})$. $\square$

Chapter 7

Diagonal Map: Central Simple Algebra

The main objective of this chapter is to study diagonal polynomial maps on central simple algebras, with a particular emphasis on their surjectivity and image structure. Our focus is on understanding how the image of a diagonal map depends both on the underlying algebra and on the base field over which it is defined. The results presented here are based on our article [49].

We begin with a detailed analysis of diagonal maps on matrix algebras, which form the foundational and most tractable class of central simple algebras. Specifically, we study matrix algebras over finite fields (of sufficiently large order), the complex field, and the real field.

As a consequence of the results over finite fields, we also obtain corresponding statements for quaternion algebras over finite fields, which are split and hence are isomorphic to $M_2(\mathbb{F}_q)$. The study over the complex and real fields leads to a complete description of diagonal maps over real division algebras. In particular, since the only finite-dimensional division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathscr{O}$, our results collectively cover all possible real division algebra settings.

The chapter culminates in surjectivity results for diagonal maps on matrix algebras over $\mathbb{C}$ and $\mathbb{R}$, quaternion matrix algebras over $\mathbb{H}$, and division octonion algebras over $\mathbb{R}$. Together, these results provide a unified treatment of diagonal polynomial maps across associative and non-associative algebras and establish sharp conditions under which such maps are surjective.

We now summarise the main results of this chapter.

Theorem 7.1.1. *Let $\mathbb{F}=\mathbb{C}$ and $k_1,k_2 \geq 1$ be integers and β be a non-zero element in $\mathbb{C}$. Then, the map $\omega\colon M_n(\mathbb{C})_s \times M_n(\mathbb{C})_s \to M_n(\mathbb{C})$ given by $\omega(x_1,x_2) = x_1^{k_1} + \beta x_2^{k_2}$, where $M_n(\mathbb{C})_s$ is the set of semisimple matrices, is surjective.*

Theorem 7.1.2. *Let $\mathbb{F}=\mathbb{R}$, $k_1 \geq k_2 \geq k_3 \geq 1$ be integers and β,γ be non-zero elements in $\mathbb{R}$. Then, the map $\omega\colon M_n(\mathbb{R})^3 \to M_n(\mathbb{R})$ given by $\omega(x_1,x_2,x_3) = x_1^{k_1} + \beta x_2^{k_2} + \gamma x_3^{k_3}$ is surjective.*

Theorem 7.1.3. *Let $\mathbb{F}=\mathbb{R}$, $k_1 \geq k_2 \geq 1$ be integers and $\beta > 0$ in $\mathbb{R}$. Then, the map $\omega\colon M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$ given by $\omega(x_1,x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective if and only if one of the following holds*

(i) n is even,

(ii) n is odd and one of the k_1 or k_2 is odd.

Further, when n is odd and k_1,k_2 both are even the image is $M_n(\mathbb{R}) \setminus \{\lambda I_n \mid \lambda < 0\}$.

Theorem 7.1.4. *Let $0 \neq \beta \in \mathbb{H}$ and let k_1,k_2 be positive integers. Then the map $\omega(x_1,x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective on $M_n(\mathbb{H})$. In particular, any matrix in $M_n(\mathbb{H})$ can be written as a sum of two k-th powers.*

As a corollary, we obtain the following result over octonion algebra.

Corollary 7.1.5. Let $\mathscr{O}$ be a division octonion algebra over $\mathbb{R}$, and $\beta \in \mathbb{R}$. Let k_1,k_2 be positive integers. Then the map $\omega(x_1,x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective on $\mathscr{O}$. In particular, any element in $\mathscr{O}$ can be written as a sum of two k-th powers.

Theorem 7.1.6. *Let $k_1, k_2 \geq 1$ and $n \geq 2$ be integers and β be a non-zero element in the finite field $\mathbb{F}_q$. Consider the map $\omega \colon M_n(\mathbb{F}_q) \times M_n(\mathbb{F}_q) \to M_n(\mathbb{F}_q)$ given by $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$. Then, there exists a constant $\mathscr{K}(k_1, k_2)$ (which depends only on k_1 and k_2) such that for all $q > \mathscr{K}(k_1, k_2)$, the map ω is surjective.*

7.2 Reduction of solution

In this section, we develop a method to find solutions over the base field by using the existence of solutions over an extension field. We start with the following proposition.

Proposition 7.2.1. *Let $\mathbb{F}$ be a field and $A \in M_n(\mathbb{F})$ has a separable characteristic polynomial. Then, the equation $f(X_1, X_2, \ldots, X_m) = A$ in $M_n(\mathbb{F})$ has a solution if $f(X_1, X_2, \ldots, X_m) = J_{\alpha,l}$ has a solution in $M_l(\mathbb{F}(\alpha))$ for all eigenvalues α of A over $\overline{\mathbb{F}}$.*

We use this proposition to reduce the problem of solving Equation (1.2.1) to that for Jordan matrices. We demonstrate this with an example.

Example 7.2.1. Consider $\begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ & & 0 & -1 \\ & & 1 & 0 \end{pmatrix}$ over $\mathbb{R}$ with characteristic polynomial $(T^2+1)^2$, which becomes $\begin{pmatrix} \iota & 1 \\ & \iota \end{pmatrix}$ over $\mathbb{C}$. We can write this as a sum of squares over $\mathbb{C}$ as follows:

$$\begin{pmatrix} \imath & 1 \\ 0 & \imath \end{pmatrix} = \begin{pmatrix} \zeta_8 & \zeta_8^{-1} \\ 0 & 0 \end{pmatrix}^2 + \begin{pmatrix} 0 & 0 \\ 0 & \zeta_8 \end{pmatrix}^2,$$

where $\zeta_8 = e^{\frac{2\pi}{8}\imath} = \cos\frac{2\pi}{8} + \imath\sin\frac{2\pi}{8} = c + \imath s$, which gives us

$$\begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ & & 0 & -1 \\ & & 1 & 0 \end{pmatrix} = \begin{pmatrix} c & -s & c & s \\ s & c & -s & c \\ & & 0 & 0 \\ & & 0 & 0 \end{pmatrix}^2 + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ & & c & -s \\ & & s & c \end{pmatrix}^2.$$

Our goal is to solve the Equation (1.2.1). This requires us to determine whether the equation $\delta X^{k_1} + \beta Y^{k_2} = A$ has a solution in $M_n(\mathbb{F})$ for a given $A \in M_n(\mathbb{F})$. Note that without loss of generality, we may assume $\delta = 1$. Thus, because of Proposition 7.2.1, our problem is reduced to the following:

Proposition 7.2.2. *The equation $X^{k_1} + \beta Y^{k_2} = A$ has a solution in $M_n(\mathbb{F})$ if the equation $X^{k_1} + \beta Y^{k_2} = J_{\alpha,l}$ has a solution in $M_l(\mathbb{F}(\alpha))$ for all eigenvalues α of A over $\overline{\mathbb{F}}$ and $l \geq 1$.*

In view of this proposition, in Section 7.3 we will explicitly deal with the equation

$$X^{k_1} + \beta Y^{k_2} = A \tag{7.2.1}$$

with $\beta \neq 0$ in the following two cases for A and $n \geq 1$:

- the invertible case when $A = J_{\alpha,n}$ where $\alpha \neq 0$, and
- the nilpotent case when $A = J_{0,n}$.

We further note that A is invertible if and only if none of the $p_i(x)$, appearing in the factorization of the minimal polynomial, is the polynomial x; i.e., in the extension fields, we must work with $J_{\alpha,n}$ with $\alpha \neq 0$.

7.3 Diagonal word

In this section, we look at Equation (7.2.1). Without loss of generality, we may assume $k_1 \geq k_2 \geq 2$. Note that if one of the $k_i = 1$, then the equation always has a solution.

7.3.1 Invertible Elements in the image

We begin with considering $A = J_{\alpha,n}$, i.e, if $X^{k_1} + \beta Y^{k_2} = J_{\alpha,n}$ with $\alpha \neq 0$ has a solution in $M_n(\mathbb{F})$. We have the following:

Lemma 7.3.1. *Let k_1, k_2 be integers and $\alpha \in \mathbb{F}^*$. Suppose the equation $X^{k_1} + \beta Y^{k_2} = \alpha$ has two solutions (a,b) and (c,d) satisfying $a^{k_1} \neq c^{k_1}$ and $b^{k_2} \neq d^{k_2}$ over $\mathbb{F}$. Then, the equation $X^{k_1} + \beta Y^{k_2} = J_{\alpha,n}$ has a solution in $M_n(\mathbb{F})_s$ where $M_n(\mathbb{F})_s$ denotes the set of diagonalisable matrices.*

Proof. For $n = 1$, we are already given the required solution. So, we may assume $n \geq 2$. We have solutions $(a,b),(c,d) \in \mathbb{F} \times \mathbb{F}$ such that $a^{k_1} \neq c^{k_1}$ and $b^{k_2} \neq d^{k_2}$, $\alpha = a^{k_1} + \beta b^{k_2}$ and $\alpha = c^{k_1} + \beta d^{k_2}$. With this in mind, we consider the following block diagonal matrices,

1. When n is even,

$$\mathscr{G}_n = \bigoplus_{n/2} \begin{pmatrix} a^{k_1} & 1 \\ 0 & c^{k_1} \end{pmatrix}, \qquad \mathscr{H}_n = (\beta b^{k_2}) \bigoplus_{(n-2)/2} \begin{pmatrix} \beta d^{k_2} & 1 \\ 0 & \beta b^{k_2} \end{pmatrix} \bigoplus (\beta d^{k_2}).$$

2. When n is odd,

$$\mathscr{G}_n = \bigoplus_{(n-1)/2} \begin{pmatrix} a^{k_1} & 1 \\ 0 & c^{k_1} \end{pmatrix} \bigoplus (a^{k_1}), \qquad \mathscr{H}_n = (\beta b^{k_2}) \bigoplus_{(n-1)/2} \begin{pmatrix} \beta d^{k_2} & 1 \\ 0 & \beta b^{k_2} \end{pmatrix}.$$

Thus, we get $J_{\alpha,n} = \mathscr{G}_n + \mathscr{H}_n$. Since $a^{k_1} \neq c^{k_1}$, and $b^{k_2} \neq d^{k_2}$, we get that $\mathscr{G}_n$ and $\mathscr{H}_n$ both are diagonalisable matrices, in fact, similar to $\text{diag}\{a^{k_1}, c^{k_1}, a^{k_1}, c^{k_1}, \dots\}$ and $\beta\text{diag}\{b^{k_2}, d^{k_2}, b^{k_2}, d^{k_2}, \dots\}$ respectively. Clearly, $\mathscr{G}_n$ is similar to a matrix which is k_1-power of a diagonal matrix, and $\mathscr{H}_n$ is similar to β times k_2-power of a diagonal matrix. Hence $J_{\alpha,n} = B^{k_1} + \beta C^{k_2}$ where $B, C \in M_n(\mathbb{F})_s$. □

We can use Proposition 7.2.1 with this lemma to show when invertible elements are in the image of the diagonal word map. Suppose $A \in M_n(\mathbb{F})$ with a separable characteristic polynomial and each eigenvalue of A over $\bar{\mathbb{F}}$ satisfies the properties of the lemma above, then A is in the image.

Remark 7.3.1. We note that the above proof also works for $\alpha = 0$ as long as we have required solutions over $\mathbb{F}$.

Here is an example that the image of ω could have nilpotent elements.

Example 7.3.2. Consider $\omega \colon M_2(\mathbb{F})_s \times M_2(\mathbb{F})_s \to M_2(\mathbb{F})$ given by $\omega(x_1, x_2) = x_1^2 + x_2^2$. Suppose $char(\mathbb{F}) \neq 2$, and $X^2 + 1 = 0$ has a solution in $\mathbb{F}$, say ι, then we can write $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}^2 + \begin{pmatrix} \iota & 0 \\ 0 & \iota \end{pmatrix}^2$. In the case $char(\mathbb{F}) = 2$ we can write $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^2 + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^2$.

7.3.2 Jordan nilpotent elements in the image when $n > 2k_1$

We are interested in getting nilpotent elements in the image of $X^{k_1} + \beta Y^{k_2}$. We assume $k_1 \geq k_2$. We show that large-size nilpotents are always present in the image. We recall the notion of the Junction matrix from Section 5 of [35].

Definition 7.3.3. Let $n \geq 1$ be a positive integer. Let $(n_1, n_2, \ldots, n_k)$ be partition of n with $1 \leq n_1 \leq n_2 \leq \ldots \leq n_k$. The Junction matrix associated with the given partition of n is

$$\mathfrak{J}_{(n_1,n_2,\ldots,n_k)} := e_{n_1,n_1+1} + e_{(n_1+n_2),(n_1+n_2+1)} + \cdots + e_{(n_1+n_2+\ldots+n_{k-1}),(n_1+n_2+\ldots+n_{k-1}+1)}$$

where $e_{i,j}$ is the matrix with 1 at ij^{th} place and 0 elsewhere.

We begin with the following:

Lemma 7.3.2. *Suppose $n \geq 2k$, and $(n_1, n_2, \ldots, n_k)$ be partition of n with all $n_i \geq 2$. Then, the junction matrix $\mathfrak{J}_{(n_1,n_2,\ldots,n_k)} = \beta.B^k$ for some $B \in M_n(\mathbb{F})$.*

Proof. Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of $\mathbb{F}^n$ and the matrix $\mathfrak{J}_{(n_1,n_2,\ldots,n_k)}$ corresponds to a linear transformation given by mapping $e_{(n_1+n_2+\cdots+n_{i+1})}$ to $e_{(n_1+n_2+\cdots+n_i)}$ and others to 0. Reordering the basis elements to

$$\{e_1, e_2, \ldots, e_{n_{i-1}}, e_{n_{i+2}}, \ldots, e_n, e_{n_1}, e_{(n_1+1)}, e_{(n_1+n_2)}, e_{(n_1+n_2+1)}, \ldots,$$
$$e_{(n_1+n_2+\cdots+n_{k-1})}, e_{(n_1+n_2+\cdots+n_{k-1}+1)}\}$$

gives a conjugate of the junction matrix $\mathfrak{J}_{(n_1,n_2,\ldots,n_k)}$, say C. The matrix

$$C = \left(\bigoplus_{n-2(k-1)} J_{0,1} \right) \bigoplus \left(\bigoplus_{k-1} J_{0,2} \right).$$

We can also see that C is conjugate to βC as C is a nilpotent matrix. Hence,

$\mathfrak{J}_{(n_1,n_2,\ldots,n_k)}$ is conjugate to βC. Now consider $B = \left(\bigoplus_{n-(2k-1)} J_{0,1}\right) \bigoplus J_{0,2k-1}$ and by Lemma 6.1 of [35], B^k is conjugate to C. Therefore, $\mathfrak{J}_{(n_1,n_2,\ldots,n_k)}$ is conjugate to βB^k. □

Theorem 7.3.4. *Let $k_1 \geq k_2 \geq 2$ be positive integers. For $n \geq 2k_1$ the Jordan nilpotent matrix $J_{0,n}$ is in the image of $f(X,Y) = X^{k_1} + \beta Y^{k_2}$.*

Proof. We begin with considering $J_{0,n}^{k_1}$. Let us denote $n' = \lfloor \frac{n}{k_1} \rfloor$ and $n'' = \lceil \frac{n}{k_1} \rceil$. Since $n \geq 2k_1$, we have $n'' \geq n' \geq 2$. We find m such that $m \equiv n \pmod{k_1}$ and $0 \leq m \leq k_1$. Then, from Miller's Lemma (Lemma 2 [60]) we get that $J_{0,n}^{k_1}$ is conjugate to

$$JF(J_{0,n}^{k_1}) = \left(\bigoplus_{k_1-m} J_{0,n'}\right) \bigoplus \left(\bigoplus_{m} J_{0,n''}\right).$$

Now, we consider $\mathfrak{J} = J_{0,n} - JF(J_{0,n}^{k_1})$. The matrix $\mathfrak{J}$ is a junction matrix associated to the following partition of n: $\left(\underbrace{n',\ldots,n'}_{k_1-m},\underbrace{n'',\ldots,n''}_{m}\right)$. From the Lemma 7.3.2 it follows that $\mathfrak{J}$ is conjugate to βB^{k_2}. This completes the proof. □

Thus, all nilpotent matrices of index $\geq 2k_1$ are in the image.

7.3.3 Nilpotent elements in the image

Now we develop another method to get nilpotent matrices in the image depending on the existence of solutions of certain equations over the base field. Let us begin with the following n-by-n matrix for $n \geq 3$ and $\varepsilon \neq 0$ in $\mathbb{F}$,

$$M = M(\varepsilon, \mathbf{x}, \mathbf{y}, z) = \begin{pmatrix} \varepsilon J_{0,(n-1)} & {}^t\mathbf{x} \\ \mathbf{y} & z \end{pmatrix}$$

over the field $\mathbb{F}$ where $\mathbf{x} = (x_1, x_2, \ldots, x_{n-1})$ and $\mathbf{y} = (y_1, y_2, \ldots, y_{n-1})$ are elements of $\mathbb{K}^{n-1}$ and ${}^t\mathbf{x}$ denotes the transpose of the vector $\mathbf{x}$. Note that $\varepsilon J_{0,(n-1)}$ is conjugate to $J_{0,(n-1)}$. The characteristic polynomial of M is

$$\begin{aligned}\chi_M(T) &= T^n - zT^{n-1} - \left(\sum_{i=1}^{n-1} x_i y_i\right) T^{n-2} - \varepsilon \left(\sum_{i=2}^{n-1} x_i y_{i-1}\right) T^{n-3} - \cdots \\ &- \varepsilon^{j-2} \left(\sum_{i=j-1}^{n-1} x_i y_{i-j+2}\right) T^{n-j} - \cdots - \varepsilon^{n-3}(x_{n-2}y_1 + x_{n-1}y_2)T - \varepsilon^{n-2} x_{n-1} y_1.\end{aligned} \tag{7.3.1}$$

We wish to understand when M is a k-power regular semisimple element (i.e., with distinct diagonal entries). We recall that the elementary symmetric polynomials are

$$\mathscr{E}_i(X_1, \ldots, X_n) = \sum_{1 \le j_1 < j_2 < \cdots < j_i \le n} X_{j_1} X_{j_2} \cdots X_{j_i}$$

and $\prod_{r=1}^{n}(T - x_r) = T^n - \mathscr{E}_1(x_1, \ldots, x_n)T^{n-1} + \cdots + (-1)^n \mathscr{E}_n(x_1, \ldots, x_n)$.

Definition 7.3.5. Let $\mathbb{F}$ be a field and $\lambda_1, \ldots, \lambda_n$ be a solution of $X_1^k + X_2^k + \cdots + X_n^k = \alpha$ over $\mathbb{F}$. We say the solution is **regular** if $\lambda_i^k \neq \lambda_j^k$ for $i \neq j$. Further, if none of the λ_i are 0, we say the solution is a non-zero regular solution.

Note that if 0 appears in a regular solution, it can appear at most once. We have the following,

Lemma 7.3.3. *Let $n \geq 3$, and $\mu_1, \ldots, \mu_n$ be a regular solution of $X_1^k + X_2^k + \cdots + X_n^k = z$ over $\mathbb{F}$. Then, for a given $\mathbf{y}$ with $y_1 \neq 0$ (similarly for a given $\mathbf{x}$ with $x_{n-1} \neq 0$) there exists $\mathbf{x}$ (respectively $\mathbf{y}$) such that the matrix $M(\varepsilon, \mathbf{x}, \mathbf{y}, z)$ is conjugate to the k power regular semisimple element* $\mathrm{diag}(\mu_1^k, \ldots, \mu_n^k)$.

Proof. Let us write $M = M(\varepsilon, \mathbf{x}, \mathbf{y}, z)$ and we have expression for $\chi_M(T)$ in the Equation 7.3.1. We require $\chi_M(T) = (T - \mu_1^k)(T - \mu_2^k) \cdots (T - \mu_n^k)$, which leads to having a solution to the following system of equations:

$$\begin{aligned}
z &= \mathscr{E}_1(\mu_1^k,\ldots,\mu_n^k) = \sum \mu_i^k \\
x_1y_1+\cdots+x_{n-1}y_{n-1} &= -\mathscr{E}_2(\mu_1^k,\ldots,\mu_n^k) \\
x_2y_1+x_3y_2+\cdots+x_{n-1}y_{n-2} &= \varepsilon^{-1}\mathscr{E}_3(\mu_1^k,\ldots,\mu_n^k) \\
\vdots &= \vdots \\
x_{n-2}y_1+x_{n-1}y_2 &= (-1)^n\varepsilon^{-(n-3)}\mathscr{E}_{n-1}(\mu_1^k,\ldots,\mu_n^k) \\
x_{n-1}y_1 &= (-1)^{n+1}\varepsilon^{-(n-2)}\mathscr{E}_n(\mu_1^k,\ldots,\mu_n^k) \\
&= (-1)^{n+1}\varepsilon^{-(n-2)}\prod \mu_i^k.
\end{aligned}$$

Clearly, the first equation is satisfied. We write the remaining equations in matrix form as follows:

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_{n-1} \\ 0 & y_1 & \cdots & y_{n-2} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & y_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} -\mathscr{E}_2(\mu_1^k,\ldots,\mu_n^k) \\ \\ \vdots \\ (-1)^{n+1}\varepsilon^{-(n-2)}\mu_1^k\cdots\mu_n^k \end{pmatrix}.$$

Since $y_1 \neq 0$, the matrix is invertible and we have a solution.

Now, we can also write the above equations, taking the bottom one first and thinking of y_i as variables, as follows:

$$\begin{pmatrix} x_{n-1} & 0 & \cdots & 0 \\ x_{n-2} & x_{n-1} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_{n-1} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} (-1)^{n+1}\varepsilon^{-(n-2)}\mu_1^k\cdots\mu_n^k \\ \\ \vdots \\ -\mathscr{E}_2(\mu_1^k,\ldots,\mu_n^k) \end{pmatrix}.$$

Since $x_{n-1} \neq 0$, we have a solution to this equation. □

Corollary 7.3.6. Let $n \geq 3$, $\lambda_1, \dots, \lambda_n$ be a regular solution of $X_1^k + X_2^k + \cdots + X_n^k = 1$ over $\mathbb{F}$ and $\varepsilon \in \mathbb{F}^*$ with $\varepsilon \neq -1$. Then, the matrix $M(\varepsilon, (1+\varepsilon)\mathbf{e}_{n-1}, \mathbf{y}, 1)$ is a k power regular semisimple element where $y_i = \frac{(-1)^{n-i}}{(1+\varepsilon)} \mathscr{E}_{n-i+1}(\lambda_1^k, \dots, \lambda_{n-1}^k)$ and $\mathbf{e}_{n-1} = (0, \dots, 0, 1)$.

Corollary 7.3.7. Let $n \geq 3$, $\mu_1, \dots, \mu_n$ be a regular solution of $X_1^k + X_2^k + \cdots + X_n^k = -\frac{1}{\beta}$ over $\mathbb{F}$ with $\mu_n = 0$. Then, for a given $\mathbf{y}$ with $y_1 \neq 0$ there exists $\mathbf{x}$ with $x_{n-1} = 0$ such that the matrix $M(1, \mathbf{x}, \mathbf{y}, -1)$ is conjugate to β times a k power regular semisimple element.

Proof. The proof follows along the similar lines as for the Lemma 7.3.3 by equating $\chi_M(T)$ to $(T - \beta\mu_1^k) \cdots (T - \beta\mu_n^k)$ by noting that the equation $x_{n-1}y_1 = 0$ will ensure $x_{n-1} = 0$. □

Theorem 7.3.8. *Let $n \geq 3$ and $\mathbb{F}$ be a field with $|\mathbb{F}| > 2$ and suppose*

1. *the equations $X_1^{k_1} + X_2^{k_1} + \cdots + X_n^{k_1} = 1$ has a regular solution, and*
2. *in addition, for $n \geq 3$, $X_1^{k_2} + X_2^{k_2} + \cdots + X_{n-1}^{k_2} = -\frac{1}{\beta}$ has a non-zero regular solution.*

Then, the nilpotent matrix $J_{0,n}$ is in the image of $f(X,Y) = X^{k_1} + \beta Y^{k_2}$.

Proof. It is enough to show that the Jordan nilpotent matrix $J_{0,n}$ is in the image of $f(X,Y) = X^{k_1} + \beta Y^{k_2}$. Let $\varepsilon \in \mathbb{F}^*$ such that $1 + \varepsilon \neq 0$. We write

$$M(\varepsilon, (1+\varepsilon)\mathbf{e}_{n-1}, \mathbf{y}, 1) + M(1, \mathbf{x}, -\mathbf{y}, -1) = \begin{pmatrix} (1+\varepsilon)J_{0,(n-1)} & (1+\varepsilon)\ {}^t\mathbf{e}_{n-1} + {}^t\mathbf{x} \\ 0 & 0 \end{pmatrix}$$

where $\mathbf{e}_{n-1} = (0, \dots, 0, 1)$ and $\mathbf{x} = (x_1, \dots, x_{n-2}, 0)$. The matrix on the right-hand side is conjugate to $J_{0,n}$. The first matrix on the left side is a k_1th power of a diagonalisable matrix (follows from Corollary 7.3.6). It also ensures $y_1 \neq 0$. The second matrix is β times a k_2th power of a diagonalisable matrix (follows from Corollary 7.3.7). □

Theorem 7.3.9. *Let $\mathbb{F}$ be a field with $|\mathbb{F}| > 2$. Suppose the equation $X_1^{k_2} + X_2^{k_2} = -\frac{1}{\beta}$ has a regular solution. Then, the nilpotent matrix $J_{0,2}$ is in the image of $f(X,Y) = X^{k_1} + \beta Y^{k_2}$.*

Proof. Let (λ_1, λ_2) be a regular solution of $X_1^{k_1} + X_2^{k_1} = -\frac{1}{\beta}$. We write

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ y & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & \frac{1}{\beta} \\ -\frac{y}{\beta} & -\frac{1}{\beta} \end{pmatrix}.$$

Now, the first matrix on the right is an idempotent and hence is a k_1-th power of itself. For the second matrix to conjugate to $\mathrm{diag}(\lambda_1^{k_2}, \lambda_2^{k_2})$, we need $\lambda_1^{k_2} + \lambda_2^{k_2} = -\frac{1}{\beta}$ and $\lambda_1^{k_2}\lambda_2^{k_2} = \frac{y}{\beta}$. Determining y gives us the desired result. □

We demonstrate an application of our result over $\mathbb{C}$, which we require in the later section.

Theorem 7.3.10. *Let $\mathbb{F} = \mathbb{C}$ and $k_1, k_2 \geq 1$ be integers and β be a non-zero element in $\mathbb{C}$. Then, the map $\omega\colon M_n(\mathbb{C})_s \times M_n(\mathbb{C})_s \to M_n(\mathbb{C})$ given by $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$, where $M_n(\mathbb{C})_s$ is the set of semisimple matrices, is surjective.*

Proof. Let A be in $M_n(\mathbb{C})$. Then, the Jordan canonical form of A is the direct sum of the Jordan matrices $J_{\alpha,l}$ where $\alpha \in \mathbb{C}$. Now, we look at the equation $X^{k_1} + \beta Y^{k_2} = \alpha$ over $\mathbb{C}$. Take any $a, c \in \mathbb{C}$ such that $a^{k_1} \neq c^{k_1}$ and consider the equations $\beta Y^{k_2} = \alpha - a^{k_1}$ and $\beta Y^{k_2} = \alpha - c^{k_1}$. We can easily find solutions required in the Lemma 7.3.1, thus $J_{\alpha,l}$ is in the image. The same argument works for Jordan nilpotent matrices in the view of remark 7.3.1 (or alternatively we can use Theorem 7.3.8, 7.3.9 for the $\alpha = 0$ case). This proves the required result. □

7.4 Image of diagonal word over $\mathbb{R}$

In this section, we consider the diagonal polynomials with coefficients in $\mathbb{R}$ and look at their image over $M_n(\mathbb{R})$. Our main theorems in this section are as follows:

Theorem 7.4.1. *Let $\mathbb{F} = \mathbb{R}$, $k_1 \geq k_2 \geq k_3 \geq 1$ be integers and β, γ be non-zero elements in $\mathbb{R}$. Then, the map $\omega\colon M_n(\mathbb{R})^3 \to M_n(\mathbb{R})$ given by $\omega(x_1,x_2,x_3) = x_1^{k_1} + \beta x_2^{k_2} + \gamma x_3^{k_3}$ is surjective.*

Theorem 7.4.2. *Let $\mathbb{F} = \mathbb{R}$, $k_1 \geq k_2 \geq 1$ be integers and $\beta > 0$ in $\mathbb{R}$. Then, the map $\omega\colon M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$ given by $\omega(x_1,x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective if and only if one of the following holds*

(i) n is even,

(ii) n is odd and one of the k_1 or k_2 is odd.

Further, when n is odd and k_1, k_2 both are even the image is $M_n(\mathbb{R}) \setminus \{\lambda I_n \mid \lambda < 0\}$.

We may assume that all of the coefficients of the diagonal polynomial are positive. That is, we are dealing with $\delta_1 x_1^{k_1} + \delta_2 x_2^{k_2} + \cdots + \delta_m x_m^{k_m}$ where $\delta_i > 0$ real for all i. Because for $x_1^{k_1} + \beta x_2^{k_2}$ with $\beta < 0$, following a similar argument as for $\mathbb{C}$ in Section 7.1.1, the equations required in the Theorem 7.3.8, 7.3.9 have solutions over $\mathbb{R}$, hence the map given by $x_1^{k_1} + \beta x_2^{k_2}$ would be surjective. In fact, without loss of generality, we may assume that $\delta_i = 1$ as $\delta_i > 0$ has a k_i-th root. Thus, in what follows, we will be dealing with the map given by $x_1^{k_1} + x_2^{k_2} + \cdots + x_m^{k_m}$. The rest of the section is devoted to the proof of these statements.

We begin by recalling a result from Richman (see Theorem 6 [53]), which also uses the work of Griffin and Krusemeyer from:

Theorem 7.4.3 (Richman, Griffin-Krusemeyer)**.** *Let k be a field with characteristic not equal to 2 and let n be odd. Then, a scalar matrix $cI_n \in M_n(\mathbb{F})$ is a sum of two squares if and only if c is a sum of two squares in $\mathbb{F}$.*

Thus, in view of this, we have,

Corollary 7.4.4. Let n be odd and k_1, k_2 both even. Suppose $\beta > 0$ is a real number. Then, a scalar matrix $\lambda I_n \in M_n(\mathbb{R})$ for $\lambda < 0$ can not be written as $A^{k_1} + \beta B^{k_2}$ where $A, B \in M_n(\mathbb{R})$.

Proof. If we can write $\lambda I_n \in M_n(\mathbb{R})$ as $A^{k_1} + \beta B^{k_2}$ then $\lambda I_n \in M_n(\mathbb{R})$ is also a sum of two squares in $M_n(\mathbb{R})$, and then by the above Theorem of Richman λ is a sum of two squares. This is not possible for $\lambda < 0$. □

The rest of the proof is devoted to essentially showing that these are the only exceptions. The proof will be divided into three cases:

Case 1: When one of the k_1 or k_2 is odd.

Case 2: Both k_1 and k_2 are even and n is even.

Case 3: Both k_1 and k_2 are even and n is odd.

7.4.1 Case 1 when one of the k_1 or k_2 is odd

The proof when k_1 or k_2 is odd is simpler. Let $A \in M_n(\mathbb{R})$. Then, A is conjugate to the direct sum of Jordan blocks

1. $J_{\alpha,l}$ where $\alpha \geq 0$ in $\mathbb{R}$,
2. $J_{\alpha,l}$ where $\alpha < 0$ in $\mathbb{R}$, and
3. $J_{p(x),l}$ where $p(x)$ is degree 2 irreducible polynomial over $\mathbb{R}$.

Using Proposition 7.2.1 we can realize $J_{p(x),l}$ of the kind $J_{\lambda,l}$ for some $\lambda \in \mathbb{C}$ where we can use Theorem 7.1.1 to prove the result. Thus, we need to deal with $J_{\alpha,l}$ where $\alpha \in \mathbb{R}$. We may assume k_2 is odd. Note that when $\alpha \neq 0$, we are done using Lemma 7.3.1 as the equation $X^{k_1} + Y^{k_2} = \alpha$ has required solutions (in the view of k_2 being odd).

When $\alpha = 0$, we can use Theorem 7.3.8 and 7.3.9 to get the result as we have solutions of the required kind over $\mathbb{R}$ (once again in view of k_2 being odd).

7.4.2 Case 2: n is even and k_1, k_2 both even

In this case, we wish to show that the equation $p(x_1,x_2) = x_1^{k_1} + x_2^{k_2} = A$ where A is in $M_n(\mathbb{R})$ and n is even, always has a solution in $M_n(\mathbb{R})$. Since the equation is closed under conjugation, we may consider A in its canonical form. The blocks appearing in the canonical form of A will be as follows:

1. $J_{\alpha,2m}$ where $\alpha \in \mathbb{R}$ and $m \geq 1$.
2. $J_{\alpha_1,2m_1-1} \oplus J_{\alpha_2,2m_2-1}$ for $\alpha_1, \alpha_2 \in \mathbb{R}$ and $m_1, m_2 \geq 1$ (because n is even).
3. $J_{f(x),m}$ where f is an irreducible polynomial of degree 2 over $\mathbb{R}$, and the particular case $m = 1$ refers to the 2×2 companion matrix $\mathfrak{C}_f$.

Denote $\tau_m = \underbrace{(\tau \oplus \tau \oplus \cdots \oplus \tau)}_{m-times}$ where $\tau = \begin{pmatrix} \cos\frac{\pi}{k_1} & -\sin\frac{\pi}{k_1} \\ \sin\frac{\pi}{k_1} & \cos\frac{\pi}{k_1} \end{pmatrix}$ then $\tau_m^{k_1} = -I_{2m}$. Thus, for a real number $\xi > 0$, $\left(\xi^{\frac{1}{k_1}} \tau_m\right)^{k_1} = -\xi I_{2m}$. Also, we note that for $\xi > 0$ and $\eta = \xi^{\frac{1}{k_2}}$, the Jordan matrix $J_{\eta,s}$ has the property that $(J_{\eta,s})^{k_2}$ is conjugate to $J_{\xi,s}$.

In view of the discussion above, we need to deal with the three kinds of blocks. In each of these cases, we show that the matrix is in the image of $\omega(x_1,x_2)$.

1. When $\alpha > 0$, we know $J_{\alpha,2m}$ has k_1-th root, so we are done. When $\alpha \leq 0$ pick $\psi > 0$ such that $\psi + \alpha > 0$, we write

$$J_{\alpha,2m} = -\psi I_{2m} + J_{(\alpha+\psi),2m} = \left(\psi^{\frac{1}{k_1}} T_m\right)^{k_1} + J_{(\alpha+\psi),2m}.$$

   Now, note that $J_{\alpha+\psi,2m}$ is a k_2-th power as it is a conjugate of $\left(J_{\sqrt[k_2]{(\alpha+\psi)},2m}\right)^{k_2}$.

2. In the case of $J_{\alpha_1,2m_1-1} \oplus J_{\alpha_2,2m_2-1}$ if both α_1, α_2 are positive it has k_1-th root. Else, we pick $\psi > 0$ such that $\psi + \alpha_1 > 0$ and $\psi + \alpha_2 > 0$ and write

$$J_{\alpha_1,2m_1-1} \oplus J_{\alpha_2,2m_2-1} = -\psi I_{2(m_1+m_2-1)} + (J_{\alpha_1+\psi,2m_1-1} \oplus J_{\alpha_2+\psi,2m_2-1}).$$

Now, $-\psi I_{2(m_1+m_2-1)}$ is a k_1-th root and $J_{\alpha_1+\psi,2m_1-1} \oplus J_{\alpha_2+\psi,2m_2-1}$ is a k_2-th root.

3. In this case we can use Lemma 7.3.1 and go over the extension $\mathbb{C}$ where we can find the solution and realise it over $\mathbb{R}$.

7.4.3 Case 3: n is odd and k_1, k_2 both even

In this case, we wish to show that the equation $p(x_1,x_2) = x_1^{k_1} + x_2^{k_2} = A$ where A is in $M_n(\mathbb{R})$ and n is odd, has a solution in $M_n(\mathbb{R})$ except when A is a negative scalar matrix. The fact that negative scalars are not in the image follows from Corollary 7.4.4. Since the equation is closed under conjugation, we can work with the canonical form of A. The canonical form of A is a direct sum of the following:

1. $J_{\alpha_1,l_1} \oplus J_{\alpha_2,l_2} \oplus \cdots \oplus J_{\alpha_r,l_r}$ where $\alpha_i \in \mathbb{R}$ and l_i even.
2. $J_{\alpha_1,l_1} \oplus J_{\alpha_2,l_2} \oplus \cdots \oplus J_{\alpha_s,l_s}$ where $\alpha_i \in \mathbb{R}$ and l_i odd (s must be odd as n is odd).
3. $\oplus J_{f_i(x),m}$ where f_i is an irreducible polynomial of degree 2 over $\mathbb{R}$. The particular case $i = 1, m = 1$ refers to the 2×2 companion matrix $\mathfrak{C}_f$.

We first prove some lemmas.

Lemma 7.4.1. *Let $m \geq 3$. Let $d_1, d_2, \ldots, d_m$ be non-negative reals. Consider the matrix*

$$M = \begin{pmatrix} -d_1 & 1 & & & & \\ & -d_2 & 1 & & & \\ & & -d_3 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & -d_{m-1} & 1 \\ -a_1 & -a_2 & \cdots & \cdots & -a_{m-1} & -(d_m+1) \end{pmatrix} \in M_m(\mathbb{R}).$$

Then, a_j's can be chosen in such a way that the matrix M is a k-th power of some *diagonalizable matrix in* $M_m(\mathbb{R})$.

Proof. The characteristic polynomial of M is given by $\chi_M(T)$

$$
\begin{aligned}
=&T^m + T^{m-1}\left(1+\sum_{1\le i\le m} d_i\right) + T^{m-2}\left(\sum_{1\le i_1<i_2\le m} d_{i_1}d_{i_2} + \sum_{1\le i\le m-1} d_i + a_{m-1}\right) + \\
&T^{m-3}\left(\sum_{1\le i_1<i_2<i_3\le m} d_{i_1}d_{i_2}d_{i_3} + \sum_{1\le i_1<i_2\le m-1} d_{i_1}d_{i_2} + a_{m-1}\left(\sum_{1\le i\le m-2} d_i\right) + a_{m-2}\right) + \\
&\cdots + T\left(\sum_{1\le i_1<i_2<i_3<\ldots<i_{m-1}\le m} d_{i_1}d_{i_2}\ldots d_{i_{m-1}} + \cdots + a_3\,(d_1+d_2) + a_2\right) + \\
&(d_1d_2\ldots d_{m-1}(d_m+1) + d_1d_2\ldots d_{m-2}a_{m-1} + \ldots + d_1a_2 + a_1).
\end{aligned}
$$

We claim that we can choose $\lambda_1,\ldots,\lambda_{m-2}$ positive reals and λ_{m-1},λ_m, a pair of non-real complex conjugates such that $\chi_M(T) = (T-\lambda_1)\cdots(T-\lambda_{m-2})(T-\lambda_{m-1})(T-\lambda_m)$ and $\lambda_i \neq \lambda_j$ for all $i \neq j$. This will help us ensure that M is conjugate to a regular semisimple element which is a k-th power. For this, we need to solve the following system of equations,

$$
\begin{aligned}
\sum_{1\le i\le m} d_i + 1 &= -\mathscr{E}_1(\lambda_1,\lambda_2\ldots,\lambda_m), \\
\sum_{1\le i_1<i_2\le m} d_{i_1}d_{i_2} + \sum_{1\le I\le m-1} d_i + a_{m-1} &= (-1)^2\mathscr{E}_2(\lambda_1,\ldots,\lambda_{m-1},\lambda_m), \\
\sum_{1\le i_1<i_2<i_3\le m} d_{i_1}d_{i_2}d_{i_3} + \sum_{1\le i_1<i_2\le m-1} d_{i_1}d_{i_2} + a_{m-1}\left(\sum_{1\le i\le m-2} d_i\right) + a_{m-2}& \\
&= (-1)^3\mathscr{E}_3(\lambda_1,\ldots,\lambda_{m-1},\lambda_m), \\
&\vdots \\
\sum_{1\le i_1<i_2<i_3<\ldots<i_{m-1}\le m} d_{i_1}d_{i_2}\ldots d_{i_{m-1}} + \ldots + a_3\,(d_1+d_2) + a_2&
\end{aligned}
$$

$$= (-1)^{m-1}\mathscr{E}_{m-1}(\lambda_1,\ldots,\lambda_{m-1},\lambda_m),$$

$$d_1d_2\ldots d_m + d_1d_2\ldots d_{m-1} + \ldots + a_2d_1 + a_1 = (-1)^m\mathscr{E}(\lambda_1,\ldots,\lambda_{m-1},\lambda_m).$$

Now, we simply pick $\lambda_1,\ldots,\lambda_{m-2}$ positive reals and λ_{m-1},λ_m a pair of non-real complex conjugate such that

$$-\sum_{1\leq i\leq m} d_i - 1 = \mathscr{E}_1(\lambda_1,\ldots,\lambda_{m-2},\lambda_{m-1},\lambda_m) = \sum_{i=1}^{m-2}\lambda_i + \lambda_{m-1} + \lambda_m$$

and $\lambda_i \neq \lambda_j$ for all $i \neq j$. This ensures each λ_i for $1 \leq i \leq m-2$ has a real k-th root. Further note that when k is even, the real part of λ_m has to be negative, and we will see that the requirement of k-th root forces a pair of complex conjugates.

We note that since λ_{m-1} and λ_m are a pair of complex conjugate the $\mathscr{E}_j(\lambda_1,\ldots,\lambda_m)$ are in $\mathbb{R}$. Thus, the above equations determine the value of a_j, which are real numbers. This proves that we can choose a_j in such a way that M is conjugate to $\tilde{M} = diag(\lambda_1,\ldots,\lambda_{m-2}) \oplus \mathfrak{C}_f$ where λ_i are positive reals and $\mathfrak{C}_f$ is companion matrix of the real polynomial $T^2 - (\lambda_{m-1}+\lambda_m)T + \lambda_{m-1}\lambda_m$. Clearly, $\tilde{M}$ has a k-th root in $M_m(\mathbb{R})$ because scalars on the diagonal are positive and the 2×2 block can be thought of as an element in $\mathbb{C}$. Hence, M is a k-th power. □

Next, we have the following,

Lemma 7.4.2. *Let $n \geq 3$ and $d_1,\ldots,d_n$ be non-negative reals. Consider the matrix $T \in M_n(\mathbb{R})$ as follows:*

$$T = \begin{pmatrix} -d_1 & 1 & & & & \\ & -d_2 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & -d_{n-1} & 1 \\ & & & & & -d_n \end{pmatrix}.$$

Then, there exist matrices B and C in $M_n(\mathbb{R})$ such that $T = B^{k_1} + C^{k_2}$.

Proof. Let us consider a matrix $B = \begin{pmatrix} 0 & 0 & \cdots & & 0 \\ 0 & 0 & \cdots & & 0 \\ a_1 & a_2 & \cdots & a_{n-1} & 1 \end{pmatrix}$ and observe that $B^{k_1} = B$. Now, from Lemma 7.4.1 we can choose $a_1, \ldots, a_{n-1}$ such that the matrix $T - B$, which is in the required form, is k_2-th power, say C^{k_2}. Thus $T = B^{k_1} + C^{k_2}$. □

Proposition 7.4.1. *Let n be odd and $A \in M_n(\mathbb{R})$. Suppose the odd-size Jordan blocks appearing in the canonical form of A are all of the size ≥ 3. Then, A is in the image of $x_1^{k_1} + x_2^{k_2}$.*

Proof. The Jordan blocks of the kind $J_{\alpha,2l}$ and $J_{f(x),l}$ can be taken care of as in Subsection 7.4.2 as they are of even size. For the Jordan blocks of the kind $J_{\alpha,2l-1}$, we can use Lemma 7.4.2 as they are of size ≥ 3. □

Lemma 7.4.3. *All diagonal matrices in $M_m(\mathbb{R})$, when m is even, and the diagonal matrices with at least 2 distinct diagonals, when m is odd, are in the image of $x_1^{k_1} + x_2^{k_2}$.*

Proof. When m is even, the diagonal matrices belonging to the image are covered in Section 7.4.2. Now, for m odd, we are done if even one of the diagonal entries is positive. So, we may assume all of the diagonal entries are negative. We need to deal with $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \ldots)$ where $\lambda_i < 0$ for all i and $\lambda_1 \neq \lambda_2$. Note that $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \ldots)$ is similar to $\begin{pmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \ddots \end{pmatrix}$ (because $\lambda_1 \neq \lambda_2$).

Now, we consider the matrix $L = \begin{pmatrix} 0 & 0 \\ a_1 & a_0 \end{pmatrix}$ and note that

$$M := \begin{pmatrix} \lambda_1 & 1 \\ & \lambda_2 \end{pmatrix} - L^{k_1} = \begin{pmatrix} \lambda_1 & 1 \\ -a_1 a_0^{k_1 - 1} & \lambda_2 - a_0^{k_1} \end{pmatrix}.$$

We claim that we can choose a_0 and a_1 such that the characteristic polynomial of the above matrix M is $\chi_M(T) = (T + \lambda_1)(T - \mu)$ with $\mu \neq -\lambda_1$. For this we need to have $-\lambda_1 + \mu = \lambda_1 + \lambda_2 - a_0^{k_1}$ and $-\lambda_1 \mu = \lambda_1(\lambda_2 - a_0^{k_1}) + a_1 a_0^{k_1 - 1}$. The first equation would require $a_0^{k_1} = 2\lambda_1 + \lambda_2 - \mu$ which can be solved by choosing μ so that $2\lambda_1 + \lambda_2 - \mu > 0$ (this means $\mu < 2\lambda_1 + \lambda_2 < 0 \leq -\lambda_1$). The second equation gives a_1. Thus, we can make M similar to $\mathrm{diag}(-\lambda_1, \mu)$ with $\mu \neq -\lambda_1$. Hence,

$$\tilde{M} := \begin{pmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & & \\ & & \lambda_3 & \\ & & & \ddots \end{pmatrix} - \begin{pmatrix} L & & \\ & 0 & \\ & & \ddots \end{pmatrix}^{k_1}$$

is similar to $\mathrm{diag}(-\lambda_1, \mu, \lambda_3, \ldots)$ where $-\lambda_1 > 0$. Now, $\tilde{M}$ has one of the diagonal entries positive (and the remaining part is even size), so it is a k_2-th power by the earlier argument. This completes the proof. □

Lemma 7.4.4. *Let $l > 1$ and $\alpha, \xi \in \mathbb{R}$. The matrices $J_{\alpha,2l} \oplus (\xi)$ and $J_{f,l} \oplus (\xi)$ are in the image of $x_1^{k_1} + x_2^{k_2}$.*

Proof. First we deal with $J_{\alpha,2l} \oplus (\xi)$. From the argument in Section 7.4.2 we know that $J_{\alpha,2l}$ is in the image of $x_1^{k_1} + x_2^{k_2}$ and hence if $\xi > 0$ (it has roots) we are done. Thus, we may assume $\xi < 0$.

Write $m = 2l$ for simplicity and consider $w_1, w_2, \ldots, w_m \in \mathbb{R}$ such that $w_1 \neq 0$. Take L to be a matrix with first $m-1$ rows 0 and the last row to be $(w_m, w_{m-1}, \ldots, w_1)$. Then,

$$J_{\alpha,m} - L^{k_1} = \begin{pmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 0 & \alpha & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \alpha & 1 \\ -w_m w_1^{k_1-1} & -w_{m-1} w_1^{k_1-1} & \cdots & \cdots & \alpha - w_1^{k_1} \end{pmatrix}.$$

We claim that we can choose $w_1, w_2, \ldots, w_m$ in such a way that the characteristic polynomial of $J_{\alpha,m} - L^{k_1}$ is $(X+\xi)^{m-1}(X-\lambda)$ with $\lambda \neq -\xi$. Following a similar calculation as in the proof of Lemma 7.4.1, we need to ensure $tr(J_{\alpha,m} - L^{k_1}) = -(m-1)\xi + \lambda$, that is, $m\alpha - w_1^{k_1} = -(m-1)\xi + \lambda$. Thus, we need to have a solution for $w_1^{k_1} = m\alpha + (m-1)\xi - \lambda$ which can be insured by choosing $\lambda < 0$, in fact take $\lambda < \xi < 0$. This allows us that $J_{\alpha,m} - L^{k_1}$ is conjugate to $M_1 \oplus (\lambda)$ where only eigen values of M_1 is $-\xi$ which is positive. Thus, $M_1 = M^{k_2}$ for some M. Now, let us write $u^{k_1} = \xi - \lambda > 0$ and $\lambda = -v^{k_1}$. Then,

$$\begin{pmatrix} J_{\alpha,m} & \\ & \xi \end{pmatrix} - \begin{pmatrix} L & \\ & u \end{pmatrix}^{k_1} = \begin{pmatrix} J_{\alpha,m} - L^{k_1} & \\ & \xi - u^{k_1} \end{pmatrix}$$

which is conjugate to $\begin{pmatrix} M^{k_2} & & \\ & \lambda & \\ & & \lambda \end{pmatrix}$. Now all we need to show is that λI_2 is a k_2-th power where $\lambda < 0$. Equivalently, enough to show $-I_2$ is a k_2-th power. This can be done using $-I = \begin{pmatrix} \cos\frac{\pi}{k_2} & -\sin\frac{\pi}{k_2} \\ \sin\frac{\pi}{k_2} & \cos\frac{\pi}{k_2} \end{pmatrix}^{k_2}$.

Now, let us deal with $J_{f,l} \oplus (\xi)$. Once again, from the argument in Section 7.4.2

we know that $J_{f,l}$ is in the image of $x_1^{k_1}+x_2^{k_2}$ and hence if $\xi > 0$ we are done. We need to deal with the case when $\xi < 0$.

First, we consider the case of $l=1$ with $f(x)=x^2+b_0x+b_1$ and $\mathfrak{C}_f := \begin{pmatrix} 0 & -b_1 \\ 1 & -b_0 \end{pmatrix} \in M_2(\mathbb{R})$. Let w_1, w_2 be two real numbers with w_1 non-zero. Consider the matrix $L = \begin{pmatrix} 0 & 0 \\ w_2 & w_1 \end{pmatrix}$. Then,

$$\begin{pmatrix} \mathfrak{C}_f & \\ & \xi \end{pmatrix} - \begin{pmatrix} L & \\ & u \end{pmatrix}^{k_1} = \begin{pmatrix} 0 & -b_1 & 0 \\ 1-w_2w_1^{k_1-1} & -b_0-w_1^{k_1} & 0 \\ 0 & 0 & \xi-u^{k_1} \end{pmatrix}$$

and we claim that with a choice of w_1, w_2 and u we can make this a k_2-th power. Note that the characteristic polynomial of $\mathfrak{C}_f - L^{k_1}$ is $T^2+(b_0+w_1^{k_1})T+b_1(1-w_2w_1^{k_1-1})$. To make $\mathfrak{C}_f - L^{k_1}$ conjugate to a diagonal matrix $\operatorname{diag}(-\xi, \lambda)$ with $\lambda \neq -\xi$ we need to equate trace and determinant, i.e, $-\xi+\lambda = -b_0 - w_1^{k_1}$ and $-\xi\lambda = b_1(1-w_2w_1^{k_1-1})$. We fix, $\lambda < 0$ such that $\xi - \lambda > 0$ and $-b_0+\xi-\lambda > 0$, which ensures, solution for $u^{k_1} = \xi - \lambda > 0$ (to get u) and $w_1^{k_1} = -b_0 + u^{k_1}$. The second equation gives w_2 and with this choice of u, w_1 and w_2 we get:

$$\begin{pmatrix} \mathfrak{C}_f & \\ & \xi \end{pmatrix} - \begin{pmatrix} L & \\ & u \end{pmatrix}^{k_1} = \begin{pmatrix} -\xi & & \\ & \lambda & \\ & & \lambda \end{pmatrix} =: M_2.$$

Now, $-\xi > 0$, which has k_2-th root and $\lambda < 0$, we can use the earlier trick on the 2×2 block by using the k_2-th root of $-I$ to get the job done.

Now, for $l > 1$, let L be a size 2 matrix as above for some w_1 and w_2. Consider

$$Q := \begin{pmatrix} \mathfrak{C}_f & I & & & & \\ & \mathfrak{C}_f & I & & & \\ & & \ddots & \ddots & & \\ & & & \mathfrak{C}_f & I & \\ & & & & \mathfrak{C}_f & \\ & & & & & \xi \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & L & 0 \\ 0 & 0 & 0 & \dots & 0 & u \end{pmatrix}^{k_1}.$$

Then, the characteristic polynomial of Q is $f(T)^{l-1}(T+\xi)(T-\lambda)^2$ equal to the minimal polynomial and hence Q is similar to $J_{f,l-1} \oplus M_2$. Since $J_{f,m-1}$ is k_2-th power and M_2 is k_2-th power so is Q. This completes the proof. □

7.4.4 Proof of the Theorem

Proof of Theorem 7.1.3. When n is even, the proof follows from the argument in Section 7.4.1 and 7.4.2. Now, suppose $n \geq 3$ is odd, and k_1, k_2 are both even. The negative real scalar matrices are not in the image, as follows from Corollary 7.4.4. It remains to show that every matrix in $M_n(\mathbb{R})$ is in the image of $x_1^{k_1} + x_2^{k_2}$ unless it is of the form $-\lambda I_n$ where λ is a positive real.

Let $A \in M_n(\mathbb{R})$. From Proposition 7.4.1, if all odd-size Jordan blocks appearing in the canonical form of A are of size ≥ 3, then we are done. So, we may assume that the odd-sized Jordan blocks are of size 1. Since n is odd, there has to be at least one block of size 1. Now we claim that if A has at least 1 even-size Jordan block, then A is in the image. For this, we can combine one even-size block with a size 1 block and use Lemma 7.4.4, and the remaining parts will be even-size blocks.

Thus, we are left with the case when A has no blocks of even size (and no blocks of odd size ≥ 3). That is, A is a diagonal matrix. Once again, if A has at least 2 distinct entries on the diagonal, we are done with Lemma 7.4.3. Thus, A must be a scalar matrix of the form λI_n. □

Proof of Theorem 7.1.2. From the proof above all we need to show that λI_n when $\lambda < 0$ and n odd is of the form $x_1^{k_1} + x_2^{k_2} + x_3^{k_3}$. For this we write $\lambda I_n = \mathrm{diag}(\lambda, \ldots, \lambda, 0) + \mathrm{diag}(0, \ldots, 0, \lambda)$. Here, the first one is k_1-th power using $\tau_{\frac{n-1}{2}}$ from Section 7.4.2. For the second one, again we use the argument from Section 7.4.2 on $\begin{pmatrix} 0 & \\ & \lambda \end{pmatrix}$ and write it as a sum of k_2 and k_3 powers. □

7.5 Images of diagonal polynomial over real quaternions

In this section, we look at the diagonal polynomial over $M_n(\mathbb{H})$ where $\mathbb{H}$ is Hamilton's real quaternion division algebra. We show that the map $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective when $\beta \neq 0$. This easily implies the surjectivity of the diagonal map for all $m \geq 2$. The result here is surprisingly easy to obtain due to the canonical form theory for matrices in $M_n(\mathbb{H})$. We begin with the following result due to Wiegmann and Liping (See [64, Theorem 1], also [41, Lemma 3]).

Lemma 7.5.1. *Every $n \times n$ matrix with real quaternion elements is similar under a matrix transformation with real quaternion elements to a matrix in (complex) Jordan normal form with diagonal elements of the form $a + bi$, $b \geq 0$. That is to say if $A \in M_n(\mathbb{H})$, then A is similar to a matrix of the form*

$$J(A) := J_{\lambda_1, n_1} \oplus J_{\lambda_2, n_2} \oplus \ldots \oplus J_{\lambda_k, n_k},$$

with $\lambda_k = a_k + ib_k \in \mathbb{C}$ *being right eigenvalues of* A*. Furthermore,* $b_k \in \mathbb{R}$ *can be chosen to be non-negative. In this decomposition* $J(A)$ *is uniquely determined by* A *up to the order of Jordan blocks* J_{λ_k, n_k}*, and* $J(A)$ *is said to be the Jordan canonical form of* A *corresponding to maximal subfield* $\mathbb{C}$ *of* $\mathbb{H}$*.*

This Lemma reduces the problem to look for $A \in M_n(\mathbb{H})$ as an image of diagonal polynomial to that of $A \in M_n(\mathbb{C})$. We call a matrix $A \in M_n(\mathbb{H})$ to be invertible if there exists $B \in M_n(\mathbb{H})$ such that $AB = BA = I$. Note that any matrix $A \in M_n(\mathbb{H})$ has finitely many conjugacy classes of left eigenvalues (i.e. $\alpha \in \mathbb{H}$ such that $A \cdot v = \alpha \cdot v$ for some $v \in \mathbb{H}^n$). Since the number of conjugacy classes in $\mathbb{H}$ are infinite, given any matrix $A \in M_n(\mathbb{H})$, there exists $\lambda \in \mathbb{H}$ such that λ is not a left eigenvalue of A, and consequently $A - \lambda \cdot I$ is invertible (see [54, Proposition 5.3.4]). Now we are ready to state and prove the final result of this article.

Theorem 7.5.1. *Let* $0 \neq \beta \in \mathbb{H}$ *and let* k_1, k_2 *be positive integers. Then the map* $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$ *is surjective on* $M_n(\mathbb{H})$*. In particular, any matrix in* $M_n(\mathbb{H})$ *can be written as a sum of two* k*-th powers.*

Proof. If $A \in M_n(\mathbb{H})$ is invertible it can be written as X^{k_1} for some matrix X in $M_n(\mathbb{H})$, since it is so in $M_n(\mathbb{C})$. Hence A is in the image of ω, by setting x_2 to be the zero matrix. Next, assume $A \in M_n(\mathbb{H})$ is not invertible. Choose $\lambda \in \mathbb{H}$ such that λ^{k_2} does not belong to the set of left eigenvalues of A. Fix $\eta \in \mathbb{H}$ such that $\eta^{k_2} = \beta$. This can be done as η can be conjugated to a complex number, thanks to Lemma 7.5.1. Then for $x_2 = (\lambda/\eta) \cdot I$, the matrix $A - \beta x^{k_2}$ is invertible (since 0 is not a left eigenvalue of $A - \beta x^{k_2}$) and hence the map is surjective by the previous argument. □

Corollary 7.5.2. Let $\mathscr{O}$ be a division octonion algebra over $\mathbb{R}$, and $\beta \in \mathbb{R}$. Let k_1, k_2 be positive integers. Then the map $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$ is surjective on $\mathscr{O}$. In particular, any element in $\mathscr{O}$ can be written as a sum of two k-th powers.

7.6 Image of diagonal polynomial over $\mathbb{F}_q$

We use the results obtained in the previous section to obtain some surjectivity results over finite fields for the diagonal word map. The result in this section is a generalisation of that in [35]. The proof is along similar lines, thus we keep it short. The proof relies on having enough solutions of the equation $X^{k_1} + \beta Y^{k_2} = c$ over the field $\mathbb{F}_q$, for large enough q. The solution of polynomial equations over finite fields has a long history with some fundamental results such as the Chevalley-Warning theorem and Lang-Weil bound, etc. We begin with some of these results regarding the number of solutions that will be used in the main proof. We recall a version of the Lang-Weil theorem [55, Theorem 5A].

Theorem 7.6.1. *Consider the polynomial equation $\delta_1 X_1^{k_1} + \cdots + \delta_m X_m^{k_m} = 1$ where $\delta_i \in \mathbb{F}_q^*$ and $k_i > 0$ for all i. Then the number of solutions S of this equation in $\mathbb{F}_q^m$ satisfies*

$$|S - q^{m-1}| \leq k_1 k_2 \cdots k_m q^{\frac{m-1}{2}} \left(1 - \frac{1}{q}\right)^{-m/2}.$$

The next lemma is along a similar line as Proposition A.3 [35].

Lemma 7.6.1. *For $k_1 \geq k_2 \geq 2$ and $\alpha, \beta \in \mathbb{F}_q^\times$, consider the polynomial*

$$F(X_1, X_2) = X_1^{k_1} + \beta X_2^{k_2} - \alpha.$$

Then, for $q > k_1^4 k_2^4$, there exists solutions (a,b) and (c,d) to $F(X_1, X_2) = 0$ such that $a^{k_1} \neq c^{k_1}$ and $b^{k_2} \neq d^{k_2}$.

Proof. By Theorem 7.6.1, we have the following inequality about the number of solutions S of the equation $F(X_1, X_2) = 0$ and $m = 2$,

$$|S - q| \leq k_1 k_2 \sqrt{q} \left(\frac{q}{q-1}\right).$$

Observe that $\frac{q}{q-1} \leq 2 \leq k_1$. Therefore, we have $|S-q| \leq k_1^2 k_2 \sqrt{q}$. Suppose (a,b) and (c,d) are solutions of $F(X_1,X_2)=0$. If $a^{k_1}=c^{k_1}$, then $F(X_1,X_2)=0$ has at most k_2^2 solutions as $(a,\zeta_{k_2}b)$ and $(c,\zeta_{k_2}d)$, where ζ_{k_2} refers to a root of unity if it exists, are also the possibility for solutions. Similarly, for $b^{k_2}=d^{k_2}$, there are at most k_1^2 solutions possible. So, we need to have

$$S \geq q - k_1^2 k_2 \sqrt{q} \geq k_1^2 + k_2^2 + 1$$

i.e., we want $\sqrt{q}(\sqrt{q}-k_1^2k_2) \geq k_1^2+k_2^2+1$. For this to be satisfied, it suffices to have $\sqrt{q} > k_1^2k_2^2$. In that case, we get $\sqrt{q}(\sqrt{q}-k_1^2k_2) > \sqrt{q} > k_1^2k_2^2 \geq 4k_1^2 \geq k_1^2+k_2^2+1$ as $k_2 \geq 2$. □

Corollary 7.6.2. Let $k_1, k_2 \geq 1$ be integers and $\alpha \in \mathbb{F}_q^*$. Then, there exist a constant $\mathscr{N}_1$ (depending on k_1 and k_2 only) such that for all $q > \mathscr{N}_1$, the matrix $J_{\alpha,n} \in M_n(\mathbb{F}_q)$ can be written as $B^{k_1} + \beta C^{k_2}$ for some $B, C \in M_n(\mathbb{F}_q)$ both diagonalisable.

Proof. Using Lemma 7.6.1, there exists a constant $\mathscr{N}_1$ (depending on k_1, k_2 only) such that for $q > \mathscr{N}_1$, and $\alpha \in \mathbb{F}_q^\times$, there exist solutions $(a,b),(c,d) \in \mathbb{F}_q^2$ such that $a^{k_1} \neq c^{k_1}$ and $b^{k_2} \neq d^{k_2}$, $\alpha = a^{k_1} + \beta b^{k_2}$ and $\alpha = c^{k_1} + \beta d^{k_2}$ for $q > \mathscr{N}_1$. Now we can simply use Lemma 7.3.1 to get the required solution. □

Now, we recall Proposition 2.3 from [34] and Proposition A.2 from [35], which guarantees regular solutions to certain equations over $\mathbb{F}_q$.

Lemma 7.6.2. *Let $\gamma \in \mathbb{F}_q^*$ and $n \geq 2$ be an integer. Then, there exists a constant $\mathscr{N}_2$, depending on k and n, such that for all $q > \mathscr{N}_2$ the equation $X_1^k + X_2^k + \cdots + X_n^k = \gamma$ has a regular solution over $\mathbb{F}_q$. In fact, it always has a non-zero regular solution when $n \geq 3$.*

Proposition 7.6.1. *Let $|\mathbb{F}| > 2$. For every integer $k_1 \geq k_2 \geq 1$, and $\beta \in \mathbb{F}_q^*$ there exists a constant $\mathscr{N}_3$, depending on k_1, k_2 and n only, such that for all $q > \mathscr{N}_3$ the Jordan nilpotent matrix $J_{0,n}$ is in the image of $X^{k_1} + \beta Y^{k_2}$.*

Proof. In view of Lemma above, the required hypothesis of Theorem 7.3.9 and 7.3.8 are satisfied if $q > \mathscr{N}_3$. Note that $\mathscr{N}_3$ is the maximum of the constants required in the hypothesis of the referred Theorems for various choices of k_1 and k_2 for different n. Thus, we have the required result. □

Now we are ready to prove the main result of this section,

Theorem 7.6.3. *Let $k_1, k_2 \geq 1$ and $n \geq 2$ be integers and β be a non-zero element in the finite field $\mathbb{F}_q$. Consider the map $\omega \colon M_n(\mathbb{F}_q) \times M_n(\mathbb{F}_q) \to M_n(\mathbb{F}_q)$ given by $\omega(x_1, x_2) = x_1^{k_1} + \beta x_2^{k_2}$. Then, there exists a constant $\mathscr{K}(k_1, k_2)$ (which depends only on k_1 and k_2) such that for all $q > \mathscr{K}(k_1, k_2)$, the map ω is surjective.*

Proof. In the view of Proposition 7.2.1, the problem is reduced to dealing with $J_{\alpha,l}$ for all extensions of $\mathbb{F}_q$ where $l \leq n$. The case of $\alpha \neq 0$ is covered by Corollary 7.6.2 for all $q > \mathscr{N}_1$ where $\mathscr{N}_1$ depends on k_1 and k_2 only. The case of $J_{0,l}$ for $l > 2k_1$ is covered by Theorem 7.3.4 which works for any q. For the case of $J_{0,l}$ with $l \leq 2k_1$ we use Proposition 7.6.1 which works for $q > \mathscr{N}_2$ depending on k_1, k_2 and l as well. Thus, if we take $q > \mathscr{K}$ where $\mathscr{K}$ is the maximum of $\mathscr{N}_1$ and various $\mathscr{N}_2$ for $l < 2k_1$ (which are finitely many), we get the result. Note that $\mathscr{K}$ depends on k_1 and k_2 only. □

Chapter 8

Polynomial Maps with Constants: Split Octonion

In this chapter, we study polynomial maps with constants on the split octonion algebra over an algebraically closed field. The presence of constants significantly enriches the behavior of such maps and leads to new phenomena that do not occur in the constant-free setting.

This chapter builds directly on the techniques and results developed earlier in the thesis for matrix algebras. In particular, the analysis of polynomial maps with constants on matrix algebras provides a guiding framework for the present study. The split octonion algebra forms a natural next step. It is a unique split eight-dimensional composition algebra and exhibits several matrix-like features, while simultaneously being non-associative. This combination makes it an ideal setting in which to investigate how constants influence the image of polynomial maps beyond the associative case.

Unlike division octonion algebras, the split octonion algebra contains zero divisors and admits a realisation as a Zorn vector matrix algebra. This description allows techniques from matrix algebras and split central simple algebras to be adapted to the octonion setting, while also highlighting the genuinely new difficulties caused by non-associativity. In particular, the interaction between the constants and the isotropic nature of the norm form plays a crucial role in determining the image of a polynomial map.

The main objective of this chapter is to describe the images of polynomial maps with constants when evaluated on the split octonion algebra and to identify conditions under which such maps are surjective. We analyze how the choice of constants influences the resulting image sets. Together with the preceding chapters on matrix and quaternion algebras, this chapter contributes to the study of polynomial maps on non-associative algebras and advances the understanding of image and surjectivity problems in composition algebras.

The results presented in this chapter are based on our article [48]. The structure of the split octonion algebra used here has already been described in Section 4.3 of Chapter 4.

Throughout this chapter, we assume that $\mathbb{F}$ is an algebraically closed field. We now state the main theorem of this chapter.

Theorem 8.1.1. *Let $\mathbb{F}$ be an algebraically closed field, and let k_1, k_2 be positive integers. Let $\mathscr{O}(\mathbb{F})$ be the split octonion algebra over $\mathbb{F}$. Then, the map induced by $A_1(x_1^{k_1}) + A_2(x_2^{k_2})$ on $\mathscr{O}(\mathbb{F})$, where $A_1, A_2 \in \mathscr{O}(\mathbb{F}) \setminus \{0\}$, is surjective if and only if the pair (A_1, A_2) under $G_2(\mathbb{F})$-action does not represent one of the following pairs:*

1. $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}\right)$

2. $\left(\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$

3. $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,\beta_6,0) & 0 \end{pmatrix}\right)$

4. $\left(\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} 0 & (1,0,0) \\ (0,\beta_6,0) & \beta_8 \end{pmatrix}\right)$

5. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}\right)$

6. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} \beta_1 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix}\right)$

7. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$

8. $\left(\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & \mathbf{0} \\ (0,1,0) & 0 \end{pmatrix}\right)$

where $\alpha_1, \alpha_8, \beta_1, \beta_6, \beta_8 \in \mathbb{F}$.

8.2 Orbit Representatives

As discussed in Chapter 5, to study the surjectivity of the map

$$\omega(x_1, x_2) = A_1(x_1^{k_1}) + A_2(x_2^{k_2})$$

on an algebra $\mathscr{A}$, it suffices to study the surjectivity of the map ω associated with a representative of the $\mathbf{Aut}(\mathscr{A})$-orbit of the coefficient tuple (A_1, A_2).

In this chapter, we take $\mathscr{A} = \mathscr{O}$, the split octonion algebra. The automorphism group of $\mathscr{O}$ is the exceptional algebraic group of type G_2. For background on this group and its action on $\mathscr{O}$, we refer to [58]. The orbit representatives of pairs (A_1, A_2) under the action of G_2 were studied by Lopatin and Zubkov; see [43, Theorem 4.1]. For convenience, we recall their result in the following proposition.

Proposition 8.2.1. *Let* $\mathbb{F}$ *be an algebraically closed field. Following are the orbit representative of the* $G_2(\mathbb{F})$*-action on* $\mathscr{O}(\mathbb{F}) \times \mathscr{O}(\mathbb{F})$*, via* $g \cdot (A_1, A_2) = (g(A_1), g(A_2))$*;*

(DD) $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$,

(EK_1) $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_1 \end{pmatrix}\right)$,

$$\text{(FK)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} \right) \textit{ with } \alpha_1 \neq \alpha_8,$$

$$\text{(FN)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} \right) \textit{ with } \alpha_1 \neq \alpha_8 \textit{ and } \beta_5 \neq 0,$$

$$\text{(FP)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} \right) \textit{ with } \alpha_1 \neq \alpha_8,$$

$$\text{(K}_1\text{E)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_1 \end{pmatrix} \right),$$

$$\text{(K}_1\text{F)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} \right) \textit{ with } \beta_1 \neq \beta_8,$$

$$\text{(K}_1\text{L}_1\text{)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & (\beta_2,0,0) \\ \mathbf{0} & \beta_1 \end{pmatrix} \right) \textit{ with } \beta_2 \neq 0,$$

$$\text{(K}_1\text{L}^{\text{T}}\text{)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} \right) \textit{ with } \beta_5 \neq 0,$$

$$\text{(K}_1\text{M)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & (0,1,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} \right),$$

$$\text{(K}_1\text{M}_1^{\text{T}}\text{)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ (0,1,0) & \beta_1 \end{pmatrix} \right),$$

for all $\alpha_1, \alpha_2, \beta_1, \beta_2, \beta_5, \beta_8 \in \mathbb{F}$.

8.3 Roots of an Element

In this section, we study the existence of ℓ-th roots of elements in the split octonion algebra $\mathscr{O}(\mathbb{F})$. Since $\mathscr{O}(\mathbb{F})$ is power-associative, powers of a single element are unambiguously defined. We begin by recording explicit formulas for ℓ-th powers, which will be used repeatedly in what follows.

1. Let
$$X = \begin{pmatrix} \alpha_X & \mathbf{b}_X \\ \mathbf{c}_X & \delta_X \end{pmatrix} \in \mathscr{O}(\mathbb{F}) \quad \text{with} \quad \langle \mathbf{b}_X, \mathbf{c}_X \rangle = 0.$$
Then
$$X^2 = \begin{pmatrix} \alpha_X^2 & (\alpha_X + \delta_X)\mathbf{b}_X \\ (\alpha_X + \delta_X)\mathbf{c}_X & \delta_X^2 \end{pmatrix}.$$
By induction, one verifies that for any positive integer ℓ,
$$X^\ell = \begin{pmatrix} \alpha_X^\ell & \left(\sum_{i=0}^{\ell-1} \alpha_X^i \delta_X^{\ell-1-i}\right) \mathbf{b}_X \\ \left(\sum_{i=0}^{\ell-1} \alpha_X^i \delta_X^{\ell-1-i}\right) \mathbf{c}_X & \delta_X^\ell \end{pmatrix}.$$
For convenience, we define
$$f(\alpha_X, \delta_X, \ell) = \sum_{i=0}^{\ell-1} \alpha_X^i \delta_X^{\ell-1-i},$$
so that
$$X^\ell = \begin{pmatrix} \alpha_X^\ell & f(\alpha_X, \delta_X, \ell)\mathbf{b}_X \\ f(\alpha_X, \delta_X, \ell)\mathbf{c}_X & \delta_X^\ell \end{pmatrix}.$$
2. For a general element

$$Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \\ \mathbf{c}_Y & \delta_Y \end{pmatrix} \in \mathscr{O}(\mathbb{F}),$$

its ℓ-th power has the form

$$Y^\ell = \begin{pmatrix} \tau_1 & \tau\,\mathbf{b}_Y \\ \\ \tau\,\mathbf{c}_Y & \tau_2 \end{pmatrix},$$

where τ, τ_1, and τ_2 are functions depending on α_Y, δ_Y, and $\langle \mathbf{b}_Y, \mathbf{c}_Y \rangle$. We now use these observations to establish the following lemma.

Lemma 8.3.1. *Let*

$$A = \begin{pmatrix} \alpha & \mathbf{b} \\ \\ \mathbf{c} & \delta \end{pmatrix} \in \mathscr{O}(\mathbb{F})$$

with $\alpha \neq \delta$ *and* $\langle \mathbf{b}, \mathbf{c} \rangle = 0$. *Then for any* $\alpha_1 \in \mathbb{F}^\times$ *and any* $k_1 \in \mathbb{Z}_+$, *there exists* $X \in \mathscr{O}(\mathbb{F})$ *such that*

$$A = \alpha_1 X^{k_1}.$$

Proof. Let

$$X = \begin{pmatrix} \alpha_X & \tau_1 \mathbf{b} \\ \\ \tau_2 \mathbf{c} & \delta_X \end{pmatrix},$$

where $\alpha_X, \delta_X, \tau_1$, and τ_2 are elements of $\mathbb{F}$ to be determined. Using the formula from above, we compute

$$A - \alpha_1 X^{k_1} = \begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{c} & \delta \end{pmatrix} - \begin{pmatrix} \alpha_1 \alpha_X^{k_1} & \alpha_1 f(\alpha_X, \delta_X, k_1)\tau_1 \mathbf{b} \\ \alpha_1 f(\alpha_X, \delta_X, k_1)\tau_2 \mathbf{c} & \alpha_1 \delta_X^{k_1} \end{pmatrix}$$

$$= \begin{pmatrix} \alpha - \alpha_1 \alpha_X^{k_1} & (1 - \alpha_1 f(\alpha_X, \delta_X, k_1)\tau_1)\,\mathbf{b} \\ (1 - \alpha_1 f(\alpha_X, \delta_X, k_1)\tau_2)\,\mathbf{c} & \delta - \alpha_1 \delta_X^{k_1} \end{pmatrix}.$$

Choose α_X and δ_X such that

$$\alpha = \alpha_1 \alpha_X^{k_1} \quad \text{and} \quad \delta = \alpha_1 \delta_X^{k_1}.$$

Since $\alpha \neq \delta$, we have $\alpha_X \neq \delta_X$, and hence $f(\alpha_X, \delta_X, k_1) \neq 0$. Now choose

$$\tau_1 = \tau_2 = (\alpha_1 f(\alpha_X, \delta_X, k_1))^{-1}.$$

With this choice, we obtain $A - \alpha_1 X^{k_1} = 0$, completing the proof. □

Remark 8.3.1. It is worth noting that the power map on $\mathscr{O}(\mathbb{F})$ is not surjective, as studied by Lopatin and Rybalov in [42].

8.4 When one of the coefficients is invertible

In this section, we consider polynomial maps with one invertible coefficient. In parallel with the result obtained for the split quaternion algebra in Chapter 6, we aim to establish an analogous statement for the split octonion algebra. Specifically, we prove the following theorem.

Theorem 8.4.1. *Let $A_1, A_2 \in \mathscr{O}(\mathbb{F})$ with $N(A_1) \neq 0$. Then for $A \in \mathscr{O}(\mathbb{F})$ and positive integers $k_1, k_2 \geq 2$, there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that $A = A_1 X^{k_1} + A_2 Y^{k_2}$.*

We assume $N(A_1) \neq 0$ and $N(A_2) \in \mathbb{F}$. The case where $N(A_2) \neq 0$ and $N(A_1) \in \mathbb{F}$ can be derived by considering the pair (A_2, A_1) under the action of G_2.
To prove the existence of X and Y for each $A \in \mathscr{O}(\mathbb{F})$, we consider the following representatives of (A_1, A_2) under the action of G_2 (note that we have combined a few classes which will be dealt with together):

$$\text{(I)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} \right) \qquad \text{(IV)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} \right)$$

$$\text{(II)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} \right) \qquad \text{(V)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & (\beta_2,\beta_3,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} \right)$$

$$\text{(III)} \quad \left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,\beta_6,0) & \beta_8 \end{pmatrix} \right) \qquad \text{(VI)} \quad \left(\begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ (\beta_5,\beta_6,0) & \beta_8 \end{pmatrix} \right)$$

We note that in all the cases mentioned above, the tuples have the following conditions:

1. either $\beta_2 \neq 0$ and $\beta_3 = 0$ with $\beta_1 = \beta_8$, or $\beta_2 = 0$ and $\beta_3 = 1$
2. either $\beta_5 \neq 0$ and $\beta_6 = 0$, or $\beta_5 = 0$ and $\beta_6 = 1$.

Since $N(A_1) \neq 0$, throughout this section we assume $\alpha_1, \alpha_8 \in \mathbb{F}^{\times}$. We begin with some lemmas that will be used to prove the main statement of this section.

Lemma 8.4.1. *For $\beta_1, \beta_8 \in \mathbb{F}$ not simultaneously zero and $A \in \mathscr{O}(\mathbb{F})$, there exist $Y \in \mathscr{O}(\mathbb{F})$ such that $A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}$ is either $\begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & \mathbf{0} \\ \mathbf{c} & \delta \end{pmatrix}$ with $\alpha' \neq \delta$, or $\begin{pmatrix} \alpha & \mathbf{0} \\ \mathbf{c} & \alpha_1^{-1}\alpha_8\delta' \end{pmatrix}$ with $\alpha \neq \delta'$.*

Proof. We consider two cases separately, where $\beta_1 \neq 0$ and $\beta_1 = 0$.

Case 1. For $\beta_1 \neq 0$, take $Y = \begin{pmatrix} \alpha_Y & \tau\mathbf{b} \\ 0 & 0 \end{pmatrix}$. Then,

$$A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha - \beta_1 \alpha_Y^{k_2} & (1 - \beta_1 f(\alpha_Y, \beta_Y, k_2)\tau)\mathbf{b} \\ \mathbf{c} & \delta \end{pmatrix}.$$

Choose $\alpha_Y \neq 0$ such that $\alpha - \beta_1 \alpha_Y^{k_2} = \alpha_8^{-1}\alpha_1\alpha'$ where $\alpha' \neq \delta$. Then $f(\alpha_Y, \delta_Y, k_2) \neq 0$. Choose $\tau = (\beta_1 f(\alpha_Y, \beta_Y, k_2))^{-1}$. We get, $A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & \mathbf{0} \\ \mathbf{c} & \delta \end{pmatrix}$ with $\alpha' \neq \delta$.

Case 2. Now, we take $\beta_1 = 0$ and consider $Y = \begin{pmatrix} 0 & \mathbf{0} \\ \tau\mathbf{c} & \delta_Y \end{pmatrix}$. Then,

$$A - \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & \mathbf{b} \\ (1 - \beta_8 f(\alpha_Y, \beta_Y, k_2)\tau_2)\mathbf{c} & \delta - \beta_8 \delta_Y^{k_2} \end{pmatrix}.$$

Choose $\delta_Y \neq 0$ such that $\delta - \beta_8 \delta_Y^{k_2} = \alpha_1^{-1}\alpha_8\delta'$ where $\delta' \neq \alpha$. Then $f(\alpha_Y, \delta_Y, k_2) \neq 0$ and $\tau = (\beta_8 f(\alpha_Y, \beta_Y, k_2))^{-1}$. We get, $A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & \mathbf{0} \\ \mathbf{c} & \alpha_1^{-1}\alpha_8\delta' \end{pmatrix}$ with $\alpha \neq \delta'$. □

Lemma 8.4.2. *For $\beta_1, \beta_8 \in \mathbb{F}$ and $A \in \mathscr{O}(\mathbb{F})$, there exist $Y \in \mathscr{O}(\mathbb{F})$ such that $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}$ is either $\begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & \mathbf{0} \\ \mathbf{c}' & \delta' \end{pmatrix}$ with $\alpha' \neq \delta'$, or $\begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & (0, b_2, b_3) \\ (c_1', 0, 0) & \delta' \end{pmatrix}$*

with $\alpha' \neq \delta'$.

Proof. We consider two cases, depending on $\beta_1 \neq 0$ and $\beta_1 = 0$.

Case 1. For $\beta_1 \neq 0$, take $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{0} & 0 \end{pmatrix}$. Then,

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha - \beta_1 \alpha_Y^{k_2} & \mathbf{b} - \beta_1 f(\alpha_Y, \delta_Y, k_2)\mathbf{b}_Y \\ \mathbf{c} - (1,0,0) \wedge f(\alpha_Y, \delta_Y, k_2)\mathbf{b}_Y & \delta \end{pmatrix}.$$

Choose $\alpha_Y \neq 0$ such that $\alpha - \beta_1 \alpha_Y^{k_2} = \alpha_8^{-1}\alpha_1\alpha'$ where $\alpha' \neq \delta$. Then $f(\alpha_Y, \delta_Y, k_2) \neq 0$ and take $\mathbf{b}_Y$ such that $\mathbf{b} - \beta_1 f(\alpha_Y, \delta_Y, k_2)\mathbf{b}_Y = 0$. We get, $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & \mathbf{0} \\ \mathbf{c}' & \delta \end{pmatrix}$ with $\alpha' \neq \delta$.

Case 2. For the case $\beta_1 = 0$, consider $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{b}_Y = (0, c_3, c_2)$ and $\mathbf{c}_Y = (\tau, 0, 0)$. Then,

$$A - \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha - f(\alpha_Y, \delta_Y, k_2)\langle \mathbf{c}_Y, (1,0,0)\rangle & \mathbf{b} - \delta_Y^{k_2}(1,0,0) \\ \begin{matrix} \mathbf{c} - \beta_8 f(\alpha_Y, \delta_Y, k_2)\mathbf{c}_Y \\ - f(\alpha_Y, \delta_Y, k_2)(1,0,0) \wedge \mathbf{b}_Y \end{matrix} & \delta - \beta_8 \delta_Y^{k_2} \end{pmatrix}.$$

Choose δ_Y such that $b_1 - \delta_Y^{k_2} = 0$. Choose α_Y such that $f(\alpha_Y, \delta_Y, k_2) = 1$ and hence τ such that $\alpha - \tau = \alpha_8^{-1}\alpha_1\alpha'$ where $\alpha' \neq \delta - \beta_8 b_1$. We get, $A -$

$$\begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & (0,b_2,b_3) \\ (c_1-\beta_8\tau,0,0) & \delta-\beta_8\delta_Y^{k_2} \end{pmatrix}, \text{ as desired.}$$ □

Lemma 8.4.3. *For $\beta_1,\beta_8 \in \mathbb{F}$ and $A \in \mathscr{O}(\mathbb{F})$, there exist $X,Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,\beta_6,0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & (0,b_2',b_3') \\ (c_1',0,0) & \delta' \end{pmatrix}$$

where $\alpha' \neq \delta'$.

Proof. We consider two cases, depending on $\beta_5 \neq 0, \beta_6 = 0$ and $\beta_5 = 0, \beta_6 = 1$.

Case 1. For $\beta_5 \neq 0, \beta_6 = 0$, take $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{b}_Y = (0,c_3,c_2)$ and $\mathbf{c}_Y = (\tau,0,0)$. Then, $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,\beta_6,0) & \beta_8 \end{pmatrix} Y^{k_2}$ is

$$\begin{pmatrix} \alpha-\beta_1\alpha_Y^{k_2}-f(\alpha_Y,\delta_Y,k_2)\langle \mathbf{c}_Y,(1,0,0)\rangle & \mathbf{b}-\beta_1 f(\alpha_Y,\delta_Y,k_2)\mathbf{b}_Y - \delta_Y^{k_2}(1,0,0) \\ \mathbf{c}-(1,0,0)\wedge f(\alpha_Y,\delta_Y,k_2)\mathbf{b}_Y - \alpha_Y^{k_2}(\beta_5,0,0) - \beta_8 f(\alpha_Y,\delta_Y,k_2)\mathbf{c}_Y & \delta-\beta_8\delta_Y^{k_2} \end{pmatrix}.$$

Choose δ_Y such that $b_1-\delta_Y^{k_2}=0$. Choose α_Y such that $f(\alpha_Y,\delta_Y,k_2)=1$ and hence τ such that $\alpha-\beta_1\alpha_Y^{k_2}-\tau=\alpha_8^{-1}\alpha_1\alpha'$ where $\alpha' \neq \delta-\beta_8 b_1$. We get,

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,\beta_6,0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & (0,b_2',b_3') \\ (c_1',0,0) & \delta-\beta_8\delta_Y^{k_2} \end{pmatrix}$$

where $b_2' = b_2 - \beta_1 c_3, b_3' = b_3 - \beta_1 c_2$ and $c_1' = c_1 - \beta_5 \alpha_Y^{k_2} - \beta_8 \tau$.

Case 2. For $\beta_5 = 0, \beta_6 = 1$, take $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{b}_Y = (0, c_2', c_3)$ and $\mathbf{c}_Y = (\tau, 0, 0)$. Then, $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2}$ is

$$\begin{pmatrix} \alpha - \beta_1 \alpha_Y^{k_2} & \mathbf{b} - \beta_1 f(\alpha_Y, \delta_Y, k_2)\mathbf{b}_Y \\ \quad - f(\alpha_Y, \delta_Y, k_2)\langle \mathbf{c}_Y, (1,0,0)\rangle & \quad - \delta_Y^{k_2}(1,0,0) - (0,1,0) \wedge \mathbf{c}_Y \\ \mathbf{c} - f(\alpha_Y, \delta_Y, k_2)\big((1,0,0) \wedge \mathbf{b}_Y\big) & \delta - \beta_8 \delta_Y^{k_2} \\ \quad - \alpha_Y^{k_2}(0,1,0) - \beta_8 \mathbf{c}_Y & \quad - f(\alpha_Y, \delta_Y, k_2)\langle (0,1,0), \mathbf{b}_Y \rangle \end{pmatrix}.$$

Choose δ_Y such that $b_1 - \delta_Y^{k_2} = 0$. Choose α_Y such that $f(\alpha_Y, \delta_Y, k_2) = 1$. Choose $c_2' = c_2 - \alpha_Y^{k_2}$. Then τ such that $\alpha - \beta_1 \alpha_Y^{k_2} - \tau = \alpha_8^{-1} \alpha_1 \alpha'$ where $\alpha' \neq \delta - c_2' - \beta_8 b_1$. We get,

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5, \beta_6, 0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha_8^{-1} \alpha_1 \alpha' & (0, b_2', b_3') \\ (c_1', 0, 0) & \delta - c_2' - \beta_8 \delta_Y^{k_2} \end{pmatrix}$$

where $b_2' = b_2 - \beta_1 c_3, b_3' = b_3 - \beta_1 c_2 + \tau$ and $c_1' = c_1 - \beta_8 \tau$. Note that in the second case, the above proof works for any non-zero β_6. □

Lemma 8.4.4. *For* $\alpha_1, \alpha_8 \in \mathbb{F}^\times$ *and* $A = \begin{pmatrix} \alpha_8^{-1} \alpha_1 \alpha' & \mathbf{b} \\ \mathbf{c} & \delta' \end{pmatrix}$ *such that* $\alpha' \neq \delta'$ *and* $\langle \mathbf{b}, \mathbf{c} \rangle = 0$. *Then there always exists an* $X \in \mathscr{O}(\mathbb{F})$ *such that* $A = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1}$.

Proof. We make use of Equation 3. Multiplying by conjugate of $\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}$ from left in the given equation, we get $\begin{pmatrix} \alpha_1\alpha' & \alpha_8\mathbf{b} \\ \alpha_1\mathbf{c} & \alpha_1\delta' \end{pmatrix} = \alpha_1\alpha_8 X^{k_1}$. By Lemma 8.3.1, there exist $X \in \mathscr{O}(\mathbb{F})$ such that $A - \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} = 0$. □

Theorem 8.4.2. *Let $A_1, A_2 \in \mathscr{O}(\mathbb{F}) \setminus \{0\}$ with $N(A_1) \neq 0$ and A_1 has representative $\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}$ under G_2 action. Then for $A \in \mathscr{O}(\mathbb{F})$ and positive integers $k_1, k_2 \geq 2$, there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = A_1 X^{k_1} + A_2 Y^{k_2}.$$

Proof. The representatives of (A_1, A_2) under the simultaneous action of G_2 are given by

1. $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$,
2. $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix}\right)$
3. $\left(\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}, \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5, \beta_6, 0) & \beta_8 \end{pmatrix}\right)$.

By Lemma 8.4.1, Lemma 8.4.2 and Lemma 8.4.3, $A - A_2Y^{k_2} = \begin{pmatrix} \alpha_8^{-1}\alpha_1\alpha' & \mathbf{b}' \\ \mathbf{c}' & \delta' \end{pmatrix}$ such that $\alpha' \neq \delta'$ and $\langle \mathbf{b}', \mathbf{c}' \rangle = 0$ in each case. By Lemma 8.4.4, there exist $X \in \mathscr{O}(\mathbb{F})$ such that $A = A_1X^{k_1} + A_2Y^{k_2}$. □

Proposition 8.4.1. *Let $A \in \mathscr{O}(\mathbb{F})$ and $\beta_1, \beta_8 \in \mathbb{F}$ not simultaneously zero. Then there exists $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}.$$

Proof. It is enough to consider the case $\beta_1\beta_8 = 0$. When $\beta_8 = 0$ and $\beta_1 \neq 0$, choose $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{0} & 0 \end{pmatrix}$, which gives

$$A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha - \beta_1\alpha_Y^{k_2} & \mathbf{b} - \beta_1\alpha_Y^{k_2-1}\mathbf{b}_Y \\ \mathbf{c} & \delta \end{pmatrix}.$$

Choose $\alpha_Y \neq 0$ such that $\alpha_1(\alpha - \beta_1\alpha_Y^{k_2}) - c_1 \neq \alpha_1\delta$ and $\mathbf{b}_Y$ such that $\mathbf{b}' = \mathbf{b} - \beta_1\alpha_Y^{k_2-1}\mathbf{b}_Y$ satisfies $\alpha_1\mathbf{b}' - \delta(1,0,0) = \mathbf{0}$. Then using Equation 3 we need to find solution for X such that $N(A_1)X^{k_1} = \begin{pmatrix} \alpha_1(\alpha - \beta_1\alpha_Y^{k_2}) - c_1 & \mathbf{0} \\ \alpha_1\mathbf{c} & \alpha_1\delta \end{pmatrix}$, which exists by Lemma 8.3.1.

In case $\beta_1 = 0$ and $\beta_8 \neq 0$, choose $Y = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\delta_Y \neq 0$ and satisfies $\alpha \neq \delta - \beta_8 \delta_Y^{k_2}$, and $\mathbf{c}_Y$ satisfies $\mathbf{c} - \beta_8 \delta_Y^{k_2-1} \mathbf{c}_Y = (1,0,0) \wedge \mathbf{b}$. Then

$$A - \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & \mathbf{b} \\ (1,0,0) \wedge \mathbf{b} & \delta - \beta_8 \delta_Y^{k_2} \end{pmatrix}.$$

Then again using Equation 3, we need existence of X such that

$$N(A_1) X^{k_1} = \begin{pmatrix} \alpha_1 \alpha & \mathbf{b} - (\delta - \beta_8 \delta_Y^{k_2}, 0, 0) \\ \mathbf{0} & \alpha_1 (\delta - \beta_8 \delta_Y^{k_2}) \end{pmatrix}.$$

Such an X exists because of Lemma 8.3.1. □

Proposition 8.4.2. *Let $A \in \mathscr{O}(\mathbb{F})$, $\beta_i \in \mathbb{F}$, for $i = 1,2,3,8$. Then there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (\beta_2, \beta_3, 0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}$$

when one of the following holds:

1. *$\beta_2 \neq 0$, $\beta_3 = 0$ and $\beta_1 = \beta_8$, or*
2. *$\beta_3 = 1$ and $\beta_2 = 0$.*

Proof. Consider the first case, where $\beta_2 \neq 0$, $\beta_3 = 0$, and $\beta_1 = \beta_8$.

Case 1. If $\beta_1 \neq 0$. Let $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{0} & 0 \end{pmatrix}$ where $\alpha_Y \neq 0$ and is such that $\alpha_1(\alpha -$

$\beta_1\alpha_Y^{k_2}) - c_1 \neq \alpha_1\delta$ and $\mathbf{b}_Y$ satisfies $\alpha_1(b - \beta_1\alpha_Y^{k_2-1}\mathbf{b}_Y) - \delta(1,0,0) = 0$. Then

$$A - A_2Y^{k_2} = \begin{pmatrix} \alpha - \beta_1\alpha_Y^{k_2} & \alpha_1^{-1}\delta(1,0,0) \\ (\beta_2,0,0)\wedge\mathbf{b}_Y & \delta \end{pmatrix} = A_1X^{k_1}.$$

Using Equation 3, we get

$$\begin{pmatrix} \alpha_1(\alpha - \beta_1\alpha_Y^{k_2}) - c_1 & \mathbf{0} \\ \mathbf{c}' & \alpha_1\delta \end{pmatrix} = N(A_1)X^{k_1}.$$

Case 2. If $\beta_1 = 0$. Let δ_Y be such that $b_1 - \delta = \delta_Y^{k_2}$ and α_Y be such that $f(\alpha_Y,\delta_Y,k_2) = 1$. Let $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$. Here, $\mathbf{c}_Y = ((c_Y)_1,0,0)$ and $(c_Y)_1$ is such that $\alpha_1(\alpha - (c_Y)_1\beta_2) - c_1 \neq \alpha_1\delta$ and $\mathbf{b}_Y = (0,(b_Y)_2,(b_Y)_3)$ satisfies $\alpha_1(\mathbf{c} - \beta_2\mathbf{b}_Y) - (1,0,0)\wedge\mathbf{b} = (c_1,0,0)$. Then

$$A - A_2Y^{k_2} = \begin{pmatrix} \alpha - (c_Y)_1\beta_2 & (\delta,b_2,b_3) \\ \mathbf{c} - (\beta_2,0,0)\wedge\mathbf{b}_Y & \delta \end{pmatrix} = A_1X^{k_1}.$$

Using Equation 3, we get

$$\begin{pmatrix} \alpha_1(\alpha - (c_Y)_1\beta_2) - c_1 & (0,b_2',b_3') \\ (c_1,0,0) & \alpha_1\delta \end{pmatrix} = N(A_1)X^{k_1}.$$

Consider the second case, $\beta_3 = 1$ and $\beta_2 = 0$. Let δ_Y such that $b_2 = \delta_Y^{k_2}$ and

α_Y such that $f(\alpha_Y, \delta_Y, k_2) = 1$. Let $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{b}_Y = (c_3, 0, c_1)$ and $\mathbf{c}_Y = (0, (c_Y)_2, 0)$ such that $\alpha - \beta_1 \alpha_Y^{k_2} - (c_Y)_2 \neq \delta - \beta_8 b_2$. Then

$$A - A_2 Y^{k_2} = \begin{pmatrix} \alpha - \beta_1 \alpha_Y^{k_2} - (c_Y)_2 & (b_1 - \beta_1 c_3, 0, b_3 - \beta_1 c_1) \\ (0, c_2 - \beta_8 (c_Y)_2, 0) & \delta - \beta_8 b_2 \end{pmatrix} = A_1 X^{k_1}.$$

Using Equation 3, we get

$$\begin{pmatrix} \alpha_1(\alpha - \beta_1 \alpha_Y^{k_2} - (c_Y)_2) & \alpha_1(b_1 - \beta_1 c_3, 0, b_3 - \beta_1 c_1) - (\delta - \beta_8 b_2, 0, 0) \\ (0, \alpha_1(c_2 - \beta_8 (c_Y)_2) + b_3 - \beta_1 c_1, 0) & \alpha_1(\delta - \beta_8 b_2) \end{pmatrix}$$
$$= N(A_1) X^{k_1}.$$

In all the cases discussed above, we have

$$\begin{pmatrix} \alpha' & \mathbf{b}' \\ \mathbf{c}' & \delta' \end{pmatrix} = N(A_1) X^{k_1}$$

where $\alpha' \neq \delta'$ and $\langle \mathbf{b}', \mathbf{c}' \rangle = 0$. By Lemma 8.3.1, there exist X such that $A = A_1 X^{k_1} + A_2 Y^{k_2}$. □

Proposition 8.4.3. *Let $A \in \mathscr{O}(\mathbb{F})$ and $\beta_1, \beta_8 \in \mathbb{F}$. Then there exists $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} \alpha_1 & (1,0,0) \\ \mathbf{0} & \alpha_1 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & \mathbf{0} \\ (\beta_5, \beta_6, 0) & \beta_8 \end{pmatrix} Y^{k_2}$$

where either $\beta_5 \neq 0$ *and* $\beta_6 = 0$ *or* $\beta_5 = 0$ *and* $\beta_6 = 1$.

Proof. Consider the following two cases:

Case 1. When $\beta_5 \neq 0$ and $\beta_6 = 0$. Let $\alpha_Y^{k_1} = \beta_5^{-1} c_1$ and δ_Y be such that $f(\alpha_Y, \delta_Y, k_2) = 1$. For $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{c}_Y = \beta_5^{-1}(0, b_3, b_2)$ and $\mathbf{b}_Y = (\tau, 0, 0)$ such that $\delta - \beta_5\tau - \beta_8\delta_Y^{k_2} \neq \alpha - \beta_1\beta_5^{-1}c_1$, we get

$$\begin{pmatrix} \alpha' & (b_1 - \beta_1\tau, 0, 0) \\ (0, c_2 - \beta_8\beta_5^{-1}b_3, c_3 - \beta_8\beta_5^{-1}b_2) & \delta' \end{pmatrix} = A - A_2Y^{k_2} = A_1X^{k_1}$$

where $\alpha' = \alpha - \beta_1\alpha_Y^{k_2}$, $\delta' = \delta - \beta_5\tau - \beta_8\delta_Y^{k_2}$ and $\alpha' \neq \delta'$. Multiplying the above equation by the conjugate of A_1, we get

$$\begin{pmatrix} \alpha_1\alpha' & (\alpha_1(b_1 - \beta_1\tau) - \delta', 0, 0) \\ \alpha_1(0, c_2 - \beta_8\beta_5^{-1}b_3, c_3 - \beta_8\beta_5^{-1}b_2) & \alpha_1\delta' \end{pmatrix} = N(A_1)X^{k_1}.$$

By using Lemma 8.3.1, there exist $X \in \mathscr{O}(\mathbb{F})$ such that $A = A_1X^{k_1} + A_2Y^{k_2}$.

Case 2. When $\beta_5 = 0$ and $\beta_6 = 1$. Let $\alpha_Y^{k_2} = c_2$ and δ_Y be such that $f(\alpha_Y, \delta_Y, k_2) = 1$. For $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\mathbf{c}_Y = (b_3, 0, b_2 - \tau_2)$ and $\mathbf{b}_Y = (0, \tau_1, 0)$ such that τ_1 satisfies $\alpha_1(\delta - \tau_1 - \beta_8\delta_Y^{k_2}) \neq \alpha_1(\alpha - \beta_1c_2) - c_1\beta_8b_3$ and $\tau_2 = \alpha_1^{-1}(\delta - \tau_1 - \beta_8\delta_Y^{k_2})$. We get

$$\begin{pmatrix} \alpha' & (\tau', b_2 - \beta_1\tau_1, 0) \\ (c_1 - \beta_8b_3, 0, c_3 - \beta_8(b_1 - \tau_2)) & \delta' \end{pmatrix} = A - A_2Y^{k_2} = A_1X^{k_1}$$

where $\alpha' = \alpha - \beta_1 c_2$, $\delta' = \delta - \tau_1 - \beta_8 \delta_Y^{k_2}$. Multiplying the above equation by the conjugate of A_1, we get

$$\begin{pmatrix} \alpha_1\alpha' - c_1 + \beta_8 b_3 & (0, b_2 - \beta_1\tau, 0) \\ \alpha_1(c_1 - \beta_8 b_3, 0, c_3 - \beta_8(b_1 + \tau_2) + b_1\tau) & \alpha_1\delta' \end{pmatrix} = N(A_1)X^{k_1}.$$

Since $\alpha_1\alpha' - c_1 + \beta_8 b_3 \neq \alpha_1\delta'$, using Lemma 8.3.1, there exist $X \in \mathscr{O}(\mathbb{F})$ such that $A = A_1X^{k_1} + A_2Y^{k_2}$. □

8.5 When both of the coefficients are non-invertible

In this section, we look at the case when both A_1 and A_2 are non-unit. We will consider the cases depending on the orbit representatives.

8.5.1 A_1 is non-unit and is diagonal

We first consider the case when A is a non-unit and diagonal.

Lemma 8.5.1. *Let* $\begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{0} & 0 \end{pmatrix} \in \mathscr{O}(\mathbb{F})$ *and* $\alpha_1 \in \mathbb{F}^\times$. *Then for a positive integer* k, *there exists* $X \in \mathscr{O}(\mathbb{F})$ *such that*

$$\begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^k.$$

Proof. Let $X = \begin{pmatrix} \alpha_X & \mathbf{b} \\ \mathbf{0} & \delta_X \end{pmatrix}$. Then, $\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^k$ is given by

$$\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \begin{pmatrix} \alpha_X^k & f(\alpha_X, \delta_X, k)\mathbf{b} \\ \mathbf{0} & \delta_X^k \end{pmatrix} = \begin{pmatrix} \alpha_1 \alpha_X^k & \alpha_1 f(\alpha_X, \delta_X, k)\mathbf{b} \\ \mathbf{0} & 0 \end{pmatrix}.$$

Let α_X be such that $\alpha = \alpha_1 \alpha_X^k$. Choose δ_X such that $f(\alpha_X, \delta_X, k) = \alpha_1^{-1}$. This gives the required element X. □

Lemma 8.5.2. *Let* $\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{c} & \delta \end{pmatrix} \in \mathscr{O}(\mathbb{F})$ *and* $\alpha_8 \in \mathbb{F}^\times$. *Then for a positive integer* k, *there exists* $Y \in \mathscr{O}(\mathbb{F})$ *such that*

$$\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{c} & \delta \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} Y^k.$$

Proof. Let $Y = \begin{pmatrix} \alpha_Y & \mathbf{0} \\ \mathbf{c} & \delta_Y \end{pmatrix}$. Then,

$$\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} Y^k = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} \begin{pmatrix} \alpha_Y^k & \mathbf{0} \\ f(\alpha_Y, \delta_Y, k)\mathbf{c} & \delta_Y^k \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0} \\ \alpha_8 f(\alpha_Y, \delta_Y, k)\mathbf{c} & \alpha_8 \alpha_Y^k \end{pmatrix}.$$

Let δ_Y be such that $\delta = \alpha_8 \delta_Y^k$. Choose α_Y such that $f(\alpha_Y, \delta_Y, k) = \alpha_8^{-1}$. This gives the required element Y. □

Proposition 8.5.1. *Let $\alpha_i, \beta_i \in \mathbb{F}$, where $i = 1,8$ and $A \in \mathscr{O}(\mathbb{F})$. Then there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & \mathbf{0} \\ \mathbf{0} & \beta_8 \end{pmatrix}$$

if and only if $\alpha_1\beta_8 \in \mathbb{F}^\times$ or $\alpha_8\beta_1 \in \mathbb{F}^\times$.

Proof. If $\alpha_1\beta_8 \in \mathbb{F}^\times$ or $\alpha_8\beta_1 \in \mathbb{F}^\times$, then the proof follows from Lemma 8.5.1 and Lemma 8.5.2. Suppose $\alpha_1 = \beta_1 = 0$ or $\alpha_8 = \beta_8 = 0$. Then X, Y exist only if A is of the form $\begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{0} & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{c} & \delta \end{pmatrix}$ depending on the case. □

Proposition 8.5.2. *For $\alpha_1, \beta_8 \in \mathbb{F}^\times$ and $A \in \mathscr{O}(\mathbb{F})$, there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}. \tag{8.5.1}$$

Proof. Let $Y = \begin{pmatrix} \alpha_Y & \mathbf{0} \\ \mathbf{c} & \delta_Y \end{pmatrix}$. Then,

$$A - \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{c} & \delta \end{pmatrix} - \begin{pmatrix} \langle (1,0,0), f(\alpha_Y, \delta_Y, k_2)\mathbf{c} \rangle & \delta_Y^{k_2}(1,0,0) \\ \beta_8 f(\alpha_Y, \delta_Y, k_2)\mathbf{c} & \beta_8 \delta_Y^{k_2} \end{pmatrix}.$$

Let δ_Y be such that $\delta = \beta_8\delta_Y^{k_2}$. Choose α_Y such that $f(\alpha_Y,\delta_Y,k_2) = \beta_8^{-1}$. So,

$$A - \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & \mathbf{b}' \\ \mathbf{0} & 0 \end{pmatrix},$$

for some $\alpha' \in \mathbb{F}$ and $\mathbf{b}' \in \mathbb{F}^3$. Now Lemma 8.5.1 is applicable, and we get the result. □

Proposition 8.5.3. *For $\beta_1, \alpha_8 \in \mathbb{F}^\times$ and $A \in \mathscr{O}(\mathbb{F})$. There exists $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

$$A = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}. \tag{8.5.2}$$

Proof. Let α_Y be such that $\alpha - \beta_1\alpha_Y^{k_2} = 0$ and δ_Y such that $f(\alpha_Y,\delta_Y,k_2) = 1$. Now, let $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{0} & \delta_Y \end{pmatrix}$. Then,

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} 0 & \mathbf{b} - \beta_1\mathbf{b}_Y - \delta_Y^{k_2}(1,0,0) \\ \mathbf{c} - (1,0,0)\wedge\mathbf{b}_Y & \delta - \beta_8\delta_Y^{k_2} \end{pmatrix}.$$

Choose $\mathbf{b}_Y$ such that $\mathbf{b} - \beta_1\mathbf{b}_Y - \delta_Y^{k_2}(1,0,0) = 0$. So,

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{c}' & \delta' \end{pmatrix},$$

for some $\delta' \in \mathbb{F}$ and $\mathbf{c}' \in \mathbb{F}^3$. Now Lemma 8.5.2 is applicable, and we get the result. $\square$

Proposition 8.5.4. *For $\alpha_1, \beta_5, \beta_8 \in \mathbb{F}^\times$ and $A \in \mathscr{O}(\mathbb{F})$, there exist X and Y in $\mathscr{O}(\mathbb{F})$ such that*

1.

$$A = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} Y^{k_2}.$$

2.

$$A = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2}. \tag{8.5.3}$$

Proof. (1) Since $\begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix}$ is a singular element of $\mathscr{O}(\mathbb{F})$, we have $\beta_1\beta_8 = \beta_5 \neq 0$. Take $Y = \begin{pmatrix} \alpha_Y & \mathbf{0} \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$. Then, $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} Y^{k_2}$ is

$$\begin{pmatrix} \alpha - \beta_1\alpha_Y^{k_2} - \langle (1,0,0), f(\alpha_Y,\delta_Y,k_2)\mathbf{c}_Y \rangle & \mathbf{b} - \delta_Y^{k_2}(1,0,0) \\ \mathbf{c} - \alpha_Y^{k_2}(\beta_5,0,0) - \beta_8 f(\alpha_Y,\delta_Y,k_2)\mathbf{c}_Y & \delta - \beta_8\delta_Y^{k_2} \end{pmatrix}.$$

Choose δ_Y such that $\delta - \delta_Y^{k_2} = 0$. Let α_Y such that $f(\alpha_Y,\delta_Y,k_2) = \beta_8^{-1}$. Then $\mathbf{c}_Y = \mathbf{c} - \alpha_Y^{k_2}(\beta_5,0,0)$. We have $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & \mathbf{b}' \\ \mathbf{0} & 0 \end{pmatrix}$. By

Lemma 8.5.1, there exist $X \in \mathscr{O}(\mathbb{F})$ such that $\begin{pmatrix} \alpha' & \mathbf{b}' \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}$.

(2) For $Y = \begin{pmatrix} \alpha_Y & \mathbf{0} \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ let us compute, $A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2} =$

$$\begin{pmatrix} \alpha - \beta_1 \alpha_Y^{k_2} - \langle (1,0,0), \mathbf{c}_Y \rangle & \mathbf{b} - \delta_Y^{k_2}(1,0,0) - (0,1,0) \wedge f(\alpha_Y, \delta_Y, k_2)\mathbf{b}_Y \\ \mathbf{c} - \alpha_Y^{k_2}(0,1,0) - \beta_8 f(\alpha_Y, \delta_Y, k_2)\mathbf{c}_Y & \delta - \beta_8 \delta_Y^{k_2} \end{pmatrix}.$$

Let δ_Y be such that $\delta = \beta_8 \delta_Y^{k_2}$. Choose α_Y such that $f(\alpha_Y, \delta_y, k_2) = \beta_8^{-1}$. Then, $\mathbf{c}_Y = c - \alpha_Y^{k_2}(0,1,0)$. Using Lemma 8.5.1, there exist $X \in \mathscr{O}(\mathbb{F})$ such that

$$A - \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & \mathbf{b}' \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}.$$

□

Proposition 8.5.5. *For $\alpha_8, \beta_5, \beta_1 \in \mathbb{F}^\times$ and $A \in \mathscr{O}(\mathbb{F})$, there exist X and Y in $\mathscr{O}(\mathbb{F})$ such that*

1.

$$A = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} Y^{k_2}.$$

2.

$$A = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2}. \tag{8.5.4}$$

Proof. The proof follows a similar approach to the previous proposition. □

Corollary 8.5.1. Let $\beta_1 \in \mathbb{F}$. When $A_1 = \begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix}$ or $\begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & 0 \end{pmatrix}$ the map $A_1(X^{k_1}) + A_2(Y^{k_2})$ is not surjective.

Proof. From Proposition 8.5.2 we note that if $\beta_8 = 0$ in Eq. (8.5.1), for any choice of $X,Y \in \mathscr{O}(\mathbb{F})$ we have $\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & \mathbf{b} \\ \mathbf{c} & 0 \end{pmatrix}$. Hence, in this case, the map is not surjective.

We note that in the Proposition 8.5.4 above (see Eq. (8.5.3)) if $\beta_8 = 0$, for any $X,Y \in \mathscr{O}(\mathbb{F})$ we have $\begin{pmatrix} \alpha_1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (1,0,0) \\ (0,1,0) & 0 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & \mathbf{b}' \\ (0,c_2',c_3') & \delta' \end{pmatrix}$, which shows that the map is not surjective in this case. □

Corollary 8.5.2. When $A_1 = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix}$ and A_2 is either of the form $\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix}$ or $\begin{pmatrix} 0 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix}$ the map $A_1(X^{k_1}) + A_2(Y^{k_2})$ is not surjective.

Proof. From Proposition 8.5.3 we note that if $\beta_1 = 0$ in Eq. (8.5.2), for any choice of $X,Y \in \mathscr{O}(\mathbb{F})$ we have $\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} + \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & (b_1,0,0) \\ \mathbf{c} & \delta \end{pmatrix}$. Hence, in this case, the map is not surjective. We note that in the Proposition 8.5.5

above (see Eq. (8.5.4)) if $\beta_1 = 0$, for any $X,Y \in \mathscr{O}(\mathbb{F})$ we have $\begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \alpha_8 \end{pmatrix} X^{k_1} +$ $\begin{pmatrix} 0 & (1,0,0) \\ (0,1,0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha & (b_1,0,b_3) \\ \mathbf{c} & \delta \end{pmatrix}$, which shows that the map is not surjective in this case. □

8.5.2 The Remaining non-unit cases

Now, we deal with all the remaining cases. We begin with an auxiliary lemma.

Lemma 8.5.3. *Let* $\begin{pmatrix} \alpha & (b_1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix} \in \mathscr{O}(\mathbb{F})$. *For a positive integer* k_1, *there exist* $X \in \mathscr{O}(\mathbb{F})$ *such that* $\begin{pmatrix} \alpha & (b_1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix} = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}$.

Proof. For $X = \begin{pmatrix} \alpha_X & (0,c_3,c_2) \\ (\alpha,0,0) & \delta_X \end{pmatrix}$ we get,

$$\begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} = \begin{pmatrix} \langle f(\alpha_X,\delta_X,k_1)(\alpha,0,0),(1,0,0)\rangle & \delta_X^{k_2}(1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix}.$$

Choose b_1 such that $b_1 = \delta_X^{k_2}$. Let α_X be such that $f(\alpha_X,\delta_X,k_1) = 1$. This completes the proof. □

Note that the equation $A - \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} = 0$ has a solution if and only if
$A = \begin{pmatrix} \alpha & (b_1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix} \in \mathscr{O}(\mathbb{F})$.

Proposition 8.5.6. *Let $A \in \mathscr{O}(\mathbb{F})$, then there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

1. *for $\beta_5 \in \mathbb{F}^\times$ we have*

$$A = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & \mathbf{0} \\ (\beta_5,0,0) & \beta_8 \end{pmatrix} Y^{k_2}.$$

2. $A = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} 0 & \mathbf{0} \\ (0,1,0) & 0 \end{pmatrix} Y^{k_2}$ *if and only if*

$$A = \begin{pmatrix} \alpha & (b_1,0,b_3) \\ (0,c_2,c_3) & \delta \end{pmatrix}.$$

Proof. For the proof of (1), let $Y = \begin{pmatrix} \alpha_Y & \mathbf{b}_Y \\ \mathbf{c}_Y & \delta_Y \end{pmatrix}$ where $\alpha_Y^{k_2} = \beta_5^{-1} c_1$, $\mathbf{c}_Y = \beta_5^{-1}(0,b_3,b_2)$, δ_Y be such that $f(\alpha_Y, \beta_Y, k_2) = 1$ and $\mathbf{b}_Y = \beta_5^{-1}(\delta - \beta_8 \delta_Y^{k_2}, 0, 0)$. Then,

$$A - \begin{pmatrix} \beta_1 & \mathbf{0} \\ (\beta_5, 0, 0) & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & (b'_1, 0, 0) \\ (0, c'_2, c'_3) & 0 \end{pmatrix}.$$

By Lemma 8.5.3, there exists $X \in \mathscr{O}(\mathbb{F})$ for which the solution of the equation exists.

For the proof of (2), let $Y = \begin{pmatrix} 0 & (0, \delta, 0) \\ (-b_3, 0, 0) & 1 \end{pmatrix}$. Then by Lemma 8.5.3, there exist $X \in \mathscr{O}(\mathbb{F})$ such that

$$\begin{aligned} A - \begin{pmatrix} 0 & \mathbf{0} \\ (0,1,0) & 0 \end{pmatrix} Y^{k_2} &= \begin{pmatrix} \alpha & (b_1, 0, b_3) - (0,1,0) \wedge (-b_3, 0, 0) \\ (0, c_2, c_3) & \delta - \langle (0,1,0), (0, \delta, 0) \rangle \end{pmatrix} \\ &= \begin{pmatrix} \alpha & (b_1, 0, 0) \\ (0, c_2, c_3) & 0 \end{pmatrix} = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}. \end{aligned}$$

□

Proposition 8.5.7. *Let $A \in \mathscr{O}(\mathbb{F})$, then there exist $X, Y \in \mathscr{O}(\mathbb{F})$ such that*

1. *for $\beta_8 \in \mathbb{F}^\times$, the equation $A = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2}$ holds if and only if $A = \begin{pmatrix} \alpha & (b_1, \beta_8^{-1}\delta, 0) \\ \mathbf{c} & \delta \end{pmatrix}$.*

2. *For* $\beta_1 \in \mathbb{F}^\times$*, the equation* $A = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} \beta_1 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2}$ *holds*

   if and only if $A = \begin{pmatrix} \alpha & (b_1, b_2, \beta_1 \tau) \\ (\tau, c_2, c_3) & 0 \end{pmatrix}$.

3. $A = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1} + \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2}$ *if and only if* $A = \begin{pmatrix} \alpha & (b_1, b_2, 0) \\ \mathbf{c} & 0 \end{pmatrix}$.

Proof. For the proof of (1), when $A = \begin{pmatrix} \alpha & (b_1, \beta_8^{-1}\delta, 0) \\ \mathbf{c} & \delta \end{pmatrix}$, let $Y = \begin{pmatrix} \alpha_Y & \mathbf{0} \\ \beta_8^{-1}\mathbf{c} & \delta_Y \end{pmatrix}$.

Then,

$$A - \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & \beta_8 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha - f(\alpha_Y, \delta_Y, k_2)\langle \mathbf{c}, (0,1,0)\rangle & (b_1, \delta - \beta_8^{-1}\delta_Y^{k_2}, 0) \\ \mathbf{c} - f(\alpha_Y, \delta_Y, k_2)\mathbf{c} & \delta - \beta_8 \delta_Y^{k_2} \end{pmatrix}$$

This equals $\begin{pmatrix} \alpha' & (b_1, 0, 0) \\ \mathbf{0} & 0 \end{pmatrix}$, if $\delta_Y^{k_2} = \beta_8^{-1}\delta$ and α_Y be such that $f(\alpha_Y, \delta_Y, k_2) = 1$. The existence of X follows from Lemma 8.5.3. A straightforward calculation shows that the converse is also true.

Now we prove (2). For $A = \begin{pmatrix} \alpha & (b_1, b_2, \beta_1 \tau) \\ (\tau, c_2, c_3) & 0 \end{pmatrix}$, let $Y = \begin{pmatrix} 1 & (0, \beta_1^{-1} b_2, \tau) \\ \mathbf{0} & 0 \end{pmatrix}$.

Then, by Lemma 8.5.3,

$$A - \begin{pmatrix} \beta_1 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2} = \begin{pmatrix} \alpha' & (b_1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix} = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}.$$

The other way follows easily.

Let us quickly see proof of (3). Let $Y = \begin{pmatrix} \alpha_Y & (0,0,c_1) \\ \mathbf{0} & \delta_Y \end{pmatrix}$. Choose δ_Y such that $b_2 = \delta_Y^{k_2}$ and α_Y so that $f(\alpha_Y, \delta_Y, k_2) = 1$. Then by Lemma 8.5.3 $A - \begin{pmatrix} 0 & (0,1,0) \\ \mathbf{0} & 0 \end{pmatrix} Y^{k_2} =$
$\begin{pmatrix} \alpha & (b_1,0,0) \\ (0,c_2,c_3) & 0 \end{pmatrix} = \begin{pmatrix} 0 & (1,0,0) \\ \mathbf{0} & 0 \end{pmatrix} X^{k_1}$ for some $X \in \mathscr{O}(\mathbb{F})$. □

Proof of Theorem 8.1.1. The Proposition 8.2.1 gives the representatives of the orbits of the G_2-action on $\mathscr{O}(\mathbb{F})$. Using Theorem 8.4.1, if either A_1 or A_2 is invertible, then the map is surjective. This covers the cases of (EK_1), ($\mathsf{K}_1\mathsf{E}$) fully. In the following, we summarize the cases where neither A_1 nor A_2 is invertible.

The case of (DD) follows from Proposition 8.5.1. When the orbits are of type (FK), the result follows from Proposition 8.5.2 and Proposition 8.5.3. The cases of (FN) and (FP) are treated in Proposition 8.5.4 and Proposition 8.5.5. When the orbit type is ($\mathsf{K}_1\mathsf{L}^\mathsf{T}$) has been analyzed in Proposition 8.5.6.

Finally, the result now follows from the results about non-surjectivity, which are discussed in Proposition 8.5.1, Theorem 8.5.1, Theorem 8.5.2, Lemma 8.5.3, Proposition 8.5.6, and Proposition 8.5.7. □

Chapter 9

Future Directions and Open Problems

This thesis investigates the surjectivity and image structure of polynomial maps on associative and non-associative algebras, motivated by the classical Waring problem and its algebraic analogues. The principal results provide complete answers to several natural questions concerning diagonal polynomial maps, polynomial maps with constants, and commutator-type maps on matrix algebras and on selected composition algebras. The purpose of this chapter is to outline further directions that arise naturally from the methods and results developed in the preceding chapters. Rather than proposing unrelated problems, the focus here is on extensions that preserve the central themes of this thesis: polynomial images, algebraic structure, and symmetry.

9.1.1 Polynomial Maps on Central Simple Algebras

A recurring theme in this thesis is the study of polynomial maps on matrix algebras over various base fields. Since every central simple algebra over a field is isomorphic to $M_n(D)$ for a division algebra D, matrix algebras serve as a natural starting point for investigating polynomial images on more general central simple algebras.

An important direction for future work is to study diagonal polynomial maps and polynomial maps with constants on algebras of the form $M_n(D)$, with particular emphasis on how invariants of the division algebra D, such as degree, index, and Brauer class, influence surjectivity and image structure. A fundamental question is whether surjectivity phenomena observed over matrix algebras depend only on splitting behavior, or whether they reflect deeper arithmetic properties of the base field.

9.1.2 Field Dependence and Descent of Polynomial Images

Several results in this thesis rely on assumptions about the base field, such as algebraic closedness or sufficiently large cardinality. A natural continuation is to investigate how polynomial images behave under scalar extension and descent.

Key problems include determining which properties of polynomial images are preserved when passing to splitting fields and understanding how surjectivity and image structure vary over finite fields, real closed fields, and imperfect fields. Developing field-independent criteria for surjectivity would provide a more intrinsic understanding of polynomial maps and help isolate the role of arithmetic in polynomial-image phenomena.

9.1.3 Polynomial Maps with Constants and Image Classification

Polynomial maps with constants play a central role in this thesis, particularly in connection with generalisations of the commutator map and analogies with word maps with constants in group theory. The results obtained here suggest that such maps exhibit strong structural constraints on their images.

A natural direction for further study is the systematic classification of images of polynomial maps with constants on matrix algebras and central simple algebras. This includes understanding when the image is a vector subspace, when it coincides with a natural Lie subalgebra, and when it is the entire algebra. These questions are closely related to the Lv́ov–Kaplansky conjecture and its refinements over various fields and algebraic settings.

9.1.4 Extensions to Composition and Non-Associative Algebras

The extension of polynomial-image results from associative algebras to the split octonion algebra demonstrates that surjectivity and image classification phenomena persist in certain non-associative settings. This naturally leads to the study of polynomial maps with constants on other composition algebras and related non-associative structures.

In this context, understanding the role of automorphism groups becomes particularly important. For example, the action of G_2 on the split octonion algebra governs orbit structure and plays a decisive role in image classification. Further investigation of symmetry and orbit stratification may provide a unifying framework for polynomial images in non-associative algebras.

9.1.5 Toward a Structural Theory of Polynomial Images

The results of this thesis, together with the directions outlined above, suggest the possibility of developing a structural theory of polynomial images on algebras. Such a theory would aim to classify possible images of polynomial maps on associative and non-associative algebras and to relate image behavior to algebraic invariants, automorphism group actions, and polynomial identities.

One early step in this direction was taken by Chuang, see [14], who studied the ranges of polynomials evaluated on matrix rings over finite fields $M_n(\mathbb{F}_q)$, characterizing which subsets can occur as images of nonconstant polynomials. This work provides a foundational example of how structural properties of algebras constrain polynomial images.

In particular, this perspective has the potential to contribute to a structural resolution of the Lv́ov–Kaplansky conjecture and to unify Waring-type problems, polynomial maps with constants, and width phenomena within a common conceptual framework.

Appendices

Diagonal Maps with Constants

The table below summarizes the image of the polynomial map

$$\omega(x_1,x_2) = A_1(x_1^{k_1}) + A_2(x_2^{k_2})$$

when evaluated on $M_2(\mathbb{F})$. The matrices A_1 and A_2 are taken from the representatives described in Chapter 6, chosen up to simultaneous conjugation by the automorphism group $\mathrm{PGL}_2(\mathbb{F})$. Since simultaneous conjugation preserves images up to similarity, this reduction yields a complete list of essentially distinct cases. The table records, for each choice of exponents (k_1,k_2) and representatives A_1,A_2, the corresponding image of ω. In particular, it shows that the structure of the image is governed by whether the rows of the constant matrices are zero or nonzero, rather than by the exponents themselves, thereby illustrating the behavior of this class of polynomial maps with constants.

Choice of A_1	Choice of A_2	Image of ω
$\mathrm{diag}(\lambda,\lambda)$	$\mathrm{diag}(\xi,\xi)$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\lambda)$	$\mathrm{diag}(\xi_1,\xi_2), \xi_1 \neq \xi_2$	$M_2(\mathbb{F})$

Choice of A_1	Choice of A_2	Image of ω
$\mathrm{diag}(\lambda,\lambda)$	$\begin{pmatrix} \xi & 1 \\ & \xi \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda \neq \mu$	$\mathrm{diag}(\xi,\xi)$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda \neq \mu$	$\mathrm{diag}(\xi_1,\xi_2), \xi_1\xi_2 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda \neq 0$	$\mathrm{diag}(\xi_1,\xi_2), \xi_2 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \mu \neq 0$	$\mathrm{diag}(\xi_1,\xi_2), \xi_1 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,0)$	$\mathrm{diag}(\xi,0)$	$\begin{pmatrix} * & * \\ & \end{pmatrix}$
$\mathrm{diag}(0,\mu)$	$\mathrm{diag}(0,\xi)$	$\begin{pmatrix} & \\ * & * \end{pmatrix}$
$\mathrm{diag}(\lambda,\mu), \lambda\mu \neq 0$	$\begin{pmatrix} \xi & 1 \\ & \xi \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda\mu = 0$	$\begin{pmatrix} \xi & 1 \\ & \xi \end{pmatrix}, \xi \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \mu \neq 0$	$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$	$M_2(\mathbb{F})$

Choice of A_1	Choice of A_2	Image of ω
$\mathrm{diag}(\lambda,0)$	$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$	$\begin{pmatrix} * & * \\ & \end{pmatrix}$
$\mathrm{diag}(\lambda,\mu),\lambda\mu\neq 0$	$\begin{pmatrix} \xi_1 & 1 \\ & \xi_2 \end{pmatrix},\xi_1\neq\xi_2$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu),\lambda\mu = 0$	$\begin{pmatrix} \xi_1 & 1 \\ & \xi_2 \end{pmatrix},\xi_1\xi_2\neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu),\mu\neq 0$	$\begin{pmatrix} \xi & 1 \\ & 0 \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,0)$	$\begin{pmatrix} \xi & 1 \\ & 0 \end{pmatrix}$	$\begin{pmatrix} * & * \\ & \end{pmatrix}$
$\mathrm{diag}(\lambda,0)$	$\begin{pmatrix} 0 & 1 \\ & \xi \end{pmatrix},\xi\neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(0,\mu)$	$\begin{pmatrix} 0 & 1 \\ & \xi \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu),\lambda\mu\neq 0$	$\begin{pmatrix} \xi & \\ 1 & \xi \end{pmatrix}$	$M_2(\mathbb{F})$

Choice of A_1	Choice of A_2	Image of ω
$\mathrm{diag}(\lambda,\mu),\lambda\mu=0$	$\begin{pmatrix} \xi & \\ 1 & \xi \end{pmatrix}, \xi \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu),\lambda\neq 0$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(0,\mu)$	$\begin{pmatrix} & \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} & \\ * & * \end{pmatrix}$
$\mathrm{diag}(\lambda,\mu),\lambda\mu\neq 0$	$\begin{pmatrix} \xi_1 & 0 \\ 1 & \xi_2 \end{pmatrix}, \xi_1 \neq \xi_2$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu),\lambda\mu=0$	$\begin{pmatrix} \xi_1 & 0 \\ 1 & \xi_2 \end{pmatrix}, \xi_1\xi_2 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,0)$	$\begin{pmatrix} \xi & \\ 1 & \xi_2 \end{pmatrix}$	$M_2(\mathbb{F})$
$\mathrm{diag}(0,\mu)$	$\begin{pmatrix} \xi_1 & 0 \\ 1 & \xi_2 \end{pmatrix}, \xi_1 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(0,\mu)$	$\begin{pmatrix} 0 & 0 \\ 1 & \xi \end{pmatrix}$	$\begin{pmatrix} & \\ * & * \end{pmatrix}$

Choice of A_1	Choice of A_2	Image of ω
$\mathrm{diag}(\lambda,\mu), \lambda \neq \mu$	$\begin{pmatrix} \xi_1 & \xi_2 \\ 1 & \xi_3 \end{pmatrix}, \xi_i \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda\mu \neq 0$	$\begin{pmatrix} 0 & \xi_2 \\ 1 & \xi_3 \end{pmatrix}, \xi_3 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu)$	$\begin{pmatrix} 0 & \xi_2 \\ 1 & \xi_3 \end{pmatrix}, \xi_i \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(0,\mu)$	$\begin{pmatrix} 0 & 0 \\ 1 & \xi \end{pmatrix}$	$\begin{pmatrix} & \\ * & * \end{pmatrix}$
$\mathrm{diag}(\lambda,\mu), \lambda \neq \mu$	$\begin{pmatrix} \xi_1 & \xi_2 \\ 1 & 0 \end{pmatrix}, \xi_1 \neq 0$	$M_2(\mathbb{F})$
$\mathrm{diag}(\lambda,\mu), \lambda \neq \mu$	$\begin{pmatrix} 0 & \xi \\ 1 & 0 \end{pmatrix}, \xi \neq 0$	$M_2(\mathbb{F})$
$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}, \lambda \neq 0$	$\begin{pmatrix} \xi & \\ z & \xi \end{pmatrix}, \xi z \neq 0$	$M_2(\mathbb{F})$
$\begin{pmatrix} & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} \xi_1 & \\ z & \xi_2 \end{pmatrix}, \xi_1 \neq \xi_2, z \neq 0$	$M_2(\mathbb{F})$

Choice of A_1	Choice of A_2	Image of ω
$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}, \lambda \neq 0$	$\begin{pmatrix} \xi_1 & \\ & \xi_2 \end{pmatrix}, \xi_1 \neq \xi_2$	$M_2(\mathbb{F})$
$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$	$\begin{pmatrix} & z \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} * & * \\ & \end{pmatrix}$
$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}$	$\begin{pmatrix} \xi & z \\ & \xi \end{pmatrix}, \xi \neq 0$	$M_2(\mathbb{F})$

Table .1.1: Images of $A_1 x_1^{k_1} + A_2 x_2^{k_2}$

Bibliography

[1] *"The Dniester Notebook: Unsolved Problems in the Theory of Rings and Modules"*. Fourth. MR1310114 (95k:16001). Novosibirsk, 1993.

[2] A. A. Albert and B. Muckenhoupt. "On matrices of trace zeros". In: *Michigan Math. J.* 4 (1957). MR83961, pp. 1–3.

[3] A. S. Amitsur and J. Levitzki. "Minimal identities for algebras". In: *Proc. Amer. Math. Soc.* 1 (1950), pp. 449–463. DOI: 10.2307/2032312.

[4] S. A. Amitsur and L. H. Rowen. "Elements of reduced trace 0". In: *Israel J. Math.* 87 (1994), pp. 161–179. DOI: 10.1007/BF02772992.

[5] R. Barai and A. S. Garge. "Trace and discriminant criteria for a matrix to be a sum of sixth and eighth powers of matrices". In: *Integers* 22 (2022). MR4390730, Paper No. A25 (14 pages).

[6] G. R. Belitskii and V. V. Sergeichuk. "Complexity of matrix problems". In: vol. 361. Ninth Conference of the International Linear Algebra Society(Haifa, 2001). 2003, pp. 203–222. DOI: 10.1016/S0024-3795(02)00391-9.

[7] G. Berhuy and F. Oggier. *"An introduction to central simple algebras and their applications to wireless communication"*. Vol. 191. Mathematical Surveys and Monographs. Amer. Math. Soc., Providence, RI, 2013, pp. viii+276. DOI: 10.1090/surv/191.

[8] M. Bhargava. "On the Conway-Schneeberger fifteen theorem". In: *Quadratic forms and their applications*. Vol. 272. Contemp. Math. Amer. Math. Soc., Providence, RI, 2000, pp. 27–37. DOI: 10.1090/conm/272/04395.

[9] M. Bhargava, J. E. Cremona, T. Fisher, N. G. Jones, and J. P. Keating. "What is the probability that a random integral quadratic form in n variables has an integral zero?" In: *Int. Math. Res. Not. IMRN* (2016), pp. 3828–3848. DOI: 10.1093/imrn/rnv251.

[10] M. Brešar and P. Šemrl. "The Waring problem for matrix algebras". In: *Israel J. Math.* 253 (2023), pp. 381–405. DOI: 10.1007/s11856-022-2366-7.

[11] M. Brešar. "Commutators and images of noncommutative polynomials". In: *Adv. Math.* 374 (2020), Paper no. 107346 (21 pages). DOI: 10.1016/j.aim.2020.107346.

[12] M. Brešar and P. Šemrl. "The Waring problem for matrix algebras, II". In: *Bull. Lond. Math. Soc.* 55 (2023), pp. 1880–1889. DOI: 10.1112/blms.12825.

[13] L. Carlitz. "Solution to Problem 140". In: *Canadian Mathematical Bulletin* 11 (1968), pp. 165–169.

[14] C.-L. Chuang. "On ranges of polynomials in finite matrix rings". In: *Proc. Amer. Math. Soc.* 110 (1990), pp. 293–302. DOI: 10.2307/2048069.

[15] P. M. Cohn. *Basic algebra*. "Groups, rings and fields". Springer-Verlag London, Ltd., London, 2003, pp. xii+465. DOI: 10.1007/978-0-85729-428-9.

[16] J.-L. Colliot-Thélène and J.-J. Sansuc. "La descente sur les variétés rationnelles". In: *Journées de Géometrie Algébrique d'Angers, Juillet 1979/Algebraic Geometry, Angers, 1979*. MR605344. Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown, Md., 1980, pp. 223–237.

[17] P. S. Fagundes. "The images of multilinear polynomials on strictly upper triangular matrices". In: *Linear Algebra Appl.* 563 (2019), pp. 287–301. DOI: 10.1016/j.laa.2018.11.014.

[18] I. G. Gargate and T. C. de Mello. "Images of multilinear polynomials on $n \times n$ upper triangular matrices over infinite fields". In: *Israel J. Math.* 252 (2022), pp. 337–354. DOI: 10.1007/s11856-022-2350-2.

[19] A. S. Garge. "Matrices over commutative rings as sums of fifth and seventh powers of matrices". In: *Linear Multilinear Algebra* 69 (2021), pp. 2220–2227. DOI: 10.1080/03081087.2019.1664386.

[20] P. Gille and T. Szamuely. *"Central simple algebras and Galois cohomology"*. Second. Vol. 165. Cambridge Studies in Advanced Mathematics. MR3727161. Cambridge University Press, Cambridge, 2017, pp. xi+417.

[21] F. A. Gnutov. "On the image of a word map with constants of a simple algebraic group II". In: *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 492 (2020). MR4223709, pp. 75–93.

[22] F. A. Gnutov and N. L. Gordeev. "On the image of a word map with constants of a simple algebraic group". In: *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 478 (2019). MR4030419, pp. 78–99.

[23] N. L. Gordeev, B. È. Kunyavskiĭ, and E. B. Plotkin. "Word maps and word maps with constants of simple algebraic groups". In: *Dokl. Akad. Nauk* 471 (2016), pp. 136–138. DOI: 10.1134/s1064562416060077.

[24] N. Gordeev, B. Kunyavskiĭ, and E. Plotkin. "Word maps, word maps with constants and representation varieties of one-relator groups". In: *J. Algebra* 500 (2018), pp. 390–424. DOI: 10.1016/j.jalgebra.2017.03.016.

[25] M. Griffin and M. Krusemeyer. "Matrices as sums of squares". In: *Linear and Multilinear Algebra* 5 (1977), pp. 33–44. DOI: 10.1080/03081087708817172.

[26] L. C. Grove. *"Classical groups and geometric algebra"*. Vol. 39. Graduate Studies in Mathematics. Amer. Math. Soc., Providence, RI, 2002, pp. x+169. DOI: 10.1090/gsm/039.

[27] D. Hilbert. "Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem)". In: *Math. Ann.* 67 (1909), pp. 281–300. DOI: 10.1007/BF01450405.

[28] A. Kanel-Belov, B. Kunyavskii, and E. Plotkin. "Word equations in simple groups and polynomial equations in simple algebras". In: *Vestnik St. Petersburg Univ. Math.* 46 (2013), pp. 3–13. DOI: 10.3103/S1063454113010044.

[29] A. Kanel-Belov, S. Malev, and L. Rowen. "The images of multilinear polynomials evaluated on 3×3 matrices". In: *Proc. Amer. Math. Soc.* 144 (2016), pp. 7–19. DOI: 10.1090/proc/12478.

[30] A. Kanel-Belov, S. Malev, and L. Rowen. "The images of non-commutative polynomials evaluated on 2×2 matrices". In: *Proc. Amer. Math. Soc.* 140 (2012), pp. 465–478. DOI: 10.1090/S0002-9939-2011-10963-8.

[31] A. Kanel-Belov, S. Malev, C. Pines, and L. Rowen. "The images of multilinear and semihomogeneous polynomials on the algebra of octonions". In: *Linear Multilinear Algebra* 72 (2024), pp. 178–187. DOI: 10.1080/03081087.2022.2158170.

[32] S. A. Katre and D. Krishnamurthi. "Matrices over non-commutative rings as sums of powers". In: *Linear Multilinear Algebra* 70 (2022), pp. 824–829. DOI: 10.1080/03081087.2020.1748856.

[33] S. A. Katre and K. Wadikar. "Matrices over noncommutative rings as sums of kth powers". In: *Linear Multilinear Algebra* 69 (2021), pp. 2050–2058. DOI: 10.1080/03081087.2019.1659219.

[34] K. Kishore. "Matrix Waring problem". In: *Linear Algebra Appl.* 646 (2022), pp. 84–94. DOI: 10.1016/j.laa.2022.03.022.

[35] K. Kishore and A. Singh. "Matrix Waring problem. II". In: *Israel J. Math.* 267 (2025), pp. 301–320. DOI: 10.1007/s11856-024-2704-z.

[36] J.-L. Lagrange. "Démonstration dún théorème dárithmétique". French. In: *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin* (1770), pp. 123–133.

[37] T. Y. Lam. *"Introduction to quadratic forms over fields"*. Vol. 67. Graduate Studies in Mathematics. Amer. Math. Soc., Providence, RI, 2005, pp. xxii+550. DOI: 10.1090/gsm/067.

[38] M. Larsen, A. Shalev, and P. H. Tiep. "The Waring problem for finite simple groups". In: *Ann. of Math. (2)* 174 (2011), pp. 1885–1950. DOI: 10.4007/annals.2011.174.3.10.

[39] J. Lee. "Integral matrices as diagonal quadratic forms". In: *Linear Multilinear Algebra* 66 (2018), pp. 742–747. DOI: 10.1080/03081087.2017.1320965.

[40] D. B. Leep, D. B. Shapiro, and A. R. Wadsworth. "Sums of squares in division algebras". In: *Math. Z.* 190 (1985), pp. 151–162. DOI: 10.1007/BF01160457.

[41] H. Liping. "Consimilarity of quaternion matrices and complex matrices". In: *Linear Algebra Appl.* 331 (2001), pp. 21–30. DOI: 10.1016/S0024-3795(01)00266-X.

[42] A. Lopatin and A. Rybalov. "On polynomial equations over split octonions". In: *Commun. Math.* 33 (2025), Paper No. 8 (12 pages). DOI: 10.46298/cm.14879.

[43] A. Lopatin and A. N. Zubkov. "Classification of $\mathbf{G}_2$-orbits for pairs of octonions". In: *J. Pure Appl. Algebra* 229 (2025), Paper No. 107875 (14 pages). DOI: 10.1016/j.jpaa.2025.107875.

[44] S. Malev. "The images of non-commutative polynomials evaluated on 2×2 matrices over an arbitrary field". In: *J. Algebra Appl.* 13 (2014), Paper no. 1450004 (12 pages). DOI: 10.1142/S0219498814500042.

[45] S. Malev. "The images of noncommutative polynomials evaluated on the quaternion algebra". In: *J. Algebra Appl.* 20 (2021), Paper No. 2150074 (8 pages). DOI: 10.1142/S0219498821500742.

[46] S. Panja and S. Prasad. "The image of polynomials and Waring type problems on upper triangular matrix algebras". In: *J. Algebra* 631 (2023), pp. 148–193. DOI: 10.1016/j.jalgebra.2023.04.027.

[47] S. Panja, P. Saini, and A. Singh. "Images of polynomial maps with constants". In: *Mathematika* 71 (2025), Paper No. e70031 (17 pages). DOI: 10.1112/mtk.70031.

[48] S. Panja, P. Saini, and A. Singh. "Polynomial maps with constants on split octonion algebras". In: *Comm. Algebra* 54 (2026), pp. 579–594. DOI: 10.1080/00927872.2025.2531559.

[49] S. Panja, P. Saini, and A. Singh. "Surjectivity of polynomial maps on matrices". In: *Eur. J. Math.* 11 (2025), Paper No. 62 (22 pages). DOI: 10.1007/s40879-025-00853-6.

[50] R. Parimala and V. Suresh. "Isotropy of quadratic forms over function fields of p-adic curves". In: *Publications Mathématiques de l'IHÉS* 88 (1998), pp. 129–150. DOI: 10.1007/s10240-006-0004-4.

[51] S. Pumplün. "Sums of squares in octonion algebras". In: *Proc. Amer. Math. Soc.* 133 (2005), pp. 3143–3152. DOI: 10.1090/S0002-9939-05-07917-7.

[52] D. R. Richman. "The Waring problem for matrices". In: *Linear and Multilinear Algebra* 22 (1987), pp. 171–192. DOI: 10.1080/03081088708817831.

[53] D. R. Richman. "Matrices as sums of squares: a conjecture of Griffin and Krusemeyer". In: *Linear and Multilinear Algebra* 17 (1985), pp. 289–294. DOI: 10.1080/03081088508817660.

[54] L. Rodman. *"Topics in quaternion linear algebra"*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2014, pp. xii+363. DOI: 10.1515/9781400852741.

[55] W. Schmidt. *"Equations over finite fields: an elementary approach"*. Second. MR2121285. Kendrick Press, Heber City, UT, 2004, pp. x+333.

[56] J. Schneider and A. Thom. "Word maps with constants on symmetric groups". In: *Math. Nachr.* 297 (2024), pp. 165–173. DOI: 10.1002/mana.202300152.

[57] V. K. Shoda. "Einige Sätze über Matrizen". In: *Japanese journal of mathematics : transactions and abstracts* 13 (1936), pp. 361–365. DOI: 10.4099/jjm1924.13.0_361.

[58] T. A. Springer and F. D. Veldkamp. *"Octonions, Jordan algebras and exceptional groups"*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000, pp. viii+208. DOI: 10.1007/978-3-662-12622-6.

[59] L. N. Vaserstein. "On the sum of powers of matrices". In: *Linear and Multilinear Algebra* 21 (1987), pp. 261–270. DOI: 10.1080/03081088708817800.

[60] S. M. Victor. "Counting matrices that are squares". In: *arXiv:1606.09299* (2016).

[61] J. F. Voloch. "Diagonal equations over function fields". In: *Bol. Soc. Brasil. Mat.* 16 (1985), pp. 29–39. DOI: 10.1007/BF02584799.

[62] Y. Wang, J. Zhou, and Y. Luo. "The image of polynomials on 2×2 upper triangular matrix algebras". In: *Linear Algebra Appl.* 610 (2021), pp. 560–570. DOI: 10.1016/j.laa.2020.10.009.

[63] E. Waring. *"Meditationes Algebraicae"*. Latin. Cambridge University Press, 1770.

[64] N. A. Wiegmann. "Some theorems on matrices with real quaternion elements". In: *Canadian J. Math.* 7 (1955), pp. 191–201. DOI: 10.4153/CJM-1955-024-x.

Index