DERIVATIONS AND AUTOMORPHISMS OF SOME INFINITE MATRIX ALGEBRAS

BAZI SONSUZ MATRİS CEBİRLERİNİN TÜREVLERİ VE OTOMORFİZMALARI

TUNAHAN YILMAZ

PROF. DR. FERİDE KUZUCUOĞLU
Supervisor

Submitted to
Graduate School of Science and Engineering of Hacettepe University
as a Partial Fulfillment to the Requirements
for the Award of the Degree of Master of Science
in Mathematics

2021

ABSTRACT

DERIVATIONS AND AUTOMORPHISMS OF SOME INFINITE MATRIX ALGEBRAS

Tunahan YILMAZ
Master of Science, Department of Mathematics
Supervisor: Prof. Dr. Feride KUZUCUOĞLU
(June) 2021, 75 pages

Let R be a commutative ring with identity and $\mathcal{M}_n(R)$ be the algebra (ring) of all $n \times n$ matrices over R. Note that an additive map $\mathfrak{D}$ of a ring R into itself is said to be a derivation of R if $\mathfrak{D}(xy) = \mathfrak{D}(x)y + x\mathfrak{D}(y)$ for all $x, y \in R$. Studies on automorphisms and derivations of matrix algebras and matrix rings have been actively continuing since the 1950s. In the first study on this subject, in the case of R being a field, Skolem-Noether showed that each automorphism of the matrix algebra $\mathcal{M}_n(R)$ is an inner automorphism ([10, Theorem 4.3.1]). It has also been shown that every derivation of $\mathcal{M}_n(R)$ is inner in the case when R is a field (see [10, Proposition(p100)]). Later on, these studies were extended to the subalgebras (subrings) of the matrix algebra (ring) $\mathcal{M}_n(R)$.

Since the 2000s, studies on Lie and Jordan automorphisms and Lie and Jordan derivations of $\mathcal{M}_n(R)$ matrix algebras (rings) and subalgebras (subrings) have been started to appear in the literature. This thesis aims to bring automorphism and derivation problems to infinite matrix algebras and rings. The first chapter of this thesis, which consists of five chapters, contains the historical development of the subject of this thesis and relevant information. Second chapter covers some basic definitions and theorems which will help us better understand the work to be done in the following chapters. In the third chapter, infinite matrix algebras and rings are introduced, and some of their basic properties are observed. In the fourth chapter, derivations of column finite matrix rings are discussed. In the last chapter, all Lie derivations of (upper) niltriangular infinite matrix algebras are described.

Keywords: Automorphism, Derivation, Lie derivation, Infinite matrix algebras, Ring of column-finite matrices, Niltriangular matrix.

ÖZET

BAZI SONSUZ MATRİS CEBİRLERİNİN TÜREVLERİ VE OTOMORFİZMALARI

Tunahan YILMAZ
Yüksek Lisans, Matematik Bölümü
Tez Danışmanı: Prof. Dr. Feride KUZUCUOĞLU
Haziran 2021, 75 sayfa

R birimli ve değişmeli bir halka ve $\mathcal{M}_n(R)$ kümesi R üzerinde tanımlı $n \times n$ tipindeki bütün matrislerin oluşturduğu cebir (halka) olsun. $\mathfrak{D}$ dönüşümü, R halkası üzerinde tanımlı toplamsal bir dönüşüm olmak üzere eğer her $x, y \in R$ için $\mathfrak{D}(xy) = \mathfrak{D}(x)y + x\mathfrak{D}(y)$ koşulunu sağlıyorsa bu dönüşüme R halkasının bir türev dönüşümü denir. Matris cebirleri ve matris halkalarının otomorfizmaları ve türevleri üzerine çalışmalar 1950 yıllarından beri aktif olarak devam etmektedir. Bu konudaki ilk çalışmada, R yerine F cismi alındığında $\mathcal{M}_n(F)$ matris cebirinin her otomorfizmasının bir iç otomorfizma olduğu Skolem-Noether tarafından gösterilmiştir ([10, Theorem 4.3.1]). Bununla birlikte, yine R yerine F cismi alındığında, $\mathcal{M}_n(F)$ matris cebirinin her türev dönüşümünün ise bir iç türev dönüşümü olduğu kanıtlanmıştır (bkz. [10, Proposition(p100)]). Daha sonra bu çalışmalar $\mathcal{M}_n(R)$ matris cebirinin (halkasının) alt cebirlerine (halkalarına) taşınmıştır.

2000'li yıllardan itibaren ise $\mathcal{M}_n(R)$ matris cebirlerinin (halkalarının) ve alt cebirlerinin (alt halkalarının) Lie ve Jordan otomorfizmaları ile Lie ve Jordan türevleri üzerine çalışmalar literatürde yer almaya başlamıştır. Bu tezin amacı otomorfizma ve türev problemlerini sonsuz matris cebirlerine ve halkalarına taşımaktır. Beş bölümden oluşan bu tezin ilk bölümü, tez konusun tarihsel gelişimi ve ilgili bilgilerden oluşmaktadır. İkinci bölüm sonraki bölümlerde yapılacak çalışmaların daha iyi anlaşılmasında yardımcı olacak bazı temel tanım ve teoremleri içermektedir. Üçüncü bölümde sonsuz matris cebirleri ve halkaları tanıtılmış ve bunların bazı temel özellikleri incelenmiştir. Dördüncü bölümde sütun sonlu matris halkalarının türevleri ele alınmıştır. Son bölümde ise (üst) nilüçgensel sonsuz matris cebirlerinin Lie türevlerinin karakterizasyonu verilmiştir.

Anahtar Kelimeler: Otomorfizma, Türev, Lie türev, Sonsuz matris cebirleri, Sütun sonlu matris halkası, Nilüçgensel matris.

ACKNOWLEDGEMENT

Throughout the writing of this thesis I have received a great deal of support and assistance. I would like to express my sincere appreciation to all of those people who have made important contributions to this study.

First and foremost, I would like to express my deep and sincere gratitude to my supervisor, Prof. Dr. Feride Kuzucuoğlu, for her patience, motivation, help, and understanding. This thesis would not have been possible without her insightful comments and suggestions.

Besides my supervisor, I would like to thank the rest of my thesis committee: Prof. Dr. Ali Erdoğan, Prof. Dr. Evrim Akalan, Assoc. Prof. Dr. Pınar Aydoğdu, and Assoc. Prof. Dr. Ebru Solak, for their helpful comments and precious time.

I also owe a special debt to my family. I am extremely grateful to my parents, Sakine Yılmaz and İlhan Yılmaz, for their sacrifices, unflagging support, and love at every stage of my life. They are always there for me. In addition, I would like to thank my dear sister Özlem Yılmaz for the joy and motivation she brought to my life. I also have my brother Mehmet Ali Yılmaz to thank for being such a nice friend.

I am very much thankful to my dear fiancee Melis Berçin Koyuncu for her love, encouragement, understanding, and patience. I am really grateful to her for spending most of her time with me editing and preparing this thesis. I could not have completed this thesis without her continuing support.

Last but not least, I would like to thank my dear friends, Turgay Tırnavalı and Zafer Çetin, for the sleepless nights we were studying together, and for the happy leisure time activities to rest my mind.

Tunahan YILMAZ

June 2021, Ankara

Contents

LIST OF SYMBOLS

Symbols	
F	A field
R	A commutative ring with identity
$\mathcal{A}$	An F-algebra
$\mathcal{L}$	A Lie algebra over R
$\mathcal{M}_n(R)$	The ring of all $n \times n$ matrices over R
$\mathcal{N}T_n(R)$	The set of all strictly upper triangular $n \times n$ matrices over R
$\mathcal{M}_\infty(R)$	The set of all infinite matrices whose rows and columns are indexed by $\mathbb{Z}^+ \times \mathbb{Z}^+$ over R
$\mathcal{D}_\infty(R)$	The ring of infinite diagonal matrices over R
$\mathcal{T}_\infty(R)$	The ring of all infinite upper triangular matrices over R
$\mathcal{M}_{Cf}(R)$	The ring of all infinite column-finite matrices over R
$\mathcal{M}_{Rf}(R)$	The ring of all infinite row-finite matrices over R
$\mathcal{M}_{RCf}(R)$	$\mathcal{M}_{Rf}(R) \cap \mathcal{M}_{Cf}(R)$
$\mathcal{M}_\infty^{fin}(R)$	The ring of finitely supported matrices over R
ξ_{ij}	The matrix whose only nonzero entry is 1 in the (i, j) coordinate, regardless of its dimension
O	The zero matrix, regardless of its dimension
I_∞ and I_n	The identity matrices, infinite and $n \times n$, respectively
$\mathcal{N}T_\infty(R)$	The Lie algebra of infinite niltriangular matrices over R
$\operatorname{Im} \phi$	The image of a Lie algebra homomorphism ϕ
$\ker \phi$	The kernel of a Lie algebra homomorphism ϕ
$\mathbb{Z}^+$	The set of positive integers
$\mathbb{C}$	The set of complex numbers

Chapter 1

INTRODUCTION

In this chapter, as we mentioned in the abstract part, we focus on the historical development of the issues we will cover in this thesis. There is no doubt that knowing what has been done in the past will help us grasp the importance of our work. In this sense, this chapter will guide us throughout our thesis. By the way, here we introduce a few necessary definitions and concepts to understand what is being discussed. (Of course, the definitions and theorems required in the next chapters are discussed more broadly in the following chapter.)

Definition 1.0.1. For any unit element x of a ring R, the mapping

$$\begin{array}{rccc} \alpha_x : & R & \to & R \\ & y & \to & xyx^{-1} \end{array}$$

defines an automorphism. It will be called the inner automorphism of R.

Definition 1.0.2. Let x be an arbitrary element of a ring R, then the additive map

$$\begin{array}{rccc} \mathfrak{D}_x : & R & \to & R \\ & y & \to & xy - yx \end{array}$$

turns out to be a derivation, called inner.

Before going into details, let us make a little briefing about the notations. Unless specifically stated otherwise, from here on the notations R and F denote a commutative ring with identity and a field, respectively. We are now ready to get down to the heart of the matter. In 1950, a study that inspired many mathematicians was carried out. It was shown that each automorphism of the matrix algebra $\mathcal{M}_n(F)$ of all $n \times n$ matrices with coefficients in F is inner automorphism (see [10]). Of course, this study made many researchers wonder if a similar result can be extended to some subalgebras of $\mathcal{M}_n(F)$ too. Dubisch and Perlis were one of them. What they did in 1951 was describe the automorphisms of the subalgebra $\mathcal{N}T_n(F)$ of all $n \times n$ matrices with zero on and below the main diagonal (see [8]). Also, in 1987, S.Jondrup ([13]) determined the automorphisms of the ring of upper triangular matrices with entries in F. Now, before making the following definitions, we would like to mention some studies on the ring $\mathcal{N}T_n(R)$ of all $n \times n$ (upper) niltriangular matrices over R, where R is an associative ring with identity. In 1983, Levchuk showed that every automorphism of the ring $\mathcal{N}T_n(R)$ is equal to the product of diagonal, inner and central automorphisms (see [16]). And besides, in 2006, Chun and Park proved that every derivation of $\mathcal{N}T_n(R)$ is a sum of a certain diagonal, ring and a strongly nilpotent derivation (see [6]).

Definition 1.0.3. An R-linear map φ of a Lie algebra $\mathcal{L}$ satisfying

$$\varphi([\kappa,\rho]) = [\varphi(\kappa),\rho] + [\kappa,\varphi(\rho)] \text{ for all } \kappa,\rho \in \mathcal{L}$$

is called a Lie derivation of $\mathcal{L}$.

Definition 1.0.4. An additive map d of a ring R is called a Jordan derivation if it satisfies

$$d(r \circ s) = d(rs + sr) = d(r)s + sd(r) + rd(s) + d(s)r$$

for arbitrary elements $r, s \in R$, where $r \circ s = rs + sr$.

Of course, the problem of classifying Lie and Jordan derivations (automorphisms) of matrix algebras and their subalgebras has also attracted the attention of many people. S.Ou, D.Wang and R.Yao described all Lie derivations of $\mathcal{N}T_n(R)$ in 2007 (see [17]). Jordan derivations of the ring $\mathcal{N}T_n(R)$ were also characterized by F.Kuzucuoğlu in 2011 (see [15]).

Now, let us consider the set $\mathcal{M}_\infty(F)$ of all square matrices over F indexed by $\mathbb{Z}^+ \times \mathbb{Z}^+$. That is,

$$\mathcal{M}_\infty(F) = \left\{ X = (x_{ij})_{i,j\in\mathbb{Z}^+} \mid x_{ij} \in F \ and \ i,j \in \mathbb{Z}^+ \right\}.$$

This set forms a vector space with respect to usual matrix addition and scalar multiplication. In the light of the product of finite matrices, let's define

$$(XY)_{ik} = \sum_{j=1}^{\infty} x_{ij}y_{jk}$$

where $X,\ Y \in \mathcal{M}_\infty(F)$. Clearly, in order for the product matrix XY to exist, the set $\{j \mid x_{ij}y_{jk} \neq 0\}$ must be finite for all $i,k \in \mathbb{Z}^+$. By the way, even if the matrix multiplication is defined on this set, the associativity may not hold. This is exactly what we will discuss in the third chapter. Of course, in order to extend the studies on finite matrix algebra to the context of infinite matrices, we need some subsets of $\mathcal{M}_\infty(F)$ which has the structure of a ring (algebra). So, it is important for those working on this to know when infinite matrices are defined and obey associativity law.

By $\mathcal{M}_{Cf}(R)$, we denote the set of all $\mathbb{Z}^+ \times \mathbb{Z}^+$ matrices with a finite number of nonzero entries in each column. Such matrices will be called column finite. It should be noted that $\mathcal{M}_{Cf}(F)$ is an uncountable dimensional vector space and there is a one-to-one correspondence between column finite matrices and linear endomorphisms of the vector space $F^{\mathbb{Z}^+}$ with respect to canonical basis. Besides, what makes this set valuable is that it has a ring structure. To be more precise, we can talk about the derivations and automorphisms of $\mathcal{M}_{Cf}(R)$. For example, R.Slowik ([20]) characterized the derivations of the ring $\mathcal{M}_{Cf}(R)$ in 2015 in the case of R being an associative ring with 1. In the fourth section, we discuss this article in detail.

Finally, we write $\mathcal{T}_\infty(R)$ to denote the set of all $\mathbb{Z}^+ \times \mathbb{Z}^+$ upper triangular matrices over R. This set forms a Lie algebra with $[X, Y] = XY - YX$ where $X, Y \in \mathcal{T}_\infty(R)$. More interestingly, it has a subset worth studying. What W.Hołubowski, I.Kashuba and S.Zurek did in 2007 was determine all derivations of its Lie subalgebra $\mathcal{N}T_\infty(R)$ of all strictly upper triangular matrices (see [11]). In the last section, we concentrate on this article.

Chapter 2

PRELIMINARIES

In this chapter, basic definitions and related theorems required in later chapters will be discussed. Firstly, we focus on what a Lie algebra is.

2.1 Lie Algebras

We start off this section with definition of a ring in which the associativity need not be satisfied.

Definition 2.1.1. A Lie ring L is defined as a nonassociative ring such that its multiplication "." satisfies the following conditions

1. $x\,.\,x = 0$ for all $x \in L$ (anti-commutativity)

2. $(x\,.\,y)\,.\,z + (y\,.\,z)\,.\,x + (z\,.\,x)\,.\,y = 0$ for all $x, y, z \in L$ (the Jacobi identity).

Example 2.1.2. *The set $\mathcal{N}T_n(R)$ of all (upper) niltriangular $n \times n$ matrices over any associative ring R with identity forms a Lie ring if multiplication is defined as "$x\,.\,y = xy - yx$." This operation is called* ***commutation****. Actually, what is even more interesting is that we can extend this to all associative rings. In other words, any associative ring has the structure of a Lie ring under commutation.*

Of course, a Lie ring may fail to form a ring:

Example 2.1.3. *Consider the set $\mathfrak{P} = \left\{X^T = -X \mid X \in \mathcal{M}_n(F)\right\}$ of all $n \times n$ skew matrices over a field F. It can be easily shown that $\mathfrak{P}$ is a Lie ring under commutation; however, it is not a ring with usual matrix multiplication.*

Definition 2.1.4. Let F be a field. A set $\mathcal{A}$ $(\neq \varnothing)$ is said to be an F-algebra if

- There is an operation "+" such that $(\mathcal{A}, +)$ is an abelian group,
- There is a function
$$\begin{aligned} F \times \mathcal{A} &\to \mathcal{A} \\ (\lambda, x) &\to \lambda x \end{aligned}$$
(that is, there is a multiplication by scalar) such that $\mathcal{A}$ forms a vector space over F with respect to "+" and multiplication by scalars,
- There is a multiplication "."
$$\begin{aligned} . : \ \mathcal{A} \times \mathcal{A} &\to \mathcal{A} \\ (x, y) &\to xy \end{aligned}$$
which is F-bilinear, i.e.,
$$(x+y)z = xz + yz \ \text{ and } \ x(y+z) = xy + xz \ \text{ for all } \ x, y, z \in \mathcal{A},$$
$$\lambda(xy) = (\lambda x)y = x(\lambda y) \ \text{ for all } \ x, y \in \mathcal{A}, \lambda \in F.$$

Definition 2.1.5. Let $\mathcal{A}$ be an F-algebra. We call it **associative** if
$$(xy)z = x(yz) \ \text{ for all } \ x, y, z \in \mathcal{A}.$$
Also, $\mathcal{A}$ is said to be **unital** if there exists $1_{\mathcal{A}} \in \mathcal{A}$ such that $1_{\mathcal{A}}x = x = x1_{\mathcal{A}}$ for all $x \in \mathcal{A}$.

Example 2.1.6. *One can easily see that $\mathcal{M}_n(F)$ - the vector space of all $n \times n$ matrices over a field F - forms a* ***unital associative algebra*** *with respect to matrix multiplication. Obviously, the identity matrix is the identity element in this algebra.*

This much information about Lie rings and algebras is enough for us. Now, we are ready to talk about what Lie algebra is. You will soon see why we have given these algebra and Lie ring definitions beforehand.

Definition 2.1.7. Let F be a field. A F-vector space $\mathcal{L}$ together with a F-bilinear map
$$\begin{aligned} [-,-] : \ \mathcal{L} \times \mathcal{L} &\to \mathcal{L} \\ (\kappa, \rho) &\to [\kappa, \rho], \end{aligned}$$
which is called the Lie bracket or commutator is said to be **Lie algebra** if

- ($\mathcal{L}_1$) Anti-commutativity : $[\kappa, \kappa] = 0$ for all $\kappa \in \mathcal{L}$, and
- ($\mathcal{L}_2$) The Jacobi Identity : $[\kappa, [\rho, z]] + [\rho, [z, \kappa]] + [z, [\kappa, \rho]] = 0$ for all $\kappa, \rho, z \in \mathcal{L}$.

Using bilinearity of the Lie bracket $[-, -]$, one can see that

$$0 = [\kappa + \rho, \kappa + \rho] = [\kappa, \kappa] + [\kappa, \rho] + [\rho, \kappa] + [\rho, \rho] = [\kappa, \rho] + [\rho, \kappa].$$

Therefore, due to condition ($\mathcal{L}_1$), it must be

$$[\kappa, \rho] = -[\rho, \kappa] \ \text{ for all } \ \kappa, \rho \in \mathcal{L}.$$

In fact, we can interpret the definition of Lie algebra in two ways:

- A Lie algebra is nothing but than an algebra satisfying $\mathcal{L}_1$ and $\mathcal{L}_2$ with "$[\kappa, \rho] = \kappa\rho$."
- A Lie ring that also has the structure of an algebra is a Lie algebra.

Let's take a look at a couple of Lie algebras. In all of the examples below, F will denote an arbitrary field.

1. Let gl(V) be the vector space of all linear maps from V to V, where V is a finite-dimensional F-vector space. Then, it forms a Lie algebra if the Lie bracket defined by
$$[\kappa, \rho] = \kappa \circ \rho - \rho \circ \kappa \ \text{ for } \ \kappa, \rho \in gl(V).$$
We will call it **general linear algebra**. ("$\circ$" stands for the composition of maps.)

2. Now, we are going to introduce "matrix version" of the above example. $\mathcal{M}_n(F)$ also has the structure of a Lie algebra with the Lie bracket defined by
$$[\kappa, \rho] = \kappa\rho - \rho\kappa,$$
where $\kappa\rho$ is the usual matrix multiplication.

   I would like to talk about one more thing about this Lie algebra because it will guide us in the following chapters: Let ξ_{ij} be an $n \times n$ matrix which has a "1" in the ij-th entry and all other entries are 0. It will be called **matrix unit**. What is crucial here is that these matrix units form a basis for the vector space $\mathcal{M}_n(F)$. To put it another way, the set
$$\left\{\xi_{ij} \mid 1 \leq i, j \leq n\right\}$$

is a basis for $\mathcal{M}_n(F)$. It is also useful to know that

$$[\xi_{ij}, \xi_{kl}] = \delta_{jk}\xi_{il} - \delta_{il}\xi_{kj},$$

where δ is the Kronecker delta. This formula will make calculations easier in $\mathcal{M}_n(F)$.

3. Recall that a matrix $\kappa \in \mathcal{M}_n(F)$ is called upper triangular if $\kappa_{ij} = 0$ for all $i > j$. Write $\mathcal{T}_n(F)$ for the vector subspace of all upper triangular matrices in $\mathcal{M}_n(F)$. It also becomes a Lie algebra with the same bracket in $\mathcal{M}_n(F)$.

4. Likewise, the subspace of all strictly upper triangular $n \times n$ matrices is a Lie algebra if the Lie bracket defined as in $\mathcal{M}_n(F)$. (Recall that we call a matrix κ strictly upper triangular if $\kappa_{ij} = 0$ for all $i \geq j$.) This Lie algebra is shown by $\mathcal{N}T_n(F)$.

5. If $\mathcal{A}$ is an associative algebra over F, then A becomes a Lie algebra with

$$[x, y] = xy - yx \quad \text{for all} \quad x, y \in \mathcal{A}.$$

Actually, in the light of examples (3) and (4), one can realize that subset of a Lie algebra $\mathcal{L}$ may also form a Lie algebra. In the next section, we will focus on subalgebras and ideals of a Lie algebra.

2.2 Subalgebras and Ideals

Let $\mathcal{L}$ be a Lie algebra over a field F.

Definition 2.2.1. A vector subspace $\mathcal{W} \subset \mathcal{L}$ is called a Lie subalgebra of $\mathcal{L}$ if

$$[\kappa, \rho] \in \mathcal{W} \quad \text{for all} \quad \kappa, \rho \in \mathcal{W}.$$

Now, we shall give the definition of an ideal of a Lie algebra.

Definition 2.2.2. A vector subspace $\mathcal{U} \subset \mathcal{L}$ is called an ideal of $\mathcal{L}$ if

$$[\kappa, \rho] \in \mathcal{U} \quad \text{for all} \quad \kappa \in \mathcal{L}, \rho \in \mathcal{U}.$$

Any Lie algebra $\mathcal{L}$ has at least two ideals: the trivial ideal $\{0\}$ and $\mathcal{L}$ itself. Note that a subalgebra may fail to be an ideal; however, an ideal is always a subalgebra. One can easily show that $\mathcal{T}_n(F)$ constitute an example of a subalgebra of $\mathcal{M}_n(F)$ which is not an ideal. Now, consider the centre of $\mathcal{L}$ defined as follows:

$$\mathcal{Z}(\mathcal{L}) = \Big\{\kappa \in \mathcal{L} \mid [\kappa, \rho] = 0 \ \text{ for all } \ \rho \in \mathcal{L}\Big\}.$$

It is a quick check to see that $\mathcal{Z}(\mathcal{L})$ is an ideal of $\mathcal{L}$. What is important here is that $\mathcal{L} = \mathcal{Z}(\mathcal{L})$ if and only if $\mathcal{L}$ is abelian. We must keep this fact in mind. Also, note that finding the center of a Lie algebra $\mathcal{L}$ is not as easy as it seems; it may take time to determine what $\mathcal{Z}(\mathcal{L})$ is. In the Chapter 5, we will face such a case.

Let $\mathcal{U}$ and $\mathcal{V}$ be ideals of a Lie algebra $\mathcal{L}$. We will close out this section with discussing how to construct new ideals from $\mathcal{U}$ and $\mathcal{V}$. In fact, we can do this in many ways:

- Firstly, we shall observe that the subspace $\mathcal{U} \cap \mathcal{V}$ is an ideal of $\mathcal{L}$. Let $\kappa \in \mathcal{L}$ and $\rho \in \mathcal{U} \cap \mathcal{V}$. Then, it must be $[\kappa, \rho] \in \mathcal{U} \cap \mathcal{V}$ since $\mathcal{U}$ and $\mathcal{V}$ are ideals. This is why $\mathcal{U} \cap \mathcal{V}$ forms an ideal of $\mathcal{L}$.

- Now, consider the following set

  $$\mathcal{U} + \mathcal{V} = \Big\{\kappa + \rho \mid \kappa \in \mathcal{U}, \rho \in \mathcal{V}\Big\}.$$

  It is not hard to see that this set is an ideal of $\mathcal{L}$.

- The above examples suggest that we might define a product of ideals. Let's consider the subspace

  $$[\mathcal{U}, \mathcal{V}] = \text{Span}\Big\{[\kappa, \rho] \mid \kappa \in \mathcal{U}, \rho \in \mathcal{V}\Big\}.$$

  We shall show that $[\mathcal{U}, \mathcal{V}]$ is also an ideal of $\mathcal{L}$. Now, let $\kappa \in \mathcal{U}$, $\rho \in \mathcal{V}$, and $z \in \mathcal{L}$. Then, using the Jacobi identity we have

  $$[z, [\kappa, \rho]] = [\kappa, [z, \rho]] + [[z, \kappa], \rho].$$

  Of course, $[z, \rho] \in \mathcal{V}$ as $\mathcal{V}$ is an ideal. Hence, $[\kappa, [z, \rho]] \in [\mathcal{U}, \mathcal{V}]$. Likewise, one can directly see that $[[z, \kappa], \rho] \in [\mathcal{U}, \mathcal{V}]$. Accordingly, $[z, [\kappa, \rho]] \in [\mathcal{U}, \mathcal{V}]$. Now, all we will do is check whether or not $[z, x] \in [\mathcal{U}, \mathcal{V}]$ in the case when $z \in \mathcal{L}$ and $x \in [\mathcal{U}, \mathcal{V}]$: By construction of $[\mathcal{U}, \mathcal{V}]$, it must be $x = \sum \lambda_i [\kappa_i, \rho_i]$, where λ_i are scalars and $\kappa_i \in I$ and $\rho_i \in J$. Since the Lie bracket is bilinear,

  $$[z, x] = [z, \sum \lambda_i [\kappa_i, \rho_i]] = \sum \lambda_i [z, [\kappa_i, \rho_i]].$$

As discussed above, each $[z, [\kappa_i, \rho_i] \in [\mathcal{U}, \mathcal{V}]$. Hence, $[z, x] \in [\mathcal{U}, \mathcal{V}]$. The result follows.

Maps that are structure preserving help us understand the structure of a mathematical object such as a vector space, group, or ring. For example, linear maps are helpful to understand how different vector spaces related to each other. In the following sections, our purpose is to extend this view to Lie algebras.

2.3 Homomorphisms

Let $\mathcal{L}$, $\mathcal{L}_1$ and $\mathcal{L}_2$ be Lie algebras over a field F.

Definition 2.3.1. A linear map $\phi : \mathcal{L}_1 \to \mathcal{L}_2$ is called homomorphism if

$$\phi([\kappa, \rho]) = [\phi(\kappa), \phi(\rho)] \text{ for all } \kappa, \rho \in \mathcal{L}_1.$$

Definition 2.3.2. We say that a homomorphism $\phi : \mathcal{L} \to \mathcal{L}$ is an automorphism if ϕ is bijective.

Now, we are going to introduce a significantly important homomorphism: Let $\kappa \in \mathcal{L}$. Consider the following map

$$\begin{array}{rccc} \text{ad}: & \mathcal{L} & \to & gl(\mathcal{L}) \\ & \kappa & \to & \text{ad}\kappa \end{array}$$

defined by $\text{ad}\kappa(\rho) = [\kappa, \rho]$ for all $\rho \in \mathcal{L}$.

The bilinearity of the commutator implies that $\kappa \to \text{ad}\kappa$ is linear. It can be easily shown that ad is a homomorphism between $\mathcal{L}$ and gl($\mathcal{L}$). We call it **adjoint homomorphism**. It is important to note that the kernel of ad is the centre of $\mathcal{L}$.

2.4 Derivations

Let $\mathcal{A}$ be an F-algebra, where F is a field. An F-linear map $\mathfrak{D} : \mathcal{A} \to \mathcal{A}$ is called a **derivation** of $\mathcal{A}$ if

$$\mathfrak{D}(xy) = \mathfrak{D}(x)y + x\mathfrak{D}(y) \text{ for all } x, y \in \mathcal{A}.$$

The adjoint homomorphism discussed above provides a "nice" example of a derivation. Let $\mathcal{L}$ be a Lie algebra and $\kappa \in \mathcal{L}$. It follows from the Jacobi identity that the map $\text{ad}\kappa : \mathcal{L} \to \mathcal{L}$

satisfies the necessary condition to be a derivation:

$$\mathrm{ad}\kappa([\rho, z]) = [\kappa, [\rho, z]] = [[\kappa, \rho], z] + [\rho, [\kappa, z]] = [\mathrm{ad}\kappa(\rho), z] + [\rho, \mathrm{ad}\kappa(z)]$$

for all $\rho, z \in \mathcal{L}$. We call such a derivation **inner derivation**.

Now, let's write $Der\mathcal{A}$ for the set of derivations of $\mathcal{A}$. This set is a vector subspace of $\mathrm{g}l(\mathcal{A})$ since it is closed under addition and scalar multiplication and contains the zero map. As you guessed, there is another reason why we are talking about this set besides being a vector subspace of $\mathrm{g}l(\mathcal{A})$. $Der\mathcal{A}$ also has the structure of a Lie subalgebra of $\mathrm{g}l(\mathcal{A})$. (One can deduce that if $\mathfrak{D}$ and $\mathfrak{C}$ are derivations of an algebra, then $[\mathfrak{D}, \mathfrak{C}] = \mathfrak{D}\mathfrak{C} - \mathfrak{C}\mathfrak{D}$ is a derivation as well.)

We have just seen that Lie algebras are nothing more than algebras satisfying ($\mathcal{L}_1$) and ($\mathcal{L}_2$). Now, we shall take this critical result as a definition:

Definition 2.4.1. An F-linear map φ of a Lie algebra $\mathcal{L}$ satisfying

$$\varphi([\kappa, \rho]) = [\varphi(\kappa), \rho] + [\kappa, \varphi(\rho)] \quad \text{for all} \quad \kappa, \rho \in \mathcal{L}$$

is called a **lie derivation** of $\mathcal{L}$.

2.5 Quotient Algebras and Isomorphism Theorems

Let $\mathcal{U}$ be an ideal of the Lie algebra $\mathcal{L}$. Of course, $\mathcal{U}$ has the structure of a vector subspace of $\mathcal{L}$, and so one may consider the cosets $z + \mathcal{U} = \left\{z + \rho \mid \rho \in \mathcal{U}\right\}$ where $z \in \mathcal{L}$ and the quotient vector subspace

$$\mathcal{L}/\mathcal{U} = \left\{z + \mathcal{U} \mid z \in \mathcal{L}\right\}.$$

Now, we are going to see that this quotient vector subspace also forms a Lie algebra. In the light of our "undergraduate abstract algebra" knowledge, we may define a Lie bracket on $\mathcal{L}/\mathcal{U}$ as follows

$$[\kappa + \mathcal{U}, z + \mathcal{U}] = [\kappa, z] + \mathcal{U} \ \text{ for } \ \kappa, z \in \mathcal{L}.$$

Of course, the bracket on the right-hand side is the Lie bracket defined on $\mathcal{L}$. Now, we check if the bracket on $\mathcal{L}/\mathcal{U}$ is well-defined. Let $\kappa + \mathcal{U}, z + \mathcal{U}, \kappa' + \mathcal{U}, z' + \mathcal{U} \in \mathcal{L}/\mathcal{U}$ and suppose $\kappa + \mathcal{U} = \kappa' + \mathcal{U}$ and $z + \mathcal{U} = z' + \mathcal{U}$. Then $\kappa - \kappa' \in \mathcal{U}$ and $z - z' \in \mathcal{U}$. All we have to do is show that $[\kappa', z'] + \mathcal{U} = [\kappa, z] + \mathcal{U}$. It follows from the bilinearity of the Lie bracket in $\mathcal{L}$ that

$$\begin{aligned}[\kappa', z'] &= [\kappa + (\kappa - \kappa'), z' + (z - z')] \\ &= [\kappa, z] + [\kappa - \kappa', z'] + [\kappa', z - z'] + [\kappa - \kappa', z - z'].\end{aligned}$$

Obviously, the final three summands must belong to $\mathcal{U}$ since $\mathcal{U}$ is an ideal. Thus, it must be

$$[\kappa' + \mathcal{U}, z' + \mathcal{U}] = [\rho, z] + \mathcal{U}.$$

This is exactly what we wanted to show. Also, one can check that the Lie bracket on $\mathcal{L}/\mathcal{U}$ is bilinear and satisfies the conditions $(\mathcal{L}_1)$ and $(\mathcal{L}_2)$. That is, $\mathcal{L}/\mathcal{U}$ has the structure of a Lie algebra with this bracket.

As expected, we have isomorphism theorems for Lie algebras as in vector spaces and groups. Let us state them:

Theorem 2.5.1. ***(Isomorphism Theorems)*** *Let $\mathcal{L}, \mathcal{L}_1$ and $\mathcal{L}_2$ be Lie algebras over a field F.*

1. ***(First Isomorphism Theorem)*** *If $\phi : \mathcal{L}_1 \to \mathcal{L}_2$ is a homomorphism between $\mathcal{L}_1$ and $\mathcal{L}_2$, then kerϕ and imϕ are ideals of $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. Also,*

$$\mathcal{L}_1 \big/ ker\phi \cong im\phi.$$

2. ***(Second Isomorphism Theorem)*** *Let $\mathcal{U}$ and $\mathcal{V}$ be ideals of our Lie algebra $\mathcal{L}$. Then*

$$(\mathcal{U} + \mathcal{V}) / \mathcal{V} \cong \mathcal{U} / \mathcal{U} \cap \mathcal{V}.$$

3. ***(Third Isomorphism Theorem)*** *If $\mathcal{U}$ and $\mathcal{V}$ are ideals of $\mathcal{L}$ such that $\mathcal{U} \subseteq \mathcal{V}$, then $\mathcal{V}/\mathcal{U}$ is an ideal of $\mathcal{L}/\mathcal{U}$ and*

$$\mathcal{L}/\mathcal{U} \Big/ \mathcal{V}/\mathcal{U} \cong \mathcal{L}/\mathcal{V}$$

Before moving onto the next section let's do a couple of examples:

Example 2.5.2. *The trace of a square matrix κ, denoted tr(κ), is defined to be the sum of its diagonal elements. Let's consider the following linear map*

$$\begin{array}{rccc} \phi : & \mathcal{M}_n(F) & \to & F \\ & \kappa & \to & tr(\kappa) \end{array}$$

, where F is a field. Now

$$\begin{aligned} \phi([\kappa, \rho]) &= \phi(\kappa\rho - \rho\kappa) = tr(\kappa\rho - \rho\kappa) \\ &= tr(\kappa\rho) - tr(\rho\kappa) \\ &= [tr(\kappa), tr(\rho)] \\ &= [\phi(\kappa), \phi(\rho)] = 0 \end{aligned}$$

for all $\kappa, \rho \in \mathcal{M}_n(F)$, showing that ϕ is a Lie algebra homomorphism. Clearly, ϕ is surjective and its kernel is $\mathcal{S}_n(F)$, which is the Lie subalgebra of matrices whose traces are zero. Therefore, it follows from the first isomorphism theorem that

$$\mathcal{M}_n(F) \Big/ \mathcal{S}_n(F) \cong F.$$

Example 2.5.3. *Let $\mathcal{L}$ be a Lie algebra over a field F. The set of all inner derivations of $\mathcal{L}$ is a Lie subalgebra of Der($\mathcal{L}$). We will denote it by $IDer(\mathcal{L})$. Now, let's define a map*

$$\begin{aligned} \phi: \quad L &\rightarrow \quad IDer(L) \\ \kappa &\rightarrow \quad ad\kappa \end{aligned}$$

Clearly, ϕ preserves addition and scalar multiplication. Now

$$\begin{aligned} \phi([\kappa, \rho])(z) &= [[\kappa, \rho], z] \\ &= [\kappa, [\rho, z]] - [\rho, [\kappa, z] \quad \textit{(by Jacobi identity)} \\ &= ([\phi(\kappa), \phi(\rho)])(z) \end{aligned}$$

for all $\kappa, \rho, z \in \mathcal{L}$, proving that ϕ is a Lie algebra homomorphism. It is also easy to see that ϕ is a surjective map and

$$\begin{aligned} Ker\phi &= \left\{ \kappa \in \mathcal{L} \mid \phi(\kappa) = ad\kappa = 0 \right\} \\ &= \left\{ \kappa \in \mathcal{L} \mid [\kappa, \rho] = 0 \textit{ for all } \rho \in \mathcal{L} \right\} \\ &= \left\{ \kappa \in \mathcal{L} \mid \kappa \in \mathcal{Z}(\mathcal{L}) \right\} = \mathcal{Z}(\mathcal{L}). \end{aligned}$$

Consequently, by the first isomorphism theorem

$$IDer(\mathcal{L}) \cong \mathcal{L} \Big/ \mathcal{Z}(\mathcal{L}).$$

Remark 2.5.4. As it happens, all mathematical objects and concepts which we dealt with so far were constructed over a field F. However, it should be kept in mind that we can carry all these works to a commutative ring with slight differences. For example, an algebra $\mathcal{A}$ over a commutative ring R is nothing more than an R-module with an R-bilinear multiplication map "."

$$\begin{aligned} .: \quad \mathcal{A} \times \mathcal{A} &\rightarrow \quad \mathcal{A} \\ (x, y) &\rightarrow \quad xy. \end{aligned}$$

In the last chapter we will describe all derivations of a Lie algebra over a commutative ring with identity.

We have already discussed that $\mathcal{M}_n(F)$-the vector space of all $n \times n$ matrices over a field F- becomes an unital associative F-algebra with respect to usual matrix multiplication. As we all know, matrices play a significant role in algebra. Therefore, the need to classify all derivations and automorphisms of this algebra has become inevitable. This is precisely what we will talk about in the following section.

2.6 Automorphisms and Derivations of The Matrix Algebra $\mathcal{M}_n(F)$

Throughout this section, as in others, F will be a field. Let's start off this section with the definition below:

Definition 2.6.1. A finite dimensional associative F-algebra $\mathcal{A}$ is said to be a **central simple algebra** if

- $\mathcal{A}$ is simple,
- Center of $\mathcal{A}$ is precisely F.

Example 2.6.2. *The matrix algebra $\mathcal{M}_n(F)$ provides a good example of a central simple algebra.*

In this section, our goal is to characterize automorphisms and derivations of $\mathcal{M}_n(F)$. To do this, we first state the Skolem-Noether theorem, which characterizes the automorphisms of simple rings and guides those studying central simple algebras.

Theorem 2.6.3. *[10, Theorem 4.3.1](Skolem-Noether) Every automorphism of a finite dimensional central simple algebra is inner.*

It immediately follows from the Skolem-Noether theorem that every automorphism of the algebra $\mathcal{M}_n(F)$ is inner. It takes a long time to give proof of the Noether-Skolem theorem to deduce this fact. Instead, we present the simplest proof of this very crucial fact.

Theorem 2.6.4. *[19, Theorem 1.1] Suppose that ϕ is a bijective linear map of $\mathcal{M}_n(F)$ such that $\phi(XY) = \phi(X)\phi(Y)$ for all $X, Y \in \mathcal{M}_n(F)$. Then there must be an invertible matrix $P \in \mathcal{M}_n(F)$ with $\phi(X) = PXP^{-1}$ for every $X \in \mathcal{M}_n(F)$.*

Proof. Let $\phi : \mathcal{M}_n(F) \to \mathcal{M}_n(F)$ be a linear map satisfying above hypotheses. Take two column vectors which are different from zero $\upsilon, \omega \in F^n$. Then, of course, one can find an element $z \in F^n$ such that $\phi(v\omega^T)z \neq 0$. Now, let's construct a map $P : F^n \to F^n$ which maps x to $\phi(x\omega^t)z$. Obviously, the linearity of ϕ implies P is a linear map. Also, P is nonzero as Pv is nonzero. Now observe that

$$\begin{aligned} PXy &= \phi(Xy\omega^T)z \\ &= \phi(X)\phi(y\omega^T)z \\ &= \phi(X)Py \end{aligned}$$

for any choice of $y \in F^n$ and $X \in \mathcal{M}_n(F)$, giving us that $PX = \phi(X)P$. As you guessed, the last job we will do is to see that the P is invertible. Now, let $u \in F^n$. Using the surjectivity of ϕ one can find $Y \in \mathcal{M}_n(F)$ such that $\phi(Y)Pv = u = PYv$. This implies that P is surjective, and so invertible. The result follows. $\square$

Theorem 2.6.5. *[10, Proposition(p100)] Let $\mathcal{A}$ be a simple algebra finite dimensional over its center F. Then any derivation of $\mathcal{A}$ is inner.*

Proof. Let $\mathfrak{D}$ be a derivation of $\mathcal{A}$. First, consider $\mathcal{A}_2$ which is the ring of all 2×2 matrices over $\mathcal{A}$. Of course, $\mathcal{A}_2$ also has the structure of a central simple algebra. Now, put

$$B = \left\{ \begin{bmatrix} x & \mathfrak{D}(x) \\ 0 & x \end{bmatrix} : x \in \mathcal{A} \right\} \text{ and } C = \left\{ \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} : x \in \mathcal{A} \right\}.$$

(Clearly, $\mathfrak{D}(\lambda) = 0$ for $\lambda \in F$.) Then, the mapping $\phi : C \to B$ defined by

$$\phi \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} = \begin{bmatrix} x & \mathfrak{D}(x) \\ 0 & x \end{bmatrix}$$

is an isomorphism between C and B leaving F elementwise fixed. Also $C \approx \mathcal{A}$. From the Noether-Skolem Theorem there must be an invertible matrix

$$\begin{bmatrix} u & y \\ z & v \end{bmatrix} \in \mathcal{A}_2$$

such that

$$\begin{bmatrix} x & \mathfrak{D}(x) \\ 0 & x \end{bmatrix} \begin{bmatrix} u & y \\ z & v \end{bmatrix} = \begin{bmatrix} u & y \\ z & v \end{bmatrix} \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}.$$

Therefore:

$$\begin{aligned} xu + \mathfrak{D}(x)z &= ux \\ xy + \mathfrak{D}(x)v &= yx \\ xz &= zx \\ xv &= vx, \end{aligned}$$

for all $x \in \mathcal{A}$. Clearly, it must be $v, z \in F$. Besides, invertibility of $\begin{bmatrix} u & y \\ z & v \end{bmatrix}$ forces one of these scalars, say z, to be nonzero. Now, put $t = uz^{-1}$. Then $\mathfrak{D}(x) = tx - xt$ for all $x \in \mathcal{A}$. Consequently, $\mathfrak{D}$ is inner. □

As a straight consequence of above theorem, we have:

Corollary 2.6.6. *Let F be a field. Then any derivation of the matrix algebra $\mathcal{M}_n(F)$ is inner.*

We have discussed how to define product of ideals $\mathcal{U}, \mathcal{V}$ of a Lie algebra $\mathcal{L}$. Now, we make use of this construction in the next section by considering the ideals

$$[\mathcal{L}, \mathcal{L}], [\mathcal{L}, [\mathcal{L}, \mathcal{L}]], [\mathcal{L}, [\mathcal{L}, [\mathcal{L}, \mathcal{L}]]], ...$$

2.7 Nilpotent Lie Algebras

The lower central series (or descending central series) of a Lie algebra $\mathcal{L}$ is defined as:

$$\Omega_1 = \mathcal{L} \text{ and } \Omega_i = [\mathcal{L}, \Omega_{i-1}] \text{ for } i \geq 2.$$

Then "$\Omega_1 \supseteq \Omega_2 \supseteq \Omega_3 \supseteq ...$". It should be noted that Ω_i is an ideal of $\mathcal{L}$ (and not just an ideal of Ω_{i-1}) for each $i \in \mathbb{N}$ since the product of ideals is an ideal. The reason of why we call this series "lower central series" is that

$$\Omega_i \big/ \Omega_{i+1} \subseteq \mathcal{Z}\left(\Omega \big/ \Omega_{i+1}\right) \text{ for any } i \in \mathbb{N}.$$

Definition 2.7.1. Let $\mathcal{L}$ be a Lie algebra. We call it **nilpotent** if $\Omega_n = 0$ for some $n \geq 1$.

Now, we shall determine if the Lie algebra $\mathcal{N}_5(F)$ of strictly upper triangular 5×5 matrices over a field F is nilpotent. To do this, we shall observe the lower central series of $\mathcal{N}_5(F)$. This observation also helps us better understand the subject we will discuss in Chapter 5.

Let $X, Y \in \mathcal{N}_5(F)$. One can easily compute that

$$XY - YX = Z = \begin{pmatrix} 0 & 0 & z_{13} & z_{14} & z_{15} \\ 0 & 0 & 0 & z_{24} & z_{25} \\ 0 & 0 & 0 & 0 & z_{35} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since $\Omega_2 = [\mathcal{N}_5(F), \mathcal{N}_5(F)] = [\Omega_1, \Omega_1]$ is defined to be the linear span of elements of the form $[X, Y] = XY - YX$, where $X, Y \in \mathcal{N}(5, F)$, we have

$$\Omega_2 = [\Omega_1, \Omega_1] = \left\{ Z \in \mathcal{N}_5(F) \mid Z = \begin{pmatrix} 0 & 0 & z_{13} & z_{14} & z_{15} \\ 0 & 0 & 0 & z_{24} & z_{25} \\ 0 & 0 & 0 & 0 & z_{35} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \right\}.$$

In other words, $\Omega_2 \subset \mathcal{N}_5(F)$ consists of matrices whose entries one above the main diagonal are all zeros. Similarly, in the case when $X \in \Omega_2$ and $Y \in \mathcal{N}_5(F)$, we obtain

$$XY - YX = Z = \begin{pmatrix} 0 & 0 & 0 & z_{14} & z_{15} \\ 0 & 0 & 0 & 0 & z_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence :

$$\Omega_3 = [\Omega_2, \Omega_1] = \left\{ Z \in \mathcal{N}_5(F) \mid Z = \begin{pmatrix} 0 & 0 & 0 & z_{14} & z_{15} \\ 0 & 0 & 0 & 0 & z_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \right\} \subset \Omega_2.$$

Using similar arguments, one can see that

$$\Omega_4 = [\Omega_3, \Omega_1] = \left\{ Z \in \mathcal{N}_5(F) \mid Z = \begin{pmatrix} 0 & 0 & 0 & 0 & z_{15} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \right\} \subset \Omega_3$$

and

$$\Omega_5 = [\Omega_4, \Omega_1] = \{O\}.$$

Hence, the Lie algebra $\mathcal{N}_5(F)$ of strictly upper triangular 5×5 matrices over a field F is nilpotent.

Now we are going to propose another interesting feature of the sets Ω_i we observed above.

Lemma 2.7.2. *The sets Ω_i, $i \geq 1$ are invariant under any derivation φ of the Lie algebra $\mathcal{N}(5, F)$.*

Proof. As mentioned in the previous section, the sets Ω_i are ideals of $\mathcal{N}(5, F)$. Now, we will show that $\varphi(\Omega_i) \subseteq \Omega_i$ for any derivation φ of $\mathcal{N}(5, F)$.

Let φ be any derivation of $\mathcal{N}(5, F)$. By definition, $\varphi(\Omega_1) \subseteq \Omega_1$. It is also clear that the set Ω_i are invariant under φ for $i \geq 5$ as $\Omega_i = [\Omega_{i-1}, \Omega_1] = \{0\}$ for $i \geq 5$. Hence, all we have to do is determine whether the sets Ω_2, Ω_3 and Ω_4 are invariant under φ. Let $i = 2$. Then

$$\begin{aligned} \varphi(\Omega_2) = \varphi([\Omega_1, \Omega_1]) &= [\varphi(\Omega_1), \Omega_1] + [\Omega_1, \varphi(\Omega_1)] \\ &\subseteq \varphi(\Omega_1)\Omega_1 - \Omega_1\varphi(\Omega_1) + \Omega_1\varphi(\Omega_1) - \varphi(\Omega_1)\Omega_1. \\ &\subseteq \Omega_1\Omega_1 - \Omega_1\Omega_1 + \Omega_1\Omega_1 - \Omega_1\Omega_1 \end{aligned}$$

Of course, $\Omega_1\Omega_1 \subseteq \Omega_2$. Hence,

$$\varphi(\Omega_2) \subseteq \Omega_2 - \Omega_2 + \Omega_2 - \Omega_2 \subseteq \Omega_2.$$

Through a similar process, one can easily see that $\varphi(\Omega_3) \subseteq \Omega_3$ and $\varphi(\Omega_4) \subseteq \Omega_4$. Thereby, the result follows. $\square$

Chapter 3

Infinite Matrices

A matrix $X = (x_{ij})$ which has infinite number of columns or rows (or both) is called an infinite matrix. In this chapter, we will deal with infinite matrices whose rows and columns are indexed by $\mathbb{Z}^+ \times \mathbb{Z}^+$ with entries in the field F of characteristic zero. We define addition and scalar multiplication on infinite matrices as follows:

$$X + Y = (x_{ij} + y_{ij})_{i,j \in \mathbb{Z}^+} \text{ and } \lambda X = (\lambda x_{ij})_{i,j \in \mathbb{Z}^+}, \text{where } \lambda \in F.$$

We denote this vector space of infinite matrices by $\mathcal{M}_\infty(F)$. As expected, infinite matrices differs significantly from finite matrices in many ways. Let's take a look at a few reasons lead to this:

- As we all know, the concept of determinant has a pretty important place in the finite matrix theory and the role determinants play cannot be ignored; however, there is no corresponding phenomenon for infinite matrices.

- Consider the ring $\mathcal{M}_n(F)$ of all $n \times n$ matrices over a field F. Let $X = (x_{ij})_{1 \leq i,j \leq n}$ and $Y = (y_{jk})_{1 \leq j,k \leq n} \in \mathcal{M}_n(F)$. Recall that their product is defined as

  $$XY = \big(\sum_{j=1}^{n} x_{ij} y_{jk}\big)_{1 \leq i,k \leq n}$$

  However, when we move this definition to infinity case, we run into a big problem. To be more precise, the multiplication of two infinite matrices $X = (x_{ij})_{i,j \in \mathbb{Z}^+}$ and $Y = (y_{jk})_{j,k \in \mathbb{Z}^+}$

  $$XY = \big(\sum_{j=1}^{\infty} x_{ij} y_{jk}\big)_{i,k \in \mathbb{Z}^+}$$

  may not exist because the above sum may diverge for some values of $i, k \in \mathbb{Z}^+$.

- Moreover, one may expect a variety of fundamental theorems in the theory of finite matrices to be valid in the case of infinite matrices. It is possible that we may obtain corresponding theorems for infinite matrices, but this is very unlikely to happen due to convergence and other difficulties to be discussed in this chapter.

Richard G.Cooke introduced some basic definitions and mentioned a couple of characteristic properties of infinite matrices in [7]. We will talk about them in the following two sections.

3.1 Some Fundamental Definitions

An infinite matrix $I_\infty \in \mathcal{M}_\infty(F)$ defined as

$$(I_\infty)_{i,j\in\mathbb{Z}^+} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

is called the **identity matrix**. Also, the matrix O whose all entries are zero is called the **zero matrix**. Clearly, $OX = XO = O$ for any choice of $X = (x_{ij})_{i,j\in\mathbb{Z}^+}$; however, it is not necessary that X or Y must be the zero matrix so that their product is zero.

A matrix $D = (d_{ij})_{i,j\in\mathbb{Z}^+} \in \mathcal{M}_\infty(F)$ such that $d_{ij} = 0$ in the case when $i \neq j$ is called a **diagonal matrix.** We will use the notation $(d_i)_{i\in\mathbb{Z}^+}$ to denote a diagonal matrix

$$D = \begin{pmatrix} d_1 & 0 & 0 & ... \\ 0 & d_2 & 0 & ... \\ 0 & 0 & d_3 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

A diagonal matrix X is said to be a **scalar matrix** if

$$X = \lambda I_\infty = \begin{pmatrix} \lambda & 0 & 0 & ... \\ 0 & \lambda & 0 & ... \\ 0 & 0 & \lambda & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \text{ where } \lambda \in F.$$

Let X be a matrix in $\mathcal{M}_\infty(F)$ such that $x_{ij} = 0$ when $j > i$. Such a matrix is called **lower triangular matrix**. Also, a matrix X is said to be an **upper triangular matrix** if $x_{ij} = 0$ when $j < i$.

We call a matrix X **column- finite** if each column of X consists of finite number of non-zero elements. The symmetric definition says that a matrix X whose each row has only a finite number of non-zero coefficients is called **row-finite**. Note that a lower triangular matrix is row-finite, and an upper triangular matrix is column-finite, as can be seen directly from their definitions. We must keep these definitions in mind because we will benefit a lot from them throughout this thesis.

3.2 A couple of properties of infinite matrices

Let $(d_i)_{i\in\mathbb{Z}^+}$ and $(\delta_i)_{i\in\mathbb{Z}^+}$ be diagonal matrices. Clearly, their product matrix is the diagonal matrix $(d_i\delta_i)_{i\in\mathbb{Z}^+}$. Note that multiplication is commutative for diagonal matrices since we have "$d_i\delta_i = \delta_i d_i$" for each $i \in \mathbb{Z}^+$. However, in general, products of matrices fails to satisfy commutative law since $(XY)_{ik} = \sum_{j=1}^{\infty} x_{ij}y_{jk}$ may be different from $(YX)_{ik} = \sum_{j=1}^{\infty} y_{ij}x_{jk}$ for some $i, k \in \mathbb{Z}^+$, even if it is supposed that both series are convergent for any choice of i and k. Even more interesting is that one may encounter matrices $X, Y \in \mathcal{M}_\infty(F)$ such that XY does not exist while YX exists. As an example for this, consider

$$X = \begin{pmatrix} 1 & 1 & 1 & ... \\ 1 & 1 & 1 & ... \\ 1 & 1 & 1 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Y = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

As can be seen easily, XY does not exist, whereas $YX = X$ exists.

Fortunately, there is no problem of existence on the addition of matrices. In other words, their sum always exist. Also, infinite matrix addition is commutative and satisfy the associativity law, unlike multiplication:

$$X + Y = Y + X, \ \ (X + Y) + Z = X + (Y + Z) \text{ for all } X, Y, Z \in \mathcal{M}_\infty(F).$$

It should also be noted that infinite matrices satisfy the distributive law

$$X(Y + Z) = XY + XZ, \ \ (Y + Z)X = YX + ZX$$

in the sense that "if XY and XZ exist, then also $X(Y+Z)$ exists and is equal to $XY+XZ$." However, notice that $X(Y + Z)$ may exist even though XY and XZ do not exist. One can check that the following matrices constitute an example for such a case.

$$X = \begin{pmatrix} 1 & 1 & 1 & ... \\ 1 & 1 & 1 & ... \\ 1 & 1 & 1 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, Y = \begin{pmatrix} -1 & 0 & 0 & \dots \\ -1 & 0 & 0 & \dots \\ -1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Z = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Now, let $D = (d_i)_{i \in \mathbb{Z}^+}, \Delta = (\delta_i)_{i \in \mathbb{Z}^+}$ and $\mathfrak{Z} = (\mathfrak{c}_i)_{i \in \mathbb{Z}^+}$ be diagonal matrices in $\mathcal{M}_\infty(F)$. Using the above observations on diagonal matrices, it is a quick check to see that

$$(D\Delta)(\mathfrak{Z}) = D(\Delta\mathfrak{Z}).$$

In other words, the multiplication of diagonal matrices satisfy associativity law, and so the subspace $\mathcal{D}_\infty(F)$ of all diagonal matrices forms a ring in contrast to $\mathcal{M}_\infty(F)$. As we see, a subspace of $\mathcal{M}_\infty(F)$ may have the structure of a ring. Now, we shall determine whether this can be extended to lower triangular matrices. Let $X = (x_{ij})_{i,j \in \mathbb{Z}^+}$ and $Y = (y_{jk})_{j,k \in \mathbb{Z}^+}$ be both lower triangular matrices. Then

$$(XY)_{ik} = \begin{cases} \sum_{j=k}^{i} x_{ij} y_{jk} & i \geq k \\ 0 & i < k \end{cases}$$

Hence, the subspace of all infinite lower triangular matrices is closed under matrix multiplication. Again, if Z is a third lower triangular matrix, then we have

$$(X(YZ))_{il} = \sum_{j=l}^{i} x_{ij} \left(\sum_{k=l}^{j} y_{jk} z_{kl} \right) \text{ and } ((XY)Z)_{il} = \sum_{k=l}^{i} \left(\sum_{j=k}^{i} x_{ij} y_{jk} \right) z_{kl}$$

One can show that above sums are equal to each other for all i, l. Thus, associativity holds for all lower triangular matrices. Also, note that row-finite and column-finite matrices also meet associativity. (In Section 3.3, we present an easier method of deciding whether or not a subset of $\mathcal{M}_\infty(F)$ satisfy associativity.)

However, associativity need not be satisfied on the multiplication of infinite matrices. Let's consider the following example :

$$X = \begin{pmatrix} 1 & 1 & 1 & ... \\ 0 & 1 & 1 & ... \\ 0 & 0 & 1 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, Y = \begin{pmatrix} 1 & -1 & 0 & \dots \\ 0 & 1 & -1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Z = \begin{pmatrix} 0 & 0 & 0 & \dots \\ -1 & 0 & 0 & \dots \\ -1 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Easy calculations show that

$$X = X(YZ) \neq (XY)Z = Z.$$

We have experienced that the products of infinite matrices may not be defined or may not meet the associativity property. Daniel P. Bossaller and Sergio R. López-Permouth conducted a study addressing these issues in [4]. In the following two section, we will concentrate on this article.

3.3 Associativity

Firstly, we shall introduce some notation. We denote the subspace of all column (row) finite matrices by $\mathcal{M}_{Cf}(F)$ (respectively, $\mathcal{M}_{Rf}(F)$). The symbol $\mathcal{M}_{RCf}(F)$ also stands for the space $\mathcal{M}_{Rf}(F) \cap \mathcal{M}_{Cf}(F)$ of row and column finite matrices where each row and column has a finite number of nonzero entries; however, note that a matrix $X \in \mathcal{M}_{RCf}(F)$ has a chance to include infinitely many nonzero entries. Finally, the subspace of finitely supported matrices, i.e., those with only finitely many nonzero entries will be denoted by $\mathcal{M}_{\infty}^{fin}(F)$. Elements in $\mathcal{M}_{\infty}^{fin}(F)$ are also known as finitary matrices.

The main goal of this section is to propose sufficient and necessary conditions for associativity. First, as you would appreciate, it should be clarified when the product of two infinite matrices exists so that we can talk about associativity. Let's take $X, Y \in \mathcal{M}_{\infty}(F)$. Then we can represent the product XY as the following array of formal sums:

$$XY = \begin{pmatrix} \sum_{j=1}^{\infty} x_{1j}y_{j1} & \sum_{j=1}^{\infty} x_{1j}y_{j2} & \sum_{j=1}^{\infty} x_{1j}y_{j3} & \cdots \\ \sum_{j=1}^{\infty} x_{2j}y_{j1} & \sum_{j=1}^{\infty} x_{2j}y_{j2} & \sum_{j=1}^{\infty} x_{2j}y_{j3} & \cdots \\ \sum_{j=1}^{\infty} x_{3j}y_{j1} & \sum_{j=1}^{\infty} x_{3j}y_{j2} & \sum_{j=1}^{\infty} x_{3j}y_{j3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In order for the product to be defined

$$(XY)_{ik} = \sum_{j=1}^{\infty} x_{ij}y_{jk}$$

must converge for any choice of $i, k \in \mathbb{Z}^+$. To put it another way, the sums in each entry of XY need to have finitely many nonzero elements. We present this crucial observation as a definition for when a matrix product is "defined."

Definition 3.3.1. Let $X, Y \in \mathcal{M}_\infty(F)$. Then we will say that their product is **defined** if $\{j \mid x_{ij}y_{jk} \neq 0\}$ is finite for any choice of $i, k \in \mathbb{Z}^+$.

Now, we shall give the proof of a well-known basic fact we use frequently in this thesis.

Lemma 3.3.2. *If $X \in \mathcal{M}_{Rf}(F)$ and $Y \in \mathcal{M}_{Cf}(F)$, then XZ and ZY are both defined for any matrix $Z \in \mathcal{M}_\infty(F)$.*

Proof. Let $i, k \in \mathbb{Z}^+$. By above definition, we need to verify $\{j \mid x_{ij}y_{jk} \neq 0\}$ is finite to show that the product ZY exists. Clearly, $\{j \mid y_{jk} \neq 0\}$ is finite as Y is column finite, and this immediately implies that the set $\{j \mid x_{ij}y_{jk} \neq 0\}$ is also finite. Hence, we are done. In a similar way, one can get definedness of XZ.

□

Before working on associativity, it would be good for us to talk about how important associativity is in matrix algebra. One of the most critical problems that linear algebra deals with is solving a linear equation system in several unknowns. Let's remember how to find the solution of a system of m linear equation in n unknowns:

1. Convert the given system to a matrix equation of the form

$$XY = Z,$$

   where X is an $m \times n$ matrix, Y and Z are columnn matrices with n and m entries, respectively.

2. Reduce the matrix X to a row reduced echelon matrix X' by applying elementary row operations $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, ..., \mathcal{E}_r$.

3. After the previous step, we get a solvable system

$$(\mathcal{E}X)Y = \mathcal{E}Z,$$

   where $\mathcal{E}$ is the product of the elemantary matrices above. (Notice that $\mathcal{E}$ is invertible since it is product of elemantary matrices.)

4. Let λ be the solution of the solvable system

$$(\mathcal{E}X)Y = \mathcal{E}Z.$$

If we multiply this new system by $\mathcal{P}$ from the left, where $\mathcal{P}$ is the inverse of $\mathcal{E}$, we obtain that

$$\mathcal{P}(\mathcal{E}X)\lambda = (\mathcal{P}\mathcal{E})X\lambda = \mathcal{P}\mathcal{E}Z.$$

Accordingly, it must be $X\lambda = Z$. That, λ is a solution of $XY = Z$.

The process mentioned above works since the product of finite matrices obeys the associativity law. As you guessed, this cannot be extended to infinite matrices since associativity may not be satisfied there. Now, let us consider the following matrices, which we have already observed in section 3.2 again:

$$X = \begin{pmatrix} 0 & 0 & 0 & ... \\ -1 & 0 & 0 & ... \\ -1 & -1 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \mathcal{E} = \begin{pmatrix} 1 & -1 & 0 & \dots \\ 0 & 1 & -1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } \mathcal{P} = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 0 & 1 & 1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Easy calculations give us that $\mathcal{E}X$, $\mathcal{P}\mathcal{E}$ are equal to identity matrix and $\mathcal{E}Z = O$, where Z is the infinite matrix whose all entries are "1." Of course, $XY = Z$ has no solution because X has a first row whose all elements are zero. Moreover, one can easily see that $(\mathcal{E}X)Y = \mathcal{E}Z$ has a solution, which is the zero vector $\lambda = O$. As we see, even though the zero vector is a solution of $(\mathcal{E}X)Y = \mathcal{E}Z$ and $\mathcal{P}\mathcal{E} = I_\infty$, we have $XO \neq Z$. The reason for this result to occur is that $\mathcal{P}(\mathcal{E}X) \neq (\mathcal{P}\mathcal{E})X$. This observation is one of the main problems we may encounter while working with infinite matrices because of associativity. In order to avoid such obstacles, we need to know under what conditions associativity is satisfied. In this section, we try to answer this crucial question. Let's take three arbitrary matrices $X, Y, Z \in \mathcal{M}_\infty(F)$. All these examples we have seen so far showed us that it might be $(XY)Z \neq X(YZ)$ even if $XY, YX, X(YZ)$, and $(XY)Z$ are defined. Worse, the product matrices XY and YZ may not even be defined. Thus, to coin some terminology will be useful for us to tackle that.

Definition 3.3.3. Let (X, Y, Z) be triple of infinite matrices. We call it **associative triple** or **associative family** if it satisfies the followings:

1. XY and YZ are defined,

2. $X(YZ)$ and $(XY)Z$ are defined, and

3. $X(YZ) = (XY)Z$.

Remark 3.3.4. The above description may seem a little strange at first glance since the statement "$X(YZ) = (XY)Z$" naturally implies that the products $XY, YZ, X(YZ)$ and $(XY)Z$ are all defined; namely, if the third condition is met, there is no need to check the others. We give such a classification because we will talk about three vector spaces built on these requirements.

Now, we observe the following triple of matrices which does not form a associative family even though the first two condition is satisfied:

$$X = \begin{pmatrix} 1 & 1 & 1 & ... \\ 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, Y = \begin{pmatrix} 0 & 1 & 0 & \dots \\ -1 & 0 & 1 & \dots \\ 0 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Z = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Then one can easily compute that

$$XY = \begin{pmatrix} -1 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } YZ = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In addition to this, $X(YZ)$ and $(XY)Z$ are also defined since XY and YZ are both row and column finite matrices. However,

$$X(YZ) = \begin{pmatrix} 1 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \neq (XY)Z = \begin{pmatrix} -1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Let (X, Y, Z) be triple of matrices. As we have just said, our primary goal is to find out under what conditions associativity is met; namely, $(XY)Z = X(YZ)$. In order to achieve our goal, we take two different approaches. Firstly, we will focus on the properties of X and Z, which enable this triple (X, Y, Z) to form an associative family. The following definitions make our job easier in the first approach.

Definition 3.3.5. Let $X, Z \in \mathcal{M}_\infty(F)$. We call Y a **link** between X and Z if XY and YZ are both defined. We say that Y is a **strong link** between X and Z if Y is a link and $X(YZ)$ and $(XY)Z$ are defined. Also, a matrix Y is said to be **associative link** between X and Z if Y is a strong link and $X(YZ) = (XY)Z$. The family of links, strong links and associative links are shown by $\Gamma_2(X,Z)$, $\Gamma_4(X,Z)$ and $\Gamma_5(X,Z)$, respectively.

It is not hard to see that $\Gamma_i(X,Z)$ has the structure of a vector subspace of $\mathcal{M}_\infty(F)$ for each $i \in${2, 4, 5}. Moreover, from definitions of our families, it immediately follows that

$$\Gamma_5(X,Z) \subseteq \Gamma_4(X,Z) \subseteq \Gamma_2(X,Z).$$

Now, we shall examine the following matrices to show that both of the containments may be proper: Let X and Z be the matrices in the previous example and put

$$Y = \begin{pmatrix} 0 & 1 & 0 & ... \\ -1 & 0 & 1 & ... \\ 0 & -1 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Y' = \begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 1 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

We have already observed that $Y \in \Gamma_4(X,Z) \setminus \Gamma_5(X,Z)$. Moreover, it is a quick check to conclude that $Y' \in \Gamma_2(X,Z) \setminus \Gamma_4(X,Z)$.

Now, we shall introduce some propositions characterizing these subspaces.

Proposition 3.3.6. $\displaystyle\bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_2(X,Z) = \mathcal{M}_{RCf}(F).$

Proof. Let's take two arbitrary elements $X, Z \in \mathcal{M}_\infty(F)$. Using Lemma 3.3.2, one can easily deduce that $\mathcal{M}_{RCf}(F) \subseteq \Gamma_2(X,Z)$. Now, in order to show the other direction assume that $Y \notin \mathcal{M}_{RCf}(F)$. Then, the matrix Y has some row or column that has nonfinite support. If we consider the matrix Z whose all entries are 1, then YZ or ZY is not defined. This shows that $Y \notin \displaystyle\bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_2(X,Z)$. Thereby, the result follows. ☐

A similar phenomenon occurs with $\Gamma_4(X,Z)$.

Proposition 3.3.7. $\displaystyle\bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_4(X,Z) = \mathcal{M}_\infty^{fin}(F).$

Proof. We can directly get one inclusion using Lemma 3.3.2: Let $Y \in \mathcal{M}_\infty^{fin}(F)$. Then XY and YZ are both defined for any choice of X and Z. Besides, $X(YZ)$ and $(XY)Z$ are also defined since $XY \in \mathcal{M}_{Rf}(F)$ and $YZ \in \mathcal{M}_{Cf}(F)$. Hence, we obtain

$$\mathcal{M}_\infty^{fin}(F) \subseteq \bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_4(X, Z).$$

For the other inclusion, let $Y \in \Gamma_4(X, Z)$ for all $X, Z \in \mathcal{M}_\infty(F)$. From above proposition, it directly follows that $Y \in \mathcal{M}_{RCf}(F)$. Now, assume that Y fails to be finitely supported matrix and consider the following two sets

$$J = \{j \mid Y_{j*} \neq 0\} \text{ and } K = \{k \mid Y_{*k} \neq 0\}.$$

Since $Y \in \mathcal{M}_{RCf}(F)$ and we assumed that $Y \notin \mathcal{M}_\infty^{fin}(F)$, at least one of the above sets need to be infinite. Without loss of generality, let $|J| = \infty$. This means that the matrix Y has infinite number of non-zero rows. Now, let Z be a matrix whose all entries are 1. Then, YZ is defined since $Y \in \mathcal{M}_{RCf}(F)$; moreover, the first column of YZ has infinitely many nonzero entries due to construction of Z. However, one can easily see that $X(YZ)$ is not defined in the case when

$$X = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Of course, this is a contradiction because Y is taken from $\bigcap\limits_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_4(X, Z)$. Hence, we are done. $\square$

Now, we shall give the following techniqual definition which will be used in the proof of the next proposition.

Definition 3.3.8. Let X be a row finite matrix. The **length** of X_{i*}, which is the i-th row of X is the smallest number $\sigma_i \geq 0$ such that $x_{ij} = 0$ for every $j > \sigma_i$. Likewise, the length of the j-th column of a column finite matrix Y is the smallest number u_j such that $y_{ij} = 0$ for every $i > u_j$.

Proposition 3.3.9. *If $X \in \mathcal{M}_{Rf}(F)$ and $Z \in \mathcal{M}_{Cf}(F)$, then it must be*

$$\Gamma_5(X, Z) = \mathcal{M}_\infty(F).$$

Proof. One inclusion is trivial, $\Gamma_5(X,Z) \subseteq \mathcal{M}_\infty(F)$. Now, all we have to do is show that $Y \in \Gamma_5(X,Z)$ for any $Y \in \mathcal{M}_\infty(F)$. From Lemma 3.3.2, we know that XY, YZ, $(XY)Z$ and $X(YZ)$ are all defined for all $Y \in \mathcal{M}_\infty$. Now, we shall determine whether or not $X(YZ) = (XY)Z$ calculating arbitrary (i,l)-th entry of both of those terms:

$$(X(YZ))_{il} = \sum_{j=1}^{\infty} x_{ij} \left(\sum_{k=1}^{\infty} y_{jk} z_{kl} \right) = \sum_{j=1}^{\sigma_i} x_{ij} \left(\sum_{k=1}^{u_l} y_{jk} z_{kl} \right)$$
$$= \sum_{j=1}^{u_l} \left(\sum_{k=1}^{\sigma_i} x_{ij} y_{jk} \right) z_{kl} = ((XY)Z)_{il},$$

where σ_i is the length of the i-th row of X and u_l is the length of the l-th column of Z.

□

Corollary 3.3.10. *Let $Y \in \mathcal{M}_\infty^{fin}(F)$. Then $X(YZ) = (XY)Z$ for any matrices X and Z. In particular,*

$$\bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_5(X,Z) = \mathcal{M}_\infty^{fin}(F).$$

Proof. Since Y is a finitely supported matrix, one may divide Y into block matrices as follows

$$Y = \left(\begin{array}{c|c} Y' & O \\ \hline O & O \end{array} \right),$$

where Y' is an $n \times n$ matrix. Now, let's write the matrices X and Z in terms of block matrices similarly:

$$X = \left(\begin{array}{c|c} X_1 & X_2 \\ \hline X_3 & X_4 \end{array} \right) \quad \text{and} \quad Z = \left(\begin{array}{c|c} Z_1 & Z_2 \\ \hline Z_3 & Z_4 \end{array} \right),$$

where X_1 and Z_1 are $n \times n$ matrices. Then,

$$X(YZ) = \left(\begin{array}{c|c} X_1(Y'Z_1) & X_1(Y'Z_2) \\ \hline X_3(Y'Z_1) & X_3(Y'Z_2) \end{array} \right) \text{ and } (XY)Z = \left(\begin{array}{c|c} (X_1Y')Z_1 & (X_1Y')Z_2 \\ \hline (X_3Y')Z_1 & (X_3Y')Z_2 \end{array} \right).$$

Since $X_1, X_3 \in \mathcal{M}_{Rf}(F)$ and $Z_1, Z_2 \in \mathcal{M}_{Cf}(F)$, it must be $X(YZ) = (XY)Z$ from the previous proposition. Moreover, the second statement follows from Proposition 3.3.7 because of the fact

$$\mathcal{M}_\infty^{fin}(F) \subseteq \bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_5(X,Z) \subseteq \bigcap_{X,Z \in \mathcal{M}_\infty(F)} \Gamma_4(X,Z).$$

□

In our second approach, we focus on the matrix Y in the statement $X(YZ) = (XY)Z$. We will investigate how it affects associativity. Here we should introduce the concept of summability to understand the work to be done in the following subsection.

Definition 3.3.11. Let $\{V_i \mid i \in I\}$ be a family of vectors indexed by the set I. Then we call this family $\{V_i \mid i \in I\}$ summable if for every $j \in \mathbb{Z}^+$, the set $\{i \mid V_i(j) \neq 0\}$ is finite, where $V_i(j)$ is the j-th entry in the i-th vector.

One may wonder how to use this new concept in our study of infinite matrices: We can extend this definition to the context of infinite matrices considering the columns (rows) of a matrix as a vector. (Note that, in such a case, our index set will be $\mathbb{Z}^+$.)

3.3.1 Summability and Associativity

Let $X, Z \in \mathcal{M}_\infty(F)$. One can easily deduce that Y is a link between them if and only if for every $k \in \mathbb{Z}^+$ the family $\{X_{*j}y_{jk} \mid j \in \mathbb{Z}^+\}$ is summable and for every $j \in \mathbb{Z}^+$, $\{y_{jk}Z_{k*} \mid k \in \mathbb{Z}^+\}$ is summable. At this point, we face a quite important question: Is ıt enough to check summability of the family $\{X_{*j}y_{jk}Z_{k*} \mid j, k \in \mathbb{Z}^+\}$ to determine whether (X, Y, Z) is an associative triple or not? Unfortunately, the answer is "No!" The following three matrices constitute a pretty good example of a triple which does not form an associative family even though $\{X_{*j}y_{jk}Z_{k*} \mid j, k \in \mathbb{Z}^+\}$ is summable. Now, put

$$X = \begin{pmatrix} 0 & 1 & 1 & ... \\ 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, Y = \begin{pmatrix} 1 & 1 & 1 & \cdots \\ 1 & 1 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Z = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Clearly, these matrices do not even satisfy the first condition to be an associative triple as XY and YZ do not exist. The surprising point here is that the family $\{X_{*j}y_{jk}Z_{k*} \mid j, k \in \mathbb{Z}^+\}$ is summable. Let's explain why:

Actually, all we have to do is multiply the columns and rows by the nonzero entries of Y which lie in the first row, first column, and $(2, 2)$ coordinate. First, we consider the y_{22}. Then we have

$$X_{*2}y_{22}Z_{2*} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & \dots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \neq O.$$

Secondly, let's look at the matrix X. As can be seen, the first column of X is zero. Thus, the family generated by any entry chosen from the first row of Y

$$\left\{X_{*1}y_{1k}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$$

must be the singleton set {O}. In addition to this, a symmetric argument holds for the family $\left\{X_{*j}y_{j1}Z_{1*} \mid j,k \in \mathbb{Z}^+\right\}$, where y_{j1} is chosen arbitrarily from the first column Y, since the entries of first row of Z are all zero. Consequently, the family $\left\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$ is summable since it has only one element which is different from zero, $X_{*2}y_{22}Z_{2*}$.

We couldn't find what we expected, and even worse, we faced the fact that summability of $\left\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$ does not force XY and YZ to be defined. On the other hand, the summability of this family is a necessary condition for associativity as our next work will show us. That is, we're not done with it yet. Hence, from this point on we will call this summability condition **requirement** ($\mathcal{S}$) for convenience. By the way, we have one more fact which is easy to see to talk about before we go into details: Let (X,Y,Z) be triple of matrices in $\mathcal{M}_\infty(F)$. Then the family $\left\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$ is summable if and only if the set $\mathcal{U} = \left\{j,k \in \mathbb{Z}^+ \mid x_{ij}y_{jk}z_{kl} \neq 0\right\}$, which is the support of arbitrarily chosen (i,l)-th entry is summable. Let us continue our discussion with the following result, which tells us that if we add one more assumption, summability of the family $\left\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$ implies associativity.

Theorem 3.3.12. *Let X,Y and Z be matrices such that XY,YZ are both defined and $\left\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\right\}$ is summable, then $X(YZ) = (XY)Z$.*

Proof. In order for the desired result to follows, we will try to show that the (i,l)-th entry of $X(YZ)$ and $(XY)Z$ are equal for any choice of $i,l \in \mathbb{Z}^+$. In other words, if we write them in terms of formal expressions, we would like to obtain the following equality

$$(X(YZ))_{il} = \sum_{j=1}^{\infty} x_{ij}\left(\sum_{k=1}^{\infty} y_{jk}z_{kl}\right) = \sum_{k=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{ij}y_{jk}\right)z_{kl} = ((XY)Z)_{il}.$$

By the definedness of YZ, we know that the set $\{k \mid y_{jk}z_{kl} \neq 0\}$ is finite. Hence, for any $j \in \mathbb{Z}^+$, we have a smallest number t_j satisfying

$$(X(YZ))_{il} = \sum_{j=1}^{\infty} x_{ij}\left(\sum_{k=1}^{\infty} y_{jk}z_{kl}\right) = \sum_{j=1}^{\infty} x_{ij}\left(\sum_{k=1}^{t_j} y_{jk}z_{kl}\right) = \sum_{j=1}^{\infty}\sum_{k=1}^{t_j} x_{ij}y_{jk}z_{kl}.$$

Moreover, each entry in the last summand above has finite support because **requirement** ($\mathcal{S}$) is satisfied by hypothesis. Hence, there is a smallest number q such that

$$\sum_{j=1}^{\infty}\sum_{k=1}^{t_j} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q}\sum_{k=1}^{t_j} x_{ij}y_{jk}z_{kl}.$$

Now, let's define $t = \mathrm{Max}\{t_j \mid 1 \leq j \leq q\}$. Then we obtain that

$$\sum_{j=1}^{\infty}\sum_{k=1}^{t_j} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q}\sum_{k=1}^{t_j} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl}.$$

A symmetric argument holds for $((XY)Z)_{il}$. The set $\{j \mid x_{ij}y_{jk} \neq 0\}$ is finite since XY is defined. Therefore, for any $k \in \mathbb{Z}^+$, there must exist a smallest number q'_k satisfying the following:

$$((XY)Z)_{il} = \sum_{k=1}^{\infty}\left(\sum_{j=1}^{\infty} x_{ij}y_{jk}\right)z_{kl} = \sum_{k=1}^{\infty}\left(\sum_{j=1}^{q'_k} x_{ij}y_{jk}\right)z_{kl} = \sum_{k=1}^{\infty}\sum_{j=1}^{q'_k} x_{ij}y_{jk}z_{kl}$$

Since $\{X_{*j}y_{jk}Z_{k*} \mid j,k \in \mathbb{Z}^+\}$ is summable, one can find a number t' such that

$$\sum_{k=1}^{\infty}\sum_{j=1}^{q'_k} x_{ij}y_{jk}z_{kl} = \sum_{k=1}^{t'}\sum_{j=1}^{q'_k} x_{ij}y_{jk}z_{kl}.$$

If we define $q' = \mathrm{Max}\{q'_k \mid 1 \leq k \leq t'\}$, then it must be

$$((XY)Z)_{il} = \sum_{k=1}^{t'}\sum_{j=1}^{q'_k} x_{ij}y_{jk}z_{kl} = \sum_{k=1}^{t'}\sum_{j=1}^{q'} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q'}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}.$$

From here on all we need to do is check whether or not

$$\sum_{j=1}^{q}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q'}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}. \qquad (*)$$

Let's assume that the above equality does not hold. Then we have the following eight cases, one of which must occur:

• $q' > q$ and $t' = t$	• $q > q'$ and $t = t'$
• $q' = q$ and $t' > t$	• $q = q'$ and $t > t'$
• $q' > q$ and $t' > t$	• $q > q'$ and $t > t'$
• $q > q'$ and $t' > t$	• $q' > q$ and $t > t'$

Now, we shall observe these cases. Let's start with the first case. If $q' > q$ and $t' = t$, then we have

$$\sum_{j=1}^{q}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl} + \sum_{j=q+1}^{q'}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q'}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}.$$

Notice that $x_{ij}y_{jk}z_{kl} = 0$ for all $j > q$ by construction of q. This forces the second summand to be zero. Hence, the equation $(*)$ holds in this case.

Second, if $q' = q$ and $t' > t$, then

$$\sum_{j=1}^{q}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl} + \sum_{j=1}^{q}\sum_{k=t+1}^{t'} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q'}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}.$$

We know that $y_{jk}z_{kl} = 0$ for all $k > t$ by definition of t, and so the second summand is zero again. Thereby, the equation $(*)$ holds in this case too. Moreover, the third case, where $q' > q$ and $t' > t$, immediately follows from the first two cases. Let's continue with the fourth case. One can write the following two equations which are equal to each other in the case when $q > q'$ and $t' > t$:

$$\sum_{j=1}^{q'}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl} + \sum_{j=q'+1}^{q}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}$$

$$\sum_{j=1}^{q}\sum_{k=1}^{t} x_{ij}y_{jk}z_{kl} + \sum_{j=1}^{q}\sum_{k=t+1}^{t'} x_{ij}y_{jk}z_{kl} = \sum_{j=1}^{q}\sum_{k=1}^{t'} x_{ij}y_{jk}z_{kl}$$

Recall that $t = \text{Max}\{t_j \mid 1 \leq j \leq q\}$ and $q' = \text{Max}\{q'_k \mid 1 \leq k \leq t'\}$. Thus, we have $x_{ij}y_{jk} = 0$ for all $j > q'$ and $y_{jk}z_{kl} = 0$ for all $k > t$. This forces the second summands of each equation above to be zero. This gives us the desired equality $(*)$. Through a symmetric process, one may prove the remaining cases. □

Let's make use of this crucial theorem. We continue with important consequences of it. In fact, the first one is nothing more than Corollary 3.3.10 with an easier proof.

Corollary 3.3.13. *Let $Y \in \mathcal{M}_{\infty}^{fin}(F)$. Then for any $X, Z \in \mathcal{M}_{\infty}(F)$, it must be*

$$X(YZ) = (XY)Z.$$

Proof. Clearly, the matrices XY and YZ are both defined as $Y \in \mathcal{M}_{\infty}^{fin}(F)$. Besides, it is not hard to see that family $\{X_{*j}y_{jk}Z_{k*} \mid j, k \in \mathbb{Z}^{+}\}$ is summable because Y consists of finitely many nonzero elements. Hence, $X(YZ) = (XY)Z$ by the above theorem.

□

Corollary 3.3.14. *If $X, Y \in \mathcal{M}_{Rf}(F)$, then $X(YZ) = (XY)Z$ for any $Z \in \mathcal{M}_{\infty}(F)$. Likewise, $X(YZ) = (XY)Z$ for any $X \in \mathcal{M}_{\infty}(F)$ in the case when Y and Z column-finite matrices.*

Proof. Let σ_i be the length of the i-th row of the matrix X, which is row-finite. Then it must be $x_{ij}y_{jk} = 0$ for every $j > \sigma_i$. Also, note that the rows Y_{j*} have finitely many nonzero elements since Y is also row-finite, and so we have only finitely many $y_{jk} \neq 0$ for $j \leq \sigma_i$. Therefore, the family $\{X_{*j}y_{jk}Z_{k*} \mid j, k \in \mathbb{Z}^{+}\}$ summable. Thus, (X, Y, Z) is an associative triple by the theorem. The second statement also follows in a similar way. □

Finally, we close out this section with the following proposition:

Proposition 3.3.15. *Let X, Y and Z be matrices such that YZ is defined and X is row finite. Then it must be $X(YZ) = (XY)Z$.*

Proof. One can immediately say that XY is defined as X is row finite. In order to see that $Y \in \Gamma_5(X, Z)$, all we will do is determine whether **requirement** ($\mathcal{S}$) is satisfied or not. Now, let's consider the set $\mathcal{U} = \{j, k \in \mathbb{Z}^{+} \mid x_{ij}y_{jk}z_{kl} \neq 0\}$, which is the support of arbitrarily chosen (i, l)-th entry. It follows from the definedness of YZ that $\{k \mid y_{jk}z_{kl} \neq 0\}$ is finite for any choice of $j \in \mathbb{Z}^{+}$. Besides, $\{j \mid x_{ij} \neq 0\}$ is also finite as X is row-finite, and so the set $\mathcal{U} = \{j, k \in \mathbb{Z}^{+} \mid x_{ij}y_{jk}z_{kl} \neq 0\}$ is finite. Hence, **requirement** ($\mathcal{S}$) is satisfied. This completes the proof. □

Remark 3.3.16. As we mentioned, $\mathcal{M}_\infty(F)$ is not a ring since it fails to satisfy associativity. But, the subspaces $\mathcal{M}_{Cf}(F)$, $\mathcal{M}_{Rf}(F)$ and $\mathcal{M}_\infty^{fin}(F)$ form an associative ring since the **requirement** ($\mathcal{S}$) satisfied in these subspaces. We should note here that the result also holds in the case when the coefficients of matrices are taken from a ring with "1" instead of a field of characteristic zero.

As we said before, in general, a variety of theorems in the context of finite matrices cannot be generalized to infinite matrices by simply letting n tend to ∞ due to obstacles such as convergence, associativity. For example, W.Hołubowski ([12]) showed that inverse of an upper triangular matrix might be lower triangular, which is a case not in the finite-matrix theory. We will deal with this article briefly to tell why we should be prepared for differences and careful while working with infinite matrices.

3.4 A Little Warning

In this section, we shall take a look at a couple of examples that show us that fundamental theorems on $n \times n$ matrices with real entries which play a key role in finite matrix theory may not hold in the case of infinite matrices.

Let X and Z be two $n \times n$ matrices. Recall that if $XZ = I_n$, then we also have $ZX = I_n$. However, the result fails to be satisfied in the case when our matrices are infinite. As an example, put

$$X = \begin{pmatrix} 0 & 0 & 0 & 0 & ... \\ 1 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Z = \begin{pmatrix} 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & ... \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Then, easy calculations give

$$XZ = \begin{pmatrix} 0 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & ... \\ 0 & 0 & 0 & 1 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } ZX = \begin{pmatrix} 1 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & ... \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = I_\infty.$$

As can be seen above, $XZ \neq I_\infty$ although $ZX = I_\infty$. Thus, this theorem does not work for infinite matrices. We also know that

- If $X = (x_{ij})_{1 \leq i,j \leq n}$ is an invertible lower (upper) triangular matrix, then its inverse X^{-1} must also be a lower (upper) triangular matrix.

- Let X be an triangular $n \times n$ matrix. Then it is invertible if and only if all its diagonal entries are invertible.

Now, we shall try to find two matrices X', Y' with entries in an infinite matrix ring to show that these results also may not work for infinite matrices.

Let $X' = \begin{pmatrix} X & Y \\ O & Z \end{pmatrix}$ be a 2×2 matrix with coefficients from the $\mathcal{M}_{RCf}(F)$, where X, Z are infinite matrices which we defined above and

$$Y = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Now, we consider $Y' = \begin{pmatrix} Z & O \\ Y & X \end{pmatrix}$. One can easily compute that $X'Y' = Y'X' = E$, where $E = \begin{pmatrix} I_\infty & 0 \\ 0 & I_\infty \end{pmatrix}$ is a unit matrix in $\mathcal{M}_2(\mathcal{M}_{RCf}(F))$. Thus, X' and Y' are both invertible. To conclude, we see that inverse of the upper triangular matrix X' that we defined above is a lower triangular matrix, which is Y'. Moreover, as can be seen easily, the diagonal entries of X' and Y' are not invertible. This is exactly what we wanted to show.

Chapter 4

Derivations of Rings of Infinite Matrices

4.1 Definitions and Statement of Results

Let R be an associative ring with 1. An additive map $\mathfrak{D}$ of this ring R is called a derivation of R if it satisfies the Leibniz Rule $\mathfrak{D}(xy) = \mathfrak{D}(x)y + x\mathfrak{D}(y)$ for all $x, y \in R$. Also, if x is any element of this ring, then the additive map

$$\begin{aligned} \mathfrak{D}_x : \quad R &\rightarrow \quad R \\ y &\rightarrow \quad xy - yx \end{aligned}$$

defines a derivation. We call such a derivation inner. Characterizing all derivations of a ring has always been of interest to many mathematicians studying algebra. Of course, one of the first ring structures that come to mind are matrix rings. First, let us briefly summarize what we know about derivations of the ring $\mathcal{M}_n(R)$ of all $n \times n$ matrices over R. Every derivation of $\mathcal{M}_n(R)$ is an inner derivation in the case of R being a field (see [10]). What one need to keep in mind is that this result does not have to work if R is not a field. We have a practical way of constructing a derivation of $\mathcal{M}_n(R)$: Let $\mathfrak{C}$ be a derivation of R. Then the map $\mathfrak{D}_{\mathfrak{C}(R)} : \ \mathcal{M}_n(R) \ \rightarrow \ \mathcal{M}_n(R)$ which is defined by

$$(\mathfrak{D}_{\mathfrak{C}(R)}(X))_{ij} = \mathfrak{C}(x_{ij}) \quad \text{for all} \ \ i, j$$

turns out to be a derivation of $\mathcal{M}_n(R)$, called induced by $\mathfrak{C}$, or shortly induced. Namely, a derivation of a ring R enable us to construct a new derivation of $\mathcal{M}_n(R)$.

Now, let us broaden our perspective. What can we say about derivations of infinite matrices? As we mentioned in the previous chapter, the set $\mathcal{M}_\infty(R)$ of all matrices over R, whose rows and columns are indexed by $\mathbb{Z}^+ \times \mathbb{Z}^+$, has no structure of a ring due to some problems such as convergence, associativity. Thus, it would be better to consider its subsets which form a ring such as the set $\mathcal{M}_{Cf}(R)$ ($\mathcal{M}_{Rf}(R)$) of all column finite matrices (row finite matrices). For example, Kolesnikov and Maltsev described all derivations of the ring $\mathcal{M}_{\infty}^{fin}(R)$ of all finitely supported matrices in [14]. They deduced that any derivation of $\mathcal{M}_{\infty}^{fin}(R)$ can be written as a sum of an inner derivation $\mathfrak{D}_X$ for some $X \in \mathcal{M}_{Cf}(R) \cap \mathcal{M}_{Rf}(R)$ and an induced derivation $\mathfrak{D}_{\mathfrak{C}(R)}$. In [20], Slowik achieved a similar result for $\mathcal{M}_{Cf}(R)$ and the ring $\mathcal{T}_{\infty}(R)$ of all upper triangular matrices. Here we will discuss this article in detail. Throughout this chapter, R denotes an associative ring with identity. Before moving onto the next section, let us state precisely what Slowik proved:

Theorem 4.1.1. *Let $\mathfrak{D}$ be a derivation of $\mathcal{M}_{Cf}(R)$. Then there exist a matrix $X \in \mathcal{M}_{Cf}(R)$ and a derivation $\mathfrak{C}$ of R such that*

$$\mathfrak{D}(Y) = \mathfrak{D}_X(Y) + \mathfrak{D}_{\mathfrak{C}(R)}(Y) \quad \textit{for all } \; Y \in \mathcal{M}_{Cf}(R).$$

The ring of infinite upper triangular matrices constitute a nice example of a subring of $\mathcal{M}_{Cf}(R)$. Same result moved here:

Theorem 4.1.2. *Let $\mathfrak{D}$ be a derivation of $\mathcal{T}_{\infty}(R)$. Then there exist a matrix $X \in \mathcal{T}_{\infty}(R)$ and a derivation $\mathfrak{C}$ of R such that*

$$\mathfrak{D}(Y) = \mathfrak{D}_X(Y) + \mathfrak{D}_{\mathfrak{C}(R)}(Y) \quad \textit{for all } \; Y \in \mathcal{T}_{\infty}(R).$$

4.2 Some Techniqual Propositions

This section will propose some techniqual propositions that help us prove our main desired results. Before working, let us introduce some notations and terminologies. As usual, we show the matrix whose only nonzero entry is 1 in the (i, j) coordinate by ξ_{ij}, regardless of its dimension. As we know, these are known as matrix units. Also, the symbols I_{∞} and I_n stand for the identity matrices, infinite and $n \times n$, respectively. We write $\mathcal{M}_{n\times\infty}(R)$ for the ring of all matrices over R with n rows and infinite number of columns (indexed by $\mathbb{Z}^+$). Likewise, the ring of all matrices over R with infinite number of rows (also indexed by $\mathbb{Z}^+$) and n columns is shown by $\mathcal{M}_{\infty\times n}(R)$.

Finally, it is time to start work. Let $\mathfrak{D}$ be a derivation of $\mathcal{R}$, where $\mathcal{R}$ is one of the rings $\mathcal{M}_{Cf}(R)$ or $\mathcal{T}_{\infty}(R)$. Firstly, we focus on the properties of $\mathfrak{D}(Y)$ if Y is finitary.

Proposition 4.2.1. *If $\mathfrak{D}$ is a derivation of $\mathcal{R}$, then there exists $Z \in \mathcal{M}_{Cf}(R)$ such that*

$$\mathfrak{D}(Y) = \mathfrak{D}_Z(Y) + \mathfrak{D}'(Y) \text{ for all } Y \in \mathcal{M}_\infty^{fin}(R) \cap \mathcal{R},$$

where for every i, $j \in \mathbb{Z}^+$ and $\lambda \in R$ one has $\mathfrak{D}'(\lambda\xi_{ij}) = \lambda'\xi_{ij}$ for some $\lambda' \in R$.

Proof. It will be more understandable if we give this proof step by step.

(1) For any $i \in \mathbb{Z}^+$ and $\lambda \in R$, all nonzero entries of $\mathfrak{D}(\lambda\xi_{ii})$ lie in the i-th row and in the i-th column.

Writing $\lambda\xi_{ii}$ as $\lambda\xi_{ii}.\xi_{ii}$, we have $\mathfrak{D}(\lambda\xi_{ii}) = \mathfrak{D}(\lambda\xi_{ii}.\xi_{ii})$. Since $\mathfrak{D}$ satisfies the Leibniz rule, we obtain

$$\mathfrak{D}(\lambda\xi_{ii}) = \mathfrak{D}(\lambda\xi_{ii})\xi_{ii} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ii}).$$

It is a quick check to see that $\mathfrak{D}(\lambda\xi_{ii})\xi_{ii}$ has nonzero coefficients only in the i-th column. On the other hand, the nonzero coefficients of $\lambda\xi_{ii}\mathfrak{D}(\xi_{ii})$ lie in the i-th row. Hence, we are done.

(2) Let $i \in \mathbb{Z}^+$ and $\lambda \in R$. It must be

$$\mathfrak{D}(\lambda\xi_{ii}) = \lambda\bigg(\sum_{j\neq i}(\mathfrak{D}(\xi_{ii}))_{ij}\xi_{ij}\bigg) + \bigg(\sum_{k\neq i}(\mathfrak{D}(\xi_{ii}))_{ki}\xi_{ki}\bigg)\lambda + \lambda'\xi_{ii}$$

for some $\lambda' \in R$.

In the previous step, we observed that

$$\mathfrak{D}(\lambda\xi_{ii}) = \mathfrak{D}(\lambda\xi_{ii})\xi_{ii} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ii}).$$

Using this, one can see that

$$\mathfrak{D}(\lambda\xi_{ii})(I_\infty - \xi_{ii}) = \lambda\xi_{ii}\mathfrak{D}(\xi_{ii}).$$

Thus, $(\mathfrak{D}(\lambda\xi_{ii}))_{ij} = (\lambda\mathfrak{D}(\xi_{ii}))_{ij}$ for all $j \neq i$. Let's take a different approach. Writing $\lambda\xi_{ii}$ as $\xi_{ii}.\lambda\xi_{ii}$, we obtain

$$\mathfrak{D}(\lambda\xi_{ii}) = \mathfrak{D}(\xi_{ii})\lambda\xi_{ii} + \xi_{ii}\mathfrak{D}(\lambda\xi_{ii}).$$

It immediately follows from the above equality that

$$(I_\infty - \xi_{ii})\mathfrak{D}(\lambda\xi_{ii}) = \mathfrak{D}(\xi_{ii})\lambda\xi_{ii}.$$

Hence, $(\mathfrak{D}(\lambda\xi_{ii}))_{ki} = (\mathfrak{D}(\xi_{ii})\lambda)_{ki}$ for all $k \neq i$. Combining these results, we get

$$\mathfrak{D}(\lambda\xi_{ii}) = \begin{array}{cc} \begin{matrix} \phantom{\lambda\mathfrak{D}(\xi_{ii})_{i1}} & & \phantom{\lambda\mathfrak{D}(\xi_{ii})_{i,i-1}} & i\text{-th column} & \phantom{\lambda\mathfrak{D}(\xi_{ii})_{i,i+1}} & \end{matrix} & \\ \begin{pmatrix} 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{1i}\lambda & 0 & ... \\ 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{2i}\lambda & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda\mathfrak{D}(\xi_{ii})_{i1} & ... & \lambda\mathfrak{D}(\xi_{ii})_{i,i-1} & \mathfrak{D}(\lambda\xi_{ii})_{ii} & \lambda\mathfrak{D}(\xi_{ii})_{i,i+1} & ... \\ 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{i+1,i} & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

This proves our claim.

Now, we shall examine that symmetric arguments hold for any $\mathfrak{D}(\lambda\xi_{ij})$, where $i, j \in \mathbb{Z}^+$ and $\lambda \in R$.

(3) For any $i, j \in \mathbb{Z}^+$ and $\lambda \in R$, all nonzero entries of $\mathfrak{D}(\lambda\xi_{ij})$ lie in the i-th row and in the j-th column.

If we write $\lambda\xi_{ij}$ as $\lambda\xi_{ij}.\xi_{jj}$, then $\mathfrak{D}(\lambda\xi_{ij}) = \mathfrak{D}(\lambda\xi_{ij}.\xi_{jj})$. Since $\mathfrak{D}$ satisfies the Leibniz rule, we obtain

$$\mathfrak{D}(\lambda\xi_{ij}.\xi_{jj}) = \mathfrak{D}(\lambda\xi_{ij})\xi_{jj} + \lambda\xi_{ij}\mathfrak{D}(\xi_{jj}).$$

Clearly, $\mathfrak{D}(\lambda\xi_{ij})\xi_{jj}$ has nonzero coefficients only in the j-th column. On the other hand, the nonzero coefficients of $\lambda\xi_{ij}\mathfrak{D}(\xi_{jj})$ lie in the i-th row. Accordingly, our claim follows.

(4) Let i, $j \in \mathbb{Z}^+$ and $\lambda \in R$. Then there exist $\lambda' \in R$ satisfying

$$\mathfrak{D}(\lambda\xi_{ij}) = \lambda\left(\sum_{k\neq j}(\mathfrak{D}(\xi_{jj}))_{jk}\xi_{ik}\right) + \left(\sum_{k\neq i}(\mathfrak{D}(\xi_{ii}))_{ki}\xi_{kj}\right)\lambda + \lambda'\xi_{ij}$$

for some $\lambda' \in R$.

In the previous step, we observed that

$$\mathfrak{D}(\lambda\xi_{ij}) = \mathfrak{D}(\lambda\xi_{ij})\xi_{jj} + \lambda\xi_{ij}\mathfrak{D}(\xi_{jj}).$$

If we write $\lambda\xi_{ij}$ as $\xi_{ii}.\lambda\xi_{ij}$, then we also have

$$\mathfrak{D}(\lambda\xi_{ij}) = \mathfrak{D}(\xi_{ii})\lambda\xi_{ij} + \xi_{ii}\mathfrak{D}(\lambda\xi_{ij}).$$

Thereby, it must be

$$\mathfrak{D}(\lambda\xi_{ij})\xi_{jj} + \lambda\xi_{ij}\mathfrak{D}(\xi_{jj}) = \mathfrak{D}(\xi_{ii})\lambda\xi_{ij} + \xi_{ii}\mathfrak{D}(\lambda\xi_{ij}).$$

From the above equality, we find

$$(\mathfrak{D}(\lambda\xi_{ij})\xi_{jj})_{kj} = (\mathfrak{D}(\xi_{ii})\lambda\xi_{ij})_{kj} \ \text{ for } \ k \neq i$$

and

$$(\xi_{ii}\mathfrak{D}(\lambda\xi_{ij}))_{ik} = (\lambda\xi_{ij}\mathfrak{D}(\xi_{jj}))_{ik} \ \text{ for } \ k \neq j.$$

Combining these results, one may write the matrix $\mathfrak{D}(\lambda\xi_{ij})$ which has nonzero entries only in the i-th row and in the j-th column as follows:

$$\mathfrak{D}(\lambda\xi_{ij}) = \begin{pmatrix} 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{1i}\lambda & 0 & ... \\ 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{2i}\lambda & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda\mathfrak{D}(\xi_{jj})_{j1} & ... & \lambda\mathfrak{D}(\xi_{jj})_{j,j-1} & \mathfrak{D}(\lambda\xi_{ij})_{ij} & \lambda\mathfrak{D}(\xi_{jj})_{j,j+1} & ... \\ 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{i+1,i}\lambda & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

This completes the proof of our claim.

(5) For any $i \neq j \in \mathbb{Z}^+$, it must be $(\mathfrak{D}(\xi_{jj}))_{ij} = -(\mathfrak{D}(\xi_{ii}))_{ij}$.

From additivity of $\mathfrak{D}$, we have $\mathfrak{D}(O) = O$. Also, $\xi_{ii}.\xi_{jj} = O$ in the case when $i \neq j$. Accordingly, we get $O = \mathfrak{D}(O) = \mathfrak{D}(\xi_{ii}.\xi_{jj}) = \mathfrak{D}(\xi_{ii})\xi_{jj} + \xi_{ii}\mathfrak{D}(\xi_{jj})$. Clearly, this give us desired result, $(\mathfrak{D}(\xi_{jj}))_{ij} = -(\mathfrak{D}(\xi_{ii}))_{ij}$.

(6) There exist a matrix Z such that for all $i, j \in \mathbb{Z}^+$ and $\lambda \in R$

$$\mathfrak{D}(\lambda\xi_{ij}) = Z.\lambda\xi_{ij} - \lambda\xi_{ij}.Z + \lambda'\xi_{ij} \ \text{ for some } \ \lambda' \in R.$$

To prove this claim, consider the matrix below

$$Z = \sum_{j\neq i}(\mathfrak{D}(\xi_{jj}))_{ij}\xi_{ij} = \begin{pmatrix} 0 & (\mathfrak{D}(\xi_{22}))_{12} & (\mathfrak{D}(\xi_{33}))_{13} & (\mathfrak{D}(\xi_{44}))_{14} & ... \\ (\mathfrak{D}(\xi_{11}))_{21} & 0 & (\mathfrak{D}(\xi_{33}))_{23} & (\mathfrak{D}(\xi_{44}))_{24} & ... \\ (\mathfrak{D}(\xi_{11}))_{31} & (\mathfrak{D}(\xi_{22}))_{32} & 0 & (\mathfrak{D}(\xi_{44}))_{34} & ... \\ (\mathfrak{D}(\xi_{11}))_{41} & (\mathfrak{D}(\xi_{22}))_{42} & (\mathfrak{D}(\xi_{33}))_{43} & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

For this Z, we have

$$Z.\lambda\xi_{ij} = \begin{array}{c c} \begin{matrix} & & & j\text{-th column} & & \end{matrix} & \\ \begin{pmatrix} 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{1i}\lambda & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & ... & 0 & \mathfrak{D}(\xi_{ii})_{i-1,i}\lambda & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & \mathfrak{D}(\xi_{i,i})_{i+1,i}\lambda & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

As can be seen above, the matrix $Z.\lambda\xi_{ij}$ has nonzero coefficients only in the j-th column. Furthermore, in step (5) we observed that $(\mathfrak{D}(\xi_{jj}))_{ij} = -(\mathfrak{D}(\xi_{ii}))_{ij}$ for all $i \neq j \in \mathbb{Z}^+$. Thus, $-Z = \sum_{j\neq i} -(\mathfrak{D}(\xi_{jj}))_{ij}\xi_{ij} = \sum_{j\neq i}(\mathfrak{D}(\xi_{ii}))_{ij}\xi_{ij}$,

$$-Z = \begin{pmatrix} 0 & (\mathfrak{D}(\xi_{11}))_{12} & (\mathfrak{D}(\xi_{11}))_{13} & (\mathfrak{D}(\xi_{11}))_{14} & ... \\ (\mathfrak{D}(\xi_{22}))_{21} & 0 & (\mathfrak{D}(\xi_{22}))_{23} & (\mathfrak{D}(\xi_{22}))_{24} & ... \\ (\mathfrak{D}(\xi_{33}))_{31} & (\mathfrak{D}(\xi_{33}))_{32} & 0 & (\mathfrak{D}(\xi_{44}))_{34} & ... \\ (\mathfrak{D}(\xi_{44}))_{41} & (\mathfrak{D}(\xi_{44}))_{42} & (\mathfrak{D}(\xi_{44}))_{43} & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Easy calculations give us that

$$-\lambda\xi_{ij}.Z = \begin{array}{c c} \begin{matrix} & & & j\text{-th column} & & \end{matrix} & \\ \begin{pmatrix} 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & ... & 0 & 0 & 0 & ... \\ \lambda\mathfrak{D}(\xi_{jj})_{j1} & ... & \lambda\mathfrak{D}(\xi_{jj})_{j,j-1} & 0 & \lambda\mathfrak{D}(\xi_{jj})_{j,j+1} & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

As we can see, the matrix Z satisfies

$$\mathfrak{D}(\lambda\xi_{ij}) = Z.\lambda\xi_{ij} - \lambda\xi_{ij}.Z + \lambda'\xi_{ij}$$

for all i, $j \in \mathbb{Z}^+$ and $\lambda \in R$. The claim also holds.

As a result, since Y can be written as a finite sum of matrix units ξ_{ij} for some i, j values and $\mathfrak{D}$ is an additive map, we obtain that

$$\mathfrak{D}(Y) = ZY - YZ + \mathfrak{D}'(Y) \ ,$$

where for every $i, j \in \mathbb{Z}^+$ and $\lambda \in R$, $\mathfrak{D}'(\lambda\xi_{ij}) = \lambda'\xi_{ij}$ for some $\lambda' \in R$. $\square$

One may ask why we started by investigating $\mathfrak{D}(Y)$, where $Y \in \mathcal{M}_\infty^{fin}(R)$. Let us explain the basic idea here. We first show that finitary matrices satisfy what we want to show, then we will extend this to whole ring $\mathcal{M}_{Cf}(R)$ ($\mathcal{T}_\infty(R)$) in a practical way as you will see. Our next job is to describe the derivations of $\mathcal{R}$ mapping $\lambda(\xi_{ij})$'s to $\lambda'\xi_{ij}$'s. Let's continue with the following remark that will help us in the proof of next proposition.

Remark 4.2.2. Suppose that $\mathfrak{D}$ is a derivation of $\mathcal{R}$ such that $\mathfrak{D}(\lambda\xi_{ij})$ is of the form $\lambda'\xi_{ij}$ for all $i, j \in \mathbb{Z}^+$. Then the maps

$$\begin{array}{rccc} \mathfrak{D}_i : & R & \rightarrow & R \\ & \lambda & \rightarrow & (\mathfrak{D}(\lambda\xi_{ii}))_{ii} \end{array}$$

are derivations of R for all $i \in \mathbb{Z}^+$.

Proof. Let $\lambda, \gamma \in R$ and $i \in \mathbb{Z}^+$. Then, using the additivity of $\mathfrak{D}$

$$\mathfrak{D}_i(\lambda + \gamma) = (\mathfrak{D}((\lambda + \gamma)\xi_{ii}))_{ii} = \mathfrak{D}(\lambda\xi_{ii})_{ii} + \mathfrak{D}(\gamma\xi_{ii})_{ii} = \mathfrak{D}_i(\lambda) + \mathfrak{D}_i(\gamma).$$

This means that $\mathfrak{D}_i$ is additive. Let's check whether or not $\mathfrak{D}_i$ satisfies Leibniz rule:

$$\begin{aligned} \mathfrak{D}_i(\lambda\gamma) &= (\mathfrak{D}(\lambda\gamma\xi_{ii}))_{ii} = (\mathfrak{D}(\lambda\xi_{ii}.\gamma\xi_{ii}))_{ii} \\ &= (\mathfrak{D}(\lambda\xi_{ii})\gamma\xi_{ii} + \lambda\xi_{ii}\mathfrak{D}(\gamma\xi_{ii}))_{ii} \\ &= \mathfrak{D}_i(\lambda)\gamma + \lambda\mathfrak{D}_i(\gamma) \end{aligned}$$

Hence, the result follows. ☐

Proposition 4.2.3. *Let $\mathfrak{D}$ be a derivation of $\mathcal{R}$ such that $\mathfrak{D}(\lambda\xi_{ij})$ is of the form $\lambda'\xi_{ij}$. Then for all finitary matrices $Y \in \mathcal{R}$*

$$\mathfrak{D}(Y) = \mathfrak{D}^{(1)}(Y) + \mathfrak{D}^{(2)}(Y),$$

where $\mathfrak{D}^{(1)}$ is an inner derivation and $\mathfrak{D}^{(2)}$ is an induced derivation.

Proof. Firstly, we construct the maps

$$\begin{array}{rccc} \mathfrak{D}_{ij} : & R & \rightarrow & R \\ & \lambda & \rightarrow & \big(\mathfrak{D}(\lambda\xi_{ij})\big)_{ij} \end{array}$$

Now, let $i, j \in \mathbb{Z}^+$ and $\lambda \in R$. We are given that $\mathfrak{D}(\lambda\xi_{ii})$ has nonzero entry only in its (i, i) position, which is $\mathfrak{D}_{ii}(\lambda)$. Thus, writing $\lambda\xi_{ij}$ as $\lambda\xi_{ii}.\xi_{ij}$, we can see that

$$\begin{aligned} \mathfrak{D}(\lambda\xi_{ij}) = \mathfrak{D}(\lambda\xi_{ii}.\xi_{ij}) &= \mathfrak{D}(\lambda\xi_{ii})\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ij}) \\ &= \mathfrak{D}_{ii}(\lambda)\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ij}). \end{aligned}$$

Of course, $\lambda\xi_{ij}$ can also be written as $\xi_{ij}.\lambda\xi_{jj}$. Therefore,

$$\begin{aligned}\mathfrak{D}(\lambda\xi_{ij}) = \mathfrak{D}(\xi_{ij}.\lambda\xi_{jj}) &= \mathfrak{D}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}(\lambda\xi_{jj}) \\ &= \mathfrak{D}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}_{jj}(\lambda).\end{aligned}$$

Combining above results, we find

$$\mathfrak{D}_{ii}(\lambda)\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ij}) = \mathfrak{D}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}_{jj}(\lambda).$$

In particular,

$$\Bigg(\mathfrak{D}_{ii}(\lambda)\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ij})\Bigg)_{ij} = \Bigg(\mathfrak{D}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}_{jj}(\lambda)\Bigg)_{ij}$$

Now, let's calculate (i, j)-th coordinate of those terms

- $\Bigg(\mathfrak{D}_{ii}(\lambda)\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}(\xi_{ij})\Bigg)_{ij} = \mathfrak{D}_{ii}(\lambda) + \lambda\mathfrak{D}(\xi_{ij})_{ij} = \mathfrak{D}_{ii}(\lambda) + \lambda\mathfrak{D}_{ij}(1),$
- $\Bigg(\mathfrak{D}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}_{jj}(\lambda)\Bigg)_{ij} = \mathfrak{D}_{jj}(\lambda) + \mathfrak{D}(\xi_{ij})_{ij}\lambda = \mathfrak{D}_{jj}(\lambda) + \mathfrak{D}_{ij}(1)\lambda.$

The right-hand side of the above two equations yields that for any $i, j \in \mathbb{Z}^+$

$$\mathfrak{D}_{jj}(\lambda) = \mathfrak{D}_{ii}(\lambda) + \lambda\mathfrak{D}_{ij}(1) - \mathfrak{D}_{ij}(1)\lambda. \qquad (*)$$

Now, we consider the diagonal matrix $\Delta = (\delta_i)_{i\in\mathbb{Z}^+}$ with $\delta_i = -\mathfrak{D}_{1i}(1)$:

$$\Delta = \sum_{i=2}^{\infty} -\mathfrak{D}_{1i}(1)\xi_{ii} = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & -\mathfrak{D}_{12}(1) & 0 & 0 & \dots \\ 0 & 0 & -\mathfrak{D}_{13}(1) & 0 & \dots \\ 0 & 0 & 0 & -\mathfrak{D}_{14}(1) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Let $\mathfrak{D}'$ be an inner derivation implemented by Δ, i.e., $\mathfrak{D}'(Y) = \Delta Y - Y\Delta$. It is obvious that $\mathfrak{D}'' := \mathfrak{D} - \mathfrak{D}'$ is a derivation of $\mathcal{R}$ as well. Now, let $i \in \mathbb{Z}^+$ and $\lambda \in R$. Then we have

$$\mathfrak{D}'_{ii}(\lambda) = (\mathfrak{D}'(\lambda\xi_{ii}))_{ii} = (\Delta.\lambda\xi_{ii} - \lambda\xi_{ii}.\Delta)_{ii} = (\Delta.\lambda\xi_{ii})_{ii} - (\lambda\xi_{ii}.\Delta)_{ii}.$$

The matrix $\Delta.\lambda\xi_{ii}$ can be seen as below

$$\Delta.\lambda\xi_{ii} = \begin{array}{cc} \begin{matrix} & & & i\text{-th column} & & \end{matrix} & \\ \begin{pmatrix} 0 & \dots & 0 & \delta_{1i}\lambda & 0 & \dots \\ 0 & \dots & 0 & \delta_{2i}\lambda & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \delta_{ii}\lambda & 0 & \dots \\ 0 & \dots & 0 & \delta_{i+1,i}\lambda & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

Since $\delta_{ni} = 0$ for all $n \neq i$ and $\delta_{ii} = -\mathfrak{D}_{1i}(1)$, it must be

$$\Delta.\lambda\xi_{ii} = \begin{array}{cc} \begin{matrix} & & & i\text{-th column} & & \end{matrix} & \\ \begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & -\mathfrak{D}_{1i}(1)\lambda & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

That is, the matrix $\Delta.\lambda\xi_{ii}$ has non-zero entry only in its (i, i) position, which is $-\mathfrak{D}_{1i}(1)\lambda$. Now, let's observe what the matrix $\lambda\xi_{ii}.\Delta$ looks like:

$$\lambda\xi_{ii}.\Delta = \begin{array}{cc} \begin{matrix} & & & i\text{-th column} & & \end{matrix} & \\ \begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -\lambda\mathfrak{D}_{1i}(1) & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{matrix} \\ \\ \\ i\text{-th row} \\ \\ \\ \end{matrix} \end{array}$$

As can be seen above, the matrix $\lambda\xi_{ii}.\Delta$ has non-zero entry only in its (i, i) position, which is $-\lambda\mathfrak{D}_{1i}(1)$. Thus, we conclude that $\mathfrak{D}'_{ii}(\lambda) = -\mathfrak{D}_{1i}(1)\lambda + \lambda\mathfrak{D}_{1i}(1)$. Using this observation and equation $(*)$, we obtain

$$\begin{aligned}\mathfrak{D}''_{ii}(\lambda) &= \mathfrak{D}_{ii}(\lambda) - \mathfrak{D}'_{ii}(\lambda)\\ &= \mathfrak{D}_{ii}(\lambda) - (-\mathfrak{D}_{1i}(1)\lambda + \lambda\mathfrak{D}_{1i}(1))\\ &= \mathfrak{D}_{ii}(\lambda) + \mathfrak{D}_{1i}(1)\lambda - \lambda\mathfrak{D}_{1i}(1)\\ &= \mathfrak{D}_{11}(\lambda)\end{aligned}$$

Since i is chosen randomly, one can immediately deduce that $\mathfrak{D}''_{ii}(\lambda) = \mathfrak{D}''_{jj}(\lambda)$ for all i, j and $\lambda \in R$. This investigation will play an important role in the rest of our work. Also, it should also be kept in mind that $\mathfrak{D}''_{ii}$ is a derivation of R by the above remark.

Now, consider the map $\mathfrak{D}''' : \mathcal{R} \rightarrow \mathcal{R}$ which is defined as follows:

$$\mathfrak{D}'''(Y) = \sum_{i,j} \mathfrak{D}''_{11}(y_{ij})\xi_{ij} = \begin{pmatrix} \mathfrak{D}''_{11}(y_{11}) & \mathfrak{D}''_{11}(y_{12}) & \mathfrak{D}''_{11}(y_{13}) & \dots \\ \mathfrak{D}''_{11}(y_{21}) & \mathfrak{D}''_{11}(y_{22}) & \mathfrak{D}''_{11}(y_{23}) & \dots \\ \mathfrak{D}''_{11}(y_{31}) & \mathfrak{D}''_{11}(y_{32}) & \mathfrak{D}''_{11}(y_{32}) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

By the construction of $\mathfrak{D}'''$, $\big(\mathfrak{D}'''(Y)\big)_{ij} = \mathfrak{D}''_{11}(y_{ij})$ for all i, j, where $\mathfrak{D}''_{11}$ is a derivation of R. That is, $\mathfrak{D}'''$ is an induced derivation. Thus, $\mathfrak{D}^{\mathrm{IV}} = \mathfrak{D}'' - \mathfrak{D}'''$ also defines a derivation of $\mathcal{R}$. Now, I would like to summarize what we have done so far. We constructed the derivations $\mathfrak{D}'$, $\mathfrak{D}''$, $\mathfrak{D}'''$ and $\mathfrak{D}^{\mathrm{IV}}$ of $\mathcal{R}$ such that

$$\mathfrak{D}(Y) = \mathfrak{D}'(Y) + \mathfrak{D}'''(Y) + \mathfrak{D}^{\mathrm{IV}}(Y) \text{ for all } Y \in \mathcal{M}^{fin}_{\infty}(R).$$

We already know that $\mathfrak{D}'$ is an inner derivation and $\mathfrak{D}'''$ is induced by $\mathfrak{D}''$. As you see, what we need to do is to verify that $\mathfrak{D}^{\mathrm{IV}}(Y)$ is an inner derivation to achieve our goal. So let's focus on $\mathfrak{D}^{\mathrm{IV}}$. Take $\lambda \in R$ and $i \in \mathbb{Z}^+$ arbitrarily. Then $\mathfrak{D}'''_{ii}(\lambda) = (\mathfrak{D}'''(\lambda\xi_{ii}))_{ii} = \mathfrak{D}''_{11}(\lambda)$ by the definition of $\mathfrak{D}'''$. We also determined that $\mathfrak{D}''{}_{11}(\lambda) = \mathfrak{D}''{}_{ii}(\lambda)$. Therefore,

$$\mathfrak{D}^{\mathrm{IV}}_{ii}(\lambda) = \mathfrak{D}''_{ii}(\lambda) - \mathfrak{D}'''_{ii}(\lambda) = \mathfrak{D}''_{ii}(\lambda) - \mathfrak{D}''_{11}(\lambda) = 0.$$

Since $i \in \mathbb{Z}^+$ and $\lambda \in R$ are randomly chosen, we have $\mathfrak{D}^{\mathrm{IV}}_{ii}(\lambda) = 0$ for all $i \in \mathbb{Z}^+$ and $\lambda \in R$. Now, consider $i, j \in \mathbb{Z}^+$ and $\lambda \in R$.

1. Clearly, we can write $\lambda\xi_{ij}$ as $\lambda\xi_{ii}.\xi_{ij}$. Then we get

$$\begin{aligned}\mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{ij}) &= \mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{ii}.\xi_{ij})\\ &= \mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{ii}).\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\\ &= \mathfrak{D}^{\mathrm{IV}}_{ii}(\lambda)\xi_{ij} + \lambda\xi_{ii}\mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\end{aligned}$$

2. Similarly, $\lambda\xi_{ij}$ can be seen as $\xi_{ij}.\lambda\xi_{jj}$. Thus, we also have the following

$$\begin{aligned}\mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{ij}) &= \mathfrak{D}^{\mathrm{IV}}(\xi_{ij}.\lambda\xi_{jj})\\ &= \mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{jj})\\ &= \mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\lambda\xi_{jj} + \xi_{ij}\mathfrak{D}^{\mathrm{IV}}_{jj}(\lambda)\end{aligned}$$

Since $\mathfrak{D}^{\mathrm{IV}}_{ii}(\lambda) = 0$ for all i and $\lambda \in R$, using (1) and (2) one can deduce that

$$\left(\mathfrak{D}^{\mathrm{IV}}(\lambda\xi_{ij})\right)_{ij} = \mathfrak{D}^{\mathrm{IV}}_{ij}(\lambda) = \lambda\mathfrak{D}^{\mathrm{IV}}_{ij}(1) = \mathfrak{D}^{\mathrm{IV}}_{ij}(1)\lambda.$$

Accordingly, $\mathfrak{D}^{\mathrm{IV}}_{ij}(1)$ is in the center of R for all pairs i, j.

Now, let $i,j,k \in \mathbb{Z}^+$. Using $\mathfrak{D}^{\mathrm{IV}}(\xi_{ik}) = \mathfrak{D}^{\mathrm{IV}}(\xi_{ij}.\xi_{jk}) = \mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\xi_{jk} + \xi_{ij}\mathfrak{D}^{\mathrm{IV}}(\xi_{jk})$, we obtain

$$\underbrace{\left(\mathfrak{D}^{\mathrm{IV}}(\xi_{ik})\right)_{ik} = \left(\mathfrak{D}^{\mathrm{IV}}(\xi_{ij})\xi_{jk} + \xi_{ij}\mathfrak{D}^{\mathrm{IV}}(\xi_{jk})\right)_{ik}}_{\mathfrak{D}^{\mathrm{IV}}_{ik}(1) = \mathfrak{D}^{\mathrm{IV}}_{ij}(1) + \mathfrak{D}^{\mathrm{IV}}_{jk}(1)}$$

Therefore, $\mathfrak{D}^{\mathrm{IV}}_{ij}$ are determined by $\mathfrak{D}^{\mathrm{IV}}_{i,i+1}(1)$ in the sense that

$$\mathfrak{D}^{\mathrm{IV}}_{i,i+k}(\lambda) = \lambda\mathfrak{D}^{\mathrm{IV}}_{i,i+k}(1) = \lambda\left(\sum_{n=0}^{k-1}\mathfrak{D}^{\mathrm{IV}}_{i+n,i+n+1}(1)\right) \ \text{ for } \ k \geq 1.$$

Of course, in the case when $\mathcal{R} = \mathcal{M}_{Cf}(R)$ we have nonzero entries below the main diagonal and so we should take this observation one step further. Using the fact that $\mathfrak{D}^{\mathrm{IV}}_{ii}(\lambda) = 0$ for all i, we can write $\mathfrak{D}^{\mathrm{IV}}_{i,i-k}(\lambda)$ as below

$$\mathfrak{D}^{\mathrm{IV}}_{i,i-k}(\lambda) = -\lambda\mathfrak{D}^{\mathrm{IV}}_{i-k,i}(1) = -\lambda\left(\sum_{n=0}^{k-1}\mathfrak{D}^{\mathrm{IV}}_{i-k+n,i-k+n+1}(1)\right) \ \text{ for } \ k \geq 1.$$

Consider now the diagonal matrix $D = (d_i)_{i\in\mathbb{Z}^+}$ which is defined as

$$D = \sum_{i=1}^{\infty} d_i\xi_{ii} = \begin{pmatrix} \mathfrak{D}^{\mathrm{IV}}_{12}(1) & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & -\mathfrak{D}^{\mathrm{IV}}_{23}(1) & 0 & \dots \\ 0 & 0 & 0 & -\mathfrak{D}^{\mathrm{IV}}_{23}(1) - \mathfrak{D}^{\mathrm{IV}}_{34}(1) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $d_1 = \mathfrak{D}^{\mathrm{IV}}_{12}(1),\ d_2 = 0,\ d_i = d_{i-1} - \mathfrak{D}^{\mathrm{IV}}_{i-1,i}(1)$ for $i \geq 3$.
One can check that

$$\mathfrak{D}^{\mathrm{IV}}(Y) = DY - YD \text{ for all } Y \in \mathcal{M}^{fin}_{\infty}(R),$$

showing that $\mathfrak{D}^{\mathrm{IV}}$ is an inner derivation. This is what we wanted to get. By the way, the above check may seem difficult at first glance but it is not. We got enough information about $\mathfrak{D}^{\mathrm{IV}}$ that make these calculations easier:

- $\mathfrak{D}^{\mathrm{IV}}_{ij}(1)$ is in the center of R for all pairs i, j and
- $\mathfrak{D}^{\mathrm{IV}}_{i,i+k}(\lambda) = \lambda\mathfrak{D}^{\mathrm{IV}}_{i,i+k}(1) = \lambda\Big(\sum_{n=0}^{k-1}\mathfrak{D}^{\mathrm{IV}}_{i+n,i+n+1}(1)\Big)$ for $k \geq 1$
- $\mathfrak{D}^{\mathrm{IV}}_{i,i-k}(\lambda) = -\lambda\mathfrak{D}^{\mathrm{IV}}_{i-k,i}(1) = -\lambda\Big(\sum_{n=0}^{k-1}\mathfrak{D}^{\mathrm{IV}}_{i-k+n,i-k+n+1}(1)\Big)$ for $k \geq 1$

Consequently, if Y is finitary , then we have

$$\mathfrak{D}(Y) = \mathfrak{D}^{'}(Y) + \mathfrak{D}^{'''}(Y) + \mathfrak{D}^{\mathrm{IV}}(Y) = \mathfrak{Z}Y - Y\mathfrak{Z} + \mathfrak{D}^{'''}(Y),$$

where $\mathfrak{Z} = \Delta + D$ and $\mathfrak{D}^{'''}$ is an induced derivation. Hence, we are done. □

Up to this point, we observed the properties of $\mathfrak{D}(Y)$ when $Y \in \mathcal{R}$ is finitary. In the following proposition, we will obverse what happens to $\mathfrak{D}(Y)$ if $Y \notin \mathcal{M}^{fin}_{\infty}(R)$.

Proposition 4.2.4. *Let $\mathfrak{D}$ be a derivation of $\mathcal{R}$ such that*

$$\mathfrak{D}(Y) = \mathfrak{D}^{'}(Y) + \mathfrak{D}^{''}(Y) \textit{ for all finitary } Y \in \mathcal{R},$$

where $\mathfrak{D}^{'}$ is an inner derivation and $\mathfrak{D}^{''}$ is an induced derivation. Then we have

$$\mathfrak{D}(Y) = \mathfrak{D}^{'}(Y) + \mathfrak{D}^{''}(Y) \textit{ for all } Y \in \mathcal{R}.$$

Proof. Assume that $\mathfrak{D}'$ is implemented by U and $\mathfrak{D}''$ is induced by $\mathfrak{C}$, which is a derivation of R. Of course, we can make such an assumption since we are given that $\mathfrak{D}'$ is an inner derivation and $\mathfrak{D}''$ is an induced derivation. If we write 1 as 1.1, then

$$\begin{aligned}\mathfrak{C}(1) &= \mathfrak{C}(1).1 + 1\mathfrak{C}(1)\\ \mathfrak{C}(1) &= \mathfrak{C}(1) + \mathfrak{C}(1)\\ \mathfrak{C}(1) &= 0\end{aligned}$$

(In a similar way, we have $\mathfrak{D}(I_\infty) = O$.) Since $\mathfrak{C}(1) = 0$ and $\mathfrak{D}''$ is induced by $\mathfrak{C}$, it must be

$$\mathfrak{D}''\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\right) = O$$

Dividing U into block matrices as below

$$\left(\begin{array}{c|c} U_1 & U_2 \\ \hline U_3 & U_4 \end{array}\right), \quad \text{where } \; U_1 \in \mathcal{M}_n(R),$$

we obtain that

$$\begin{aligned}
\mathfrak{D}\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\right) &= \mathfrak{D}'\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\right) \\
&= \left(\begin{array}{c|c} U_1 & U_2 \\ \hline U_3 & U_4 \end{array}\right) \cdot \left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right) - \left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right) \cdot \left(\begin{array}{c|c} U_1 & U_2 \\ \hline U_3 & U_4 \end{array}\right) + O \\
&= \left(\begin{array}{c|c} O & -U_2 \\ \hline U_3 & O \end{array}\right)
\end{aligned}$$

Also, from

$$O = \mathfrak{D}(I_\infty) = \mathfrak{D}\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\right) + \mathfrak{D}\left(\left(\begin{array}{c|c} O & O \\ \hline O & I_\infty \end{array}\right)\right),$$

it follows that

$$\mathfrak{D}\left(\left(\begin{array}{c|c} O & O \\ \hline O & I_\infty \end{array}\right)\right) = \left(\begin{array}{c|c} O & U_2 \\ \hline -U_3 & O \end{array}\right).$$

Now, let $Y \in \mathcal{R}$. We divide it as follows

$$Y = \left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & Y_4 \end{array}\right), \quad \text{where } \; Y_1 \in \mathcal{M}_n(R).$$

From here on we turn our attention to the block matrices $Y_1 \in \mathcal{M}_n(R), Y_2 \in \mathcal{M}_{n\times\infty}(R)$,

$Y_3 \in \mathcal{M}^{fin}_{\infty\times n}(R)$ and $Y_4 \in \mathcal{M}_{Cf}(R)$. Let's start with Y_2. We can write that

$$\begin{aligned}
\left(\begin{array}{c|c} V_1' & V_2' \\ \hline V_3' & V_4' \end{array}\right) &= \mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) = \mathfrak{D}\left(\left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\cdot\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) \\
&= \left(\begin{array}{c|c} O & -U_2 \\ \hline U_3 & O \end{array}\right)\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right) + \left(\begin{array}{c|c} I_n & O \\ \hline O & O \end{array}\right)\left(\begin{array}{c|c} V_1' & V_2' \\ \hline V_3' & V_4' \end{array}\right) \\
&= \left(\begin{array}{c|c} O & O \\ \hline O & U_3Y_2 \end{array}\right) + \left(\begin{array}{c|c} I_nV_1' & I_nV_2' \\ \hline O & O \end{array}\right) \\
&= \left(\begin{array}{c|c} V_1' & V_2' \\ \hline O & U_3Y_2 \end{array}\right).
\end{aligned}$$

On the other side, we have

$$\mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) = \mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\cdot\left(\begin{array}{c|c} O & O \\ \hline O & I_\infty \end{array}\right)\right) = \left(\begin{array}{c|c} -Y_2U_3 & V_2' \\ \hline O & V_4' \end{array}\right).$$

Accordingly, it must be

$$\begin{aligned}
\left(\begin{array}{c|c} -Y_2U_3 & V_2' \\ \hline O & U_3Y_2 \end{array}\right) &= \mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) \\
&= U\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right) - \left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)U + \left(\begin{array}{c|c} O & V_2'' \\ \hline O & O \end{array}\right).
\end{aligned}$$

It follows from the above that

$$\begin{aligned}
\left(\mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right)\right)_{ij} &= \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right)\right)_{ij} + \left((\mathfrak{D}-\mathfrak{D}')\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right)_{ij} \\
&= \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right)\right)_{ij} + ((\mathfrak{D}-\mathfrak{D}')((Y_2)_{ij}\xi_{ij}))_{ij} \\
&= \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right)\right)_{ij} + \mathfrak{C}((Y_2)_{ij})
\end{aligned}$$

for all i, j. As a result of that, for all $Y_2 \in \mathcal{M}_{n\times\infty}(R)$, we obtain

$$\mathfrak{D}\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) = \mathfrak{D}'\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} O & Y_2 \\ \hline O & O \end{array}\right)\right).$$

Now, we consider $Y_1 \in \mathcal{M}_n(R), Y_3 \in \mathcal{M}^{fin}_{\infty\times n}(R)$. From our hypothesis, we already have

$$\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & O \\ \hline Y_3 & O \end{array}\right)\right) = \mathfrak{D}'\left(\left(\begin{array}{c|c} Y_1 & O \\ \hline Y_3 & O \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} Y_1 & O \\ \hline Y_3 & O \end{array}\right)\right).$$

The additivity of $\mathfrak{D}$ implies

$$\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right) = \mathfrak{D}'\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right).$$

Finally, let's observe $Y_4 \in \mathcal{M}_{Cf}(R)$. From

$$\begin{aligned}
\left(\begin{array}{c|c} W_1' & W_2' \\ \hline W_3' & W_4' \end{array}\right) &= \mathfrak{D}\left(\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right) = \mathfrak{D}\left(\left(\begin{array}{c|c} O & O \\ \hline O & I_\infty \end{array}\right) \cdot \left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right) \\
&= \left(\begin{array}{c|c} O & U_2 \\ \hline -U_3 & O \end{array}\right)\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right) + \left(\begin{array}{c|c} O & O \\ \hline O & I_\infty \end{array}\right)\left(\begin{array}{c|c} W_1' & W_2' \\ \hline W_3' & W_4' \end{array}\right) \\
&= \left(\begin{array}{c|c} O & U_2Y_4 \\ \hline O & O \end{array}\right) + \left(\begin{array}{c|c} O & O \\ \hline W_3' & W_4' \end{array}\right) \\
&= \left(\begin{array}{c|c} O & U_2Y_4 \\ \hline W_3' & W_4' \end{array}\right),
\end{aligned}$$

it follows that $W_1' = O$. Thus, for any i, j, we get

$$\begin{aligned}
\left(\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & Y_4 \end{array}\right)\right)\right)_{ij} &= \left(\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right)\right)_{ij} + \left(\mathfrak{D}\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right)_{ij} \\
&= \left(\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right)\right)_{ij} + O.
\end{aligned}$$

Fortunately, it is time to get the result. Let $i, j \in \mathbb{Z}^+$. We just deduced that

$$\left(\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right)\right)_{ij} = \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & O \end{array}\right)\right)\right)_{ij}.$$

It is also easy to see that

$$\left(\mathfrak{D}\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right)_{ij} = O = \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} O & O \\ \hline O & Y_4 \end{array}\right)\right)\right)_{ij}$$

Thanks to these observations, we get

$$\left(\mathfrak{D}\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & Y_4 \end{array}\right)\right)\right)_{ij} = \left(\mathfrak{D}'\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & Y_4 \end{array}\right)\right) + \mathfrak{D}''\left(\left(\begin{array}{c|c} Y_1 & Y_2 \\ \hline Y_3 & Y_4 \end{array}\right)\right)\right)_{ij},$$

showing that $\mathfrak{D}(Y) = \mathfrak{D}'(Y) + \mathfrak{D}''(Y)$ for all $Y \in \mathcal{R}$. □

With these lemmas in hand we are now ready to prove our main results.

4.3 Proofs of the Main Results

Proof. Let $\mathfrak{D}$ be a derivation of $\mathcal{R}$. From Proposition 4.2.1, it immediately follows that $\mathfrak{D}(Y) = \mathfrak{D}_Z(Y) + \mathfrak{D}'(Y)$ for all $Y \in \mathcal{M}_\infty^{fin} \cap \mathcal{R}$, where for all $i, j \in \mathbb{Z}^+$ and $\lambda \in R$ we have $\mathfrak{D}'(\lambda\xi_{i,j}) = \lambda'\xi_{i,j}$ for some $\lambda' \in R$. Moreover, from Proposition 4.2.3, it must be $\mathfrak{D}'(Y) = \mathfrak{D}_{\mathfrak{Z}}(Y) + \mathfrak{D}_{\mathfrak{C}(R)}(Y)$ for all $Y \in \mathcal{M}_\infty^{fin} \cap \mathcal{R}$, where $\mathfrak{C}$ is a derivation of R. Therefore, $\mathfrak{D}(Y) = \mathfrak{D}_X(Y) + \mathfrak{D}_{\mathfrak{C}(R)}(Y)$ for all $Y \in \mathcal{M}_\infty^{fin} \cap \mathcal{R}$, where $X = Z + \mathfrak{Z}$. Finally, Proposition 4.2.4 gives the desired result

$$\mathfrak{D}(Y) = \mathfrak{D}_X(Y) + \mathfrak{D}_{\mathfrak{C}(R)}(Y) \text{ for all } Y \in \mathcal{R}.$$

Of course, we also need to verify that $X \in \mathcal{R}$ to complete the proof. In the case of $\mathcal{R} = \mathcal{M}_{Cf}(R)$, the result is trivial since the matrices $Z, \mathfrak{Z}$ are both in $\mathcal{M}_{Cf}(R)$, and so their sum must be column finite as well. Now, consider the case of $\mathcal{R} = \mathcal{T}_\infty(R)$. Assume that $X \notin \mathcal{T}_\infty(R)$. Then there exist $i, j \in \mathbb{Z}^+$ with $i > j$ such that $x_{ij} \neq 0$. In such a case,

$$\begin{aligned}
(\mathfrak{D}(\xi_{jj}))_{ij} &= (\mathfrak{D}_X(\xi_{jj}))_{ij} + \big(\mathfrak{D}_{\mathfrak{C}(R)}(\xi_{jj})\big)_{ij} \\
&= (\mathfrak{D}_X(\xi_{jj}))_{ij} + \mathfrak{C}((\xi_{jj})_{ij}) \\
&= (\mathfrak{D}_X(\xi_{jj}))_{ij} + 0 \\
&= (\mathfrak{D}_X(\xi_{jj}))_{ij} = x_{ij} \neq 0.
\end{aligned}$$

This means that $\mathfrak{D}(\xi_{jj}) \notin \mathcal{T}_\infty(R)$, which is a contradiction. Hence, we get $X \in \mathcal{T}_\infty(R)$. □

Chapter 5

Derivations of Infinite Niltriangular Lie Matrix Algebras

Let R be a commutative ring with identity. We have already observed that the set of all niltriangular $n \times n$ matrices over R whose entries are all zeros on and below the main diagonal forms an R-algebra shown by $\mathcal{N}T_n(R)$. As discussed earlier, many mathematicians have been interested in derivations and automorphisms of this algebra. Let's take a quick look at a few studies on this subject to understand how important the work to be done in this chapter is. The problem of describing automorphisms of the algebra $\mathcal{N}T_n(R)$ was investigated by Dubish and Perlis when R is a field (see $[8]$). Later, Cao and Wang also addressed this algebra in their work. What they did was describe the automorphism group of $\mathcal{N}T_n(R)$ (see [5]). Moreover, Wang, Ou and Yao gave derivations of $\mathcal{N}T_n(R)$ as a Lie algebra (see [17]).

We now consider the set $\mathcal{T}_\infty(R)$ of all infinite $\mathbb{Z}^+ \times \mathbb{Z}^+$ upper triangular matrices over R. It is no wonder that $\mathcal{T}_\infty(R)$ constructs an **associative algebra** with respect to usual matrix addition, scalar multiplication and matrix multiplication. This algebra $\mathcal{T}_\infty(R)$ was investigated by Sushkevich (see [21]) when $R = \mathbb{C}$ is a field of complex numbers. As you may recall, we observed that an associative algebra always forms a Lie algebra. In particular, our algebra $\mathcal{T}_\infty(R)$ is a Lie algebra with $[X, Y] = XY - YX$. Throughout this chapter, we will concentrate on its Lie subalgebra of strictly upper triangular matrices, which is denoted by $\mathcal{N}T_\infty(R)$. Now, let us talk about what we will do. All derivations of $\mathcal{N}T_\infty(R)$ were described by W.Hołubowski, I.Kashuba and S.Zurek in $[11]$. They proved that any derivation of $\mathcal{N}T_\infty(R)$ can be expressed as the sum of an inner derivation and a diagonal derivation. The purpose of this chapter is to discuss this article in detail.

5.1 Notations, Some Definitions and Basic Facts

We have already discussed what matrix units are, but it is better to talk about them once again since they play a vital role in this chapter.

Definition 5.1.1. An infinite matrix ξ_{ij} whose only nonzero entry is 1 in the (i, j)-th entry is called matrix unit.

Regarding these matrices, we have the following fact than can be proved easily

$$[\xi_{ij}, \xi_{kl}] = \delta_{jk}\xi_{il} - \delta_{li}\xi_{kj},$$

where δ is the Kronecker delta. As you know, it was mentioned in the second chapter that matrix units form a basis for the vector space $\mathcal{M}_n(F)$ of all $n \times n$ matrices over a field F. Well, do you think we can say the same thing for $\mathcal{N}T_\infty(R)$? (Note that our Lie algebra $\mathcal{N}T_\infty(R)$ is an uncountable dimensional vector space in the case of R being a field.) I can just hear you saying, "Of course not." Let us explain why this cannot happen. Consider the set $\left\{\xi_{ij} \mid 1 \leq i < j\right\}$, which consists of all matrix units in $\mathcal{N}T_\infty(R)$. Of course, we have no doubt that this set is linearly independent. On the other hand, one can easily find a matrix $X \in \mathcal{N}T_\infty(R)$ which can not be written as a finite linear combination matrix units. As an example, consider

$$X = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & \dots \\ 0 & 0 & 1 & 1 & 1 & 1 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \in \mathcal{N}T_\infty(R).$$

Nevertheless, the fact that matrix units do not form a basis of $\mathcal{N}T_\infty(R)$ does not prevent us from writing a matrix $X \in \mathcal{N}T_\infty(R)$ as

$$X = \sum_{1 \leq i < j} x_{ij}\xi_{ij}.$$

As you may recall, we examined the lower central series of the Lie algebra $\mathcal{N}_5(F)$ of strictly upper triangular 5×5 matrices over a field F. Now, we move the discussion we had in the second chapter here. Let's see what similarities or differences are there. Denote the lower central series of $\mathcal{N}T_\infty(R)$ by

$$\Omega_1 = \mathcal{N}T_\infty(R)$$

$$\Omega_2 = [\Omega_1, \mathcal{N}T_\infty(R)]$$

$$\vdots$$

$$\Omega_n = [\Omega_{n-1}, \mathcal{N}T_\infty(R)].$$

One can easily see that if $X, Y \in \mathcal{N}T_\infty(R)$, then

$$XY - YX = Z = \begin{pmatrix} 0 & 0 & z_{13} & z_{14} & z_{15} & \cdots \\ 0 & 0 & 0 & z_{24} & z_{25} & \cdots \\ 0 & 0 & 0 & 0 & z_{35} & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Since $\Omega_2 = [\mathcal{N}T_\infty(R), \mathcal{N}T_\infty(R)] = [\Omega_1, \Omega_1]$ is defined to be the linear span of elements of the form $[X, Y] = XY - YX$, where $X, Y \in \mathcal{N}T_\infty(R)$, we obtain

$$\Omega_2 = [\Omega_1, \Omega_1] = \left\{ Z \in \mathcal{N}T_\infty(R) \mid Z = \begin{pmatrix} 0 & 0 & z_{13} & z_{14} & z_{15} & \dots \\ 0 & 0 & 0 & z_{24} & z_{25} & \dots \\ 0 & 0 & 0 & 0 & z_{35} & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \right\} \subset \Omega_1.$$

You may have noticed that $\Omega_2 \subset \mathcal{N}T_\infty(R)$ consists of matrices whose entries one above the main diagonal are all zeros. As we will see soon, a symmetric discussion holds for all Ω_n. We now focus on Ω_3. Let $X \in \Omega_2$ and $Y \in \mathcal{N}T_\infty(R)$. Then we have

$$XY - YX = Z = \begin{pmatrix} 0 & 0 & 0 & z_{14} & z_{15} & \cdots \\ 0 & 0 & 0 & 0 & z_{25} & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

That is why

$$\Omega_3 = [\Omega_2, \Omega_1] = \left\{ Z \in \mathcal{NT}_\infty(R) \mid Z = \begin{pmatrix} 0 & 0 & 0 & z_{14} & z_{15} & \dots \\ 0 & 0 & 0 & 0 & z_{25} & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right\} \subset \Omega_2.$$

Now, we shall generalize the results we obtained above. Let $X \in \Omega_{n-1}$ and $Y \in \mathcal{NT}_\infty(R)$. Then easy calculations give us

$$XY - YX = Z = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & z_{1,n+1} & z_{1,n+2} & z_{1,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & z_{2,n+2} & z_{2,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & z_{3,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Thus, it must be

$$\Omega_n = \left\{ Z \in \mathcal{NT}_\infty(R) \mid Z = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & z_{1,n+1} & z_{1,n+2} & z_{1,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & z_{2,n+2} & z_{2,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & z_{3,n+3} & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right\}.$$

Thereby, we proved the following result:

Proposition 5.1.2. *For any $n \geq 1$,*

$$\Omega_n = \left\{ X = (x_{ij})_{i,j \in \mathbb{Z}^+} \in \mathcal{NT}_\infty(R) \mid x_{ij} = 0 \textit{ if } j < i + n \right\}.$$

Corollary 5.1.3. $\displaystyle\bigcap_{n=1}^{\infty} \Omega_n = \{O\}.$

Proof. Clearly, $O \in \Omega_n$ for all $n \geq 1$. So $O \subseteq \bigcap_{n=1}^{\infty} \Omega_n$. For the other inclusion, assume that $\bigcap_{n=1}^{\infty} \Omega_n \neq \{0\}$. Then, it follows that there exist a matrix $O \neq X = (x_{ij})_{i,j \in \mathbb{Z}^+} \in \bigcap_{n=1}^{\infty} \Omega_n$. This means that $x_{ij} \neq 0$ for some $i, j \in \mathbb{Z}^+$. For this $i, j \in \mathbb{Z}^+$, we can immediately find $n \in \mathbb{Z}^+$ such that $j < i + n$. Of course, in such a case $X \notin \Omega_n$ by Proposition 5.1.2. Hence, it must be $X \notin \bigcap_{n=1}^{\infty} \Omega_n$, and this contradicts our assumption. Thereby, the result follows. $\square$

There is one more thing we want to talk about the lower central series of $\mathcal{N}T_\infty(R)$.

Proposition 5.1.4. $[\Omega_n, \Omega_m] = \Omega_{n+m}$ *for all* $n, m \geq 0$.

Proof. Let $X \in \Omega_n$, $Y \in \Omega_m$ and $Z = (z_{ij})_{i,j\in\mathbb{Z}^+} = XY$, where $z_{ij} = \sum_{k=1}^{\infty} x_{ik}y_{kj}$. Now, assume that $j < i+n+m$. If $x_{ik} = 0$, then $x_{ik}y_{kj} = 0$. Otherwise, it must be $k \geq i+n$ because $X \in \Omega_n$, and this forces j to be less than $k+m$. Of course, this means that $y_{kj} = 0$ since $Y \in \Omega_m$. Consequently, for all $j < i+n+m$

$$z_{ij} = \sum_{k=1}^{\infty} x_{ik}y_{kj} = \sum_{k=1}^{\infty} 0 = 0.$$

Hence, $Z = (z_{ij})_{i,j\in\mathbb{Z}^+} = XY \in \Omega_{n+m}$ by above proposition. Accordingly, we find that $[\Omega_n, \Omega_m] \subseteq \Omega_{n+m}$.

For the other inclusion, we shall show that $\Omega_{n+m} = [\Omega_n, Y]$, where

$$Y = (y_{ij})_{i,j\in\mathbb{Z}^+} = \begin{cases} y_{ij} = 1 & \text{if } i = m+i \\ 0 & \textit{otherwise.} \end{cases}$$

Let us represent the matrix Y as the following rectangular array to have a better understanding of what our process is:

$$Y = \begin{pmatrix} 0 & \dots & 0 & y_{1,m+1} & 0 & 0 & 0 & \dots \\ 0 & \dots & 0 & 0 & y_{2,m+2} & 0 & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & y_{3,m+3} & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & 0 & y_{4,m+4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ with } \ y_{i,m+i} = 1 \text{ for all } i.$$

As can be seen clearly, $Y \in \Omega_m$. Now, let's take an arbitrary matrix $U \in \Omega_{n+m}$. In order to achieve our goal, it is enough to show that equation $XY - YX = U$ has a solution $X \in \Omega_n$. Evaluating (i, j)-th entry of the expression on the left, we get that

$$XY - YX = U \Rightarrow (XY)_{ij} - (YX)_{ij} = U_{ij} \Rightarrow \sum_{k=1}^{\infty} x_{ik}y_{kj} - \sum_{k=1}^{\infty} y_{ik}x_{kj} = U_{ij}.$$

If $j > m$, then $\sum_{k=1}^{\infty} x_{ik}y_{kj} = x_{i,j-m}$ as $y_{i,m+i} = 1$ for $i \geq 1$. Otherwise, $\sum_{k=1}^{\infty} x_{ik}y_{kj} = 0$. Moreover, $\sum_{k=1}^{\infty} y_{ik}x_{kj} = x_{i+m,j}$ due to construction of Y. So in other words, we obtain the following system of equations

- If $j > m$,

$$\begin{cases} x_{i,j-m} - x_{i+m,j} = u_{ij} & 1 \leq i \leq \infty. \end{cases}$$

- If $j \leq m$,

$$\begin{cases} -x_{i+m,j} = u_{ij} & 1 \leq i \leq \infty. \end{cases}$$

In view of the above observations, we now consider the matrix X, which is defined as below

$$X = (x_{ij})_{i,j \in \mathbb{Z}^+} = \begin{cases} x_{ij} = 0 & i \leq m \\ x_{ij} = x_{i-m,j-m} - u_{i-m,j} & i > m \text{ and } j > m \\ x_{ij} = 0 & i > m \text{ and } j \leq m \end{cases}$$

Observing how we defined the matrix X, one can check that $X \in \Omega_n$. Moreover, it is not hard to see that X is a solution of the above system of equations. Therefore, it must be $\Omega_{n+m} \subseteq [\Omega_n, \Omega_n]$. This completes proof. □

Remark 5.1.5. The Corollary 5.1.3 also provides us with important information about the structure of the sets Ω_n. Since the intersection of all members of the lower central series of $\mathcal{N}T_\infty(R)$ is trivial, $\mathcal{N}T_\infty(R)$ is residually nilpotent and has a trivial center $\mathcal{Z}(\mathcal{N}T_\infty(R))$.

In order to carry out our primary work in a more understandable way, it is better for us to recall the definitions and examples we mentioned earlier concentrating on the Lie algebra $\mathcal{N}T_\infty(R)$.

Definition 5.1.6. An R-linear map $\varphi : \mathcal{N}T_\infty(R) \to \mathcal{N}T_\infty(R)$ satisfying

$$\varphi([X, Y]) = [\varphi(X), Y] + [X, \varphi(Y)] \text{ for all } X, Y \in \mathcal{N}T_\infty(R)$$

is called a **derivation** of $\mathcal{N}T_\infty(R)$.

We denote the set of all derivations of $\mathcal{N}T_\infty(R)$ by $Der(\mathcal{N}T_\infty(R))$. It can be easily shown that it forms a Lie algebra under operations

$$(\varphi + \psi)(X) = \varphi(X) + \psi(X),$$

$$(\lambda.\varphi)(X) = \lambda.\varphi(X),$$

$$([\varphi, \psi])(X) = \varphi(\psi(X)) - \psi(\varphi(X)).$$

While we're on this subject, let's examine the behaviour of the sets $\Omega_n, n \geq 2$ under a derivation of $\mathcal{N}T_\infty(R)$.

Lemma 5.1.7. *For each* n, $\Omega_n = \left\{X = (x_{ij})_{i,j\in\mathbb{Z}^+} \in \mathcal{N}T_\infty(R) \mid x_{ij} = 0 \text{ for all } j < i+n\right\}$ *is invariant under any derivation* φ *of* $\mathcal{N}T_\infty(R)$.

Proof. We give a proof by induction on n. Clearly, $\varphi(\Omega_1) \subseteq \Omega_1$ by definition. Then

$$\begin{aligned}
\varphi(\Omega_{i+1}) &= \varphi([\Omega_i, \Omega_1]) = [\varphi(\Omega_i), \Omega_1] + [\Omega_i, \varphi(\Omega_1)] \\
&\subseteq \varphi(\Omega_i)\Omega_1 - \Omega_1\varphi(\Omega_i) + \Omega_i\varphi(\Omega_1) - \varphi(\Omega_1)\Omega_i. \\
&\subseteq \Omega_i\Omega_1 - \Omega_1\Omega_i + \Omega_i\Omega_1 - \Omega_1\Omega_i \text{ (by the induction hypothesis)} \\
&\subseteq \Omega_{i+1} - \Omega_{i+1} + \Omega_{i+1} - \Omega_{i+1} \subseteq \Omega_{i+1}.
\end{aligned}$$

□

Inner derivations will be an essential part here as it has been up to now. Now, we take a look at them.

Definition 5.1.8. Let $X \in \mathcal{N}T_\infty(R)$. Then the mapping

$$\begin{array}{rccc}
\mathrm{ad}X: & \mathcal{N}T_\infty(R) & \to & \mathcal{N}T_\infty(R) \\
& Y & \to & \mathrm{ad}X(Y) = [X, Y]
\end{array}$$

turns out to be a derivation of our Lie algebra $\mathcal{N}T_\infty(R)$, called **inner derivation** induced by X.

The set of all inner derivations of $\mathcal{N}T_\infty(R)$ is a Lie subalgebra of $Der(\mathcal{N}T_\infty(R))$. We will denote it by $IDer(\mathcal{N}T_\infty(R))$. Now, consider the following map

$$\begin{array}{rccc}
\phi: & \mathcal{N}T_\infty(R) & \to & IDer(\mathcal{N}T_\infty(R)) \\
& X & \to & \mathrm{ad}X
\end{array}$$

Clearly, ϕ preserves addition and scalar multiplication. Let $X \in \mathcal{N}T_\infty(R)$, then we have

$$\begin{aligned}
\phi([X, Y])(Z) &= \phi(XY - YX)(Z) \\
&= (XYZ - YXZ) - (ZXY - ZYX) \\
&= \mathrm{ad}X(\mathrm{ad}Y(Z)) - \mathrm{ad}Y(\mathrm{ad}X(Z)) \\
&= ([\phi(X), \phi(Y)])(Z).
\end{aligned}$$

As a result of that, ϕ defines a Lie algebra homomorphism. Also, it is easy to see that ϕ is a surjective map and

$$\begin{aligned} Ker\phi &= \left\{ X \in \mathcal{N}T_\infty(R) \mid \phi(X) = \mathrm{ad}X = 0 \quad \right\} \\ &= \left\{ X \in \mathcal{N}T_\infty(R) \mid XY - YX = 0 \text{ for all } Y \in \mathcal{N}T_\infty(R) \right\} \\ &= \left\{ X \in \mathcal{N}T_\infty(R) \mid X \in \mathcal{Z}(\mathcal{N}T_\infty(R)) \right\} \\ &= \mathcal{Z}(\mathcal{N}T_\infty(R)). \end{aligned}$$

By using First Isomorphism Theorem, we get

$$IDer(\mathcal{N}T_\infty(R)) \cong \mathcal{N}T_\infty(R) \Big/ \mathcal{Z}(\mathcal{N}T_\infty(R)) \cong \mathcal{N}T_\infty(R).$$

Definition 5.1.9. Let $D \in \mathcal{D}_\infty(R)$, where $\mathcal{D}_\infty(R)$ is the Lie subalgebra of diagonal matrices. Then the mapping defined as

$$\begin{array}{rccc} \mathrm{ad}D: & \mathcal{N}T_\infty(R) & \rightarrow & \mathcal{N}T_\infty(R) \\ & X & \rightarrow & [D, X] \end{array}$$

constitute another classical derivation of $\mathcal{N}T_\infty(R)$. It is known as **diagonal derivation**.

In the next section, we propose some technical lemmas that will make our job easier.

5.2 Auxiliary Lemmas

Firstly, we consider the following subsets of $\mathcal{N}T_\infty(R)$:

$$\mathfrak{h}_k = \left\{ \sum_{k<j} x_{kj}\xi_{kj} \mid x_{kj} \in R \right\}, \quad k = 1, 2, 3, \dots$$

Let's see what we have when $k = 1$:

$$\mathfrak{h}_1 = \left\{ \sum_{1<j} x_{1j}\xi_{1j} \mid x_{1j} \in R \right\} = \left\{ X \in \mathcal{N}T_\infty(R) \mid X = \begin{pmatrix} 0 & x_{12} & x_{13} & x_{14} & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right\}.$$

Lemma 5.2.1. *The subset $\mathfrak{h}_1$ is an ideal of $\mathcal{N}T_\infty(R)$ and it is invariant under the action of any derivation of $\mathcal{N}T_\infty(R)$.*

Proof. Let $X = (x_{ij})_{i,j\in\mathbb{Z}^+} \in \mathfrak{h}_1$ and $Z = (z_{ij})_{i,j\in\mathbb{Z}^+} \in \mathcal{N}T_\infty(R)$. Then, we can easily compute that $ZX = O$ and

$$XZ = \begin{pmatrix} 0 & 0 & (x_{12}z_{23}) & (x_{12}z_{24} + x_{13}z_{34}) & \sum_{j=2}^{4} x_{1j}z_{j5} & \sum_{j=2}^{5} x_{1j}z_{j6} & \ldots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

It follows from the above observations that $[Z, X] = ZX - XZ = -XZ \in \mathfrak{h}_1$, and so $\mathfrak{h}_1$ is an ideal of $\mathcal{N}T_\infty(R)$. Now, we shall try to get second statement. Let φ be a derivation of $\mathcal{N}T_\infty(R)$. The first thing we will do is show that $\varphi(\xi_{12}) \in \mathfrak{h}_1$. Since $\varphi(\xi_{12}) \in \mathcal{N}T_\infty(R)$, we can write $\varphi(\xi_{12})$ as below

$$\varphi(\xi_{12}) = \sum_{1\le i<j} x_{ij}\xi_{ij}.$$

Obviously, $\varphi(O) = O$ by the additivity of φ and the Lie product $[\xi_{12}, \xi_{k,k+1}] = 0$ for any $k > 2$. Combining these investigations,

$$\begin{aligned} \varphi(O) = \varphi([\xi_{12}, \xi_{k,k+1}]) &= [\varphi(\xi_{12}), \xi_{k,k+1}] + [\xi_{12}, \varphi(\xi_{k,k+1})] \\ O &= \varphi(\xi_{12})\xi_{k,k+1} - \xi_{k,k+1}\varphi(\xi_{12}) + \xi_{12}\varphi(\xi_{k,k+1}) - \varphi(\xi_{k,k+1})\xi_{12} \\ O &= \varphi(\xi_{12})\xi_{k,k+1} - \xi_{k,k+1}\varphi(\xi_{12}) + \xi_{12}\varphi(\xi_{k,k+1}) - 0 \quad (\text{since } \xi_{12} \in \mathfrak{h}_1) \qquad (*) \end{aligned}$$

Let's find out what the matrix $\varphi(\xi_{12})\xi_{k,k+1}$ looks like. Easy calculations give us that

$$\varphi(\xi_{12})\xi_{k,k+1} = \begin{array}{cc} \begin{array}{c} (k{+}1)\text{-th column} \end{array} & \\ \begin{pmatrix} 0 & \ldots & 0 & \varphi(\xi_{12})_{1k} & 0 & \ldots \\ 0 & \ldots & 0 & \varphi(\xi_{12})_{2k} & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & \varphi(\xi_{12})_{kk} = 0 & 0 & \ldots \\ 0 & \ldots & 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} & \begin{array}{c} \\ \\ \\ k\text{-th row} \\ \\ \\ \end{array} \end{array}$$

We focus now on the matrix

$$
\xi_{k,k+1}\varphi(\xi_{12}) = \begin{array}{cc}
\qquad\qquad\qquad (k{+}2)\text{-th column} & \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \dots \\
0 & \dots & 0 & \varphi(\xi_{12})_{k+1,k+2} & \varphi(\xi_{12})_{k+1,k+3} & \dots \\
0 & \dots & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix} & \begin{array}{l} \\ \\ \\ k\text{-th row} \\ \\ \\ \end{array}
\end{array}
$$

In view of the above observations, the equation $(*)$ can be written as follows

$$
\begin{aligned}
O &= \varphi(\xi_{12})\xi_{k,k+1} - \xi_{k,k+1}\varphi(\xi_{12}) + \xi_{12}\varphi(\xi_{k,k+1}) \\
O &= \sum_{i=1}^{k-1} x_{ik}\xi_{i,k+1} - \sum_{j=k+2}^{\infty} x_{k+1,j}\xi_{kj} + \xi_{12}\varphi(\xi_{k,k+1})
\end{aligned}
$$

From the fact that $\xi_{12}\varphi(\xi_{k,k+1}) \in \mathfrak{h}_1$ and $\varphi(\xi_{12})\xi_{k,k+1}$ has no nonzero entry in its k-th row, it follows that $\sum_{j=k+2}^{\infty} a_{k+1,j}\xi_{kj} = O$. Accordingly, it must be $\varphi(\xi_{12})\xi_{k,k+1} = -\xi_{12}\,\varphi(\xi_{k,k+1})$, and so $\varphi(\xi_{12})\xi_{k,k+1} \in \mathfrak{h}_1$. This means that $x_{ik} = 0$ for any $k \geq 3, 1 < i < k$. Hence, we get $\varphi(\xi_{12}) \in \mathfrak{h}_1$.

Now consider the set $\mathfrak{h}_1^{'}$ which is defined as below

$$
\mathfrak{h}_1^{'} = [\mathfrak{h}_1, \mathfrak{h}_2] = \left\{ X \in \mathcal{N}T_\infty(R) \mid X = \begin{pmatrix}
0 & 0 & x_{13} & x_{14} & \dots \\
0 & 0 & 0 & 0 & \dots \\
0 & 0 & 0 & 0 & \dots \\
0 & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix} \right\} \subset \Omega_2.
$$

Since $\mathfrak{h}_1$ is an ideal of $\mathcal{N}T_\infty(R)$, we can immediately say that $R.\xi_{12} + \mathfrak{h}_1^{'} \subseteq \mathfrak{h}_1$. Actually, what is more crucial is that the reverse inclusion also holds. Let's take an arbitrary matrix $Z = (z_{ij})_{i,j\in\mathbb{Z}^+} \in \mathfrak{h}_1$

$$
Z = \begin{pmatrix}
0 & z_{12} & z_{13} & z_{14} & z_{15} & \dots \\
0 & 0 & 0 & 0 & 0 & \dots \\
0 & 0 & 0 & 0 & 0 & \dots \\
0 & 0 & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}
$$

and put

$$X = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \in \mathfrak{h}_1 \text{ and } Y = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & z_{13} & z_{14} & z_{15} & \dots \\ 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \in \mathfrak{h}_2.$$

Then $Z = z_{12}\xi_{12} + XY$, and so $\mathfrak{h}_1 \subseteq R\xi_{12} + \mathfrak{h}_1^{'}$. In addition to this, from the fact that a matrix $T \neq O \in R\xi_{12}$ has nonzero entry (only) in the (1, 2) position, we have $R\xi_{12} \cap \mathfrak{h}_1^{'} = \{O\}$. Consequently, $\mathfrak{h}_1$ is a direct sum of these two subset, $\mathfrak{h}_1 = R\xi_{12} + \mathfrak{h}_1^{'}$. Therefore, we get $\varphi(\mathfrak{h}_1) = R\varphi(\xi_{12}) + \varphi(\mathfrak{h}_1^{'})$. Indeed, the idea of writing $\mathfrak{h}_1$ as a direct sum of $\mathfrak{h}_1^{'}$ and $R\xi_{12}$ will ease things along. We have just seen that $\varphi(\xi_{12}) \in \mathfrak{h}_1$, and so $R\varphi(\xi_{12}) \subseteq \mathfrak{h}_1$. Hence, from here on all we only need to do is obtain that $\varphi(\mathfrak{h}_1^{'}) \subseteq \mathfrak{h}_1$. To do this, we first show that $\varphi(\mathfrak{h}_1^{'}) \subseteq \mathfrak{h}_1 + \mathfrak{h}_2$. By the construction of $\mathfrak{h}_1'$, we have

$$\varphi(\mathfrak{h}_1^{'}) = \varphi([\mathfrak{h}_1, \mathfrak{h}_2]) \subseteq \varphi(\mathfrak{h}_1)\mathfrak{h}_2 - \mathfrak{h}_2\varphi(\mathfrak{h}_1) + \mathfrak{h}_1\varphi(\mathfrak{h}_2) - \varphi(\mathfrak{h}_2)\mathfrak{h}_1.$$

Since $\mathfrak{h}_1^{'} = [\mathfrak{h}_1, \mathfrak{h}_2] \subseteq \Omega_2$ and Ω_2 is invariant under φ, we deduce that

$$\begin{aligned} \varphi(\mathfrak{h}_1)\mathfrak{h}_2 &= (R\varphi(\xi_{12}) + \varphi(\mathfrak{h}_1^{'}))\mathfrak{h}_2 = R\varphi(\xi_{12})\mathfrak{h}_2 + \varphi(\mathfrak{h}_1^{'})\mathfrak{h}_2 \\ &\subseteq \mathfrak{h}_1\mathfrak{h}_2 + \Omega_2\mathfrak{h}_2 \subseteq \mathfrak{h}_1 + \{O\} = \mathfrak{h}_1. \end{aligned}$$

Thus,

$$\varphi(\mathfrak{h}_1^{'}) \subseteq \varphi(\mathfrak{h}_1)\mathfrak{h}_2 - \mathfrak{h}_2\varphi(\mathfrak{h}_1) + \mathfrak{h}_1\varphi(\mathfrak{h}_2) - \varphi(\mathfrak{h}_2)\mathfrak{h}_1 \subseteq \mathfrak{h}_1 - \mathfrak{h}_2 + \mathfrak{h}_1 - \{O\} \subseteq \mathfrak{h}_1 + \mathfrak{h}_2.$$

We are now ready to conclude that $\varphi(\mathfrak{h}_1^{'}) \subseteq \mathfrak{h}_1$, which will complete our proof as we said before. For any $X \in \mathfrak{h}_1^{'}$, we have

$$\begin{aligned} O &= \varphi(O) = \varphi([\xi_{12}, X]) \\ O &= \varphi(\xi_{12})X - X\varphi(\xi_{12}) + \xi_{12}\varphi(X) - \varphi(X)\xi_{12} \\ O &= O - O + \xi_{12}\varphi(X) - O = \xi_{12}\varphi(X). \end{aligned}$$

Since X is an arbitrary element of $\mathfrak{h}_1^{'}$, we obtain that $\xi_{12}\varphi(\mathfrak{h}_1^{'}) = \{O\}$. To put it another way, the equality tell us that $\varphi(X)_{2j} = 0$ for all j. Hence, it must be $\varphi(\mathfrak{h}_1^{'}) \subseteq \mathfrak{h}_1$ from the fact that $\varphi(\mathfrak{h}_1^{'}) \subseteq \mathfrak{h}_1 + \mathfrak{h}_2$. Consequently, the desired result follows. $\square$

Remark 5.2.2. In the previous lemma, we observe that $\mathfrak{h}_1$ is an ideal of $\mathcal{N}T_\infty(R)$. However, $\mathfrak{h}_k$ is not an ideal of $\mathcal{N}T_\infty(R)$ for $k \geq 2$ since $\mathfrak{h}_1 \cap \mathfrak{h}_k$ is trivial.

The following lemma is key here. It will form the basis of the proof of the main result.

Lemma 5.2.3. *Any derivation φ of $\mathcal{N}T_\infty(R)$ can be decomposed into the sum of a diagonal derivation, an inner derivation and a derivation ψ such that $\psi(\mathfrak{h}_1) = \{O\}$.*

Proof. From Lemma 5.1.7 and Lemma 5.2.1, it follows that $\varphi(\xi_{1k}) \in \mathfrak{h}_1 \cap \Omega_{k-1}$ for any derivation φ and $k \geq 2$. Thus, one may write $\varphi(\xi_{1k})$ as follows

$$\varphi(\xi_{1k}) = \sum_{j\geq k} x^k_{1j}\xi_{1j} = \begin{pmatrix} 0 & 0 & \dots & 0 & x^k_{1,k} & x^k_{1,k+1} & x^k_{1,k+2} & x^k_{1,k+3} & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \quad k \geq 2.$$

Now, let's take an arbitrary $Y = (y_{ij})_{i,j\in\mathbb{Z}^+} = \sum_{j=2}^{\infty} y_{1j}\xi_{1j} \in \mathfrak{h}_1$. Then,

$$\varphi(Y)_{1j} = \sum_{i=2}^{j} y_{1i}x^i_{1j} \text{ for } j = 2, 3, 4...$$

If we are going to represent $\varphi(Y)$ in array format:

$$\varphi(Y) = \begin{pmatrix} 0 & (y_{12}x^2_{12}) & (y_{12}x^2_{13} + y_{13}x^3_{13}) & (y_{12}x^2_{14} + y_{13}x^3_{14} + y_{14}x^4_{14}) & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \qquad (1)$$

Now, consider the matrices

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 & ... \\ 0 & x^2_{12} & 0 & 0 & ... \\ 0 & 0 & x^3_{13} & 0 & ... \\ 0 & 0 & 0 & x^4_{14} & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } X = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & x^2_{13} & x^2_{14} & x^2_{15} & \dots \\ 0 & 0 & 0 & x^3_{14} & x^3_{15} & ... \\ 0 & 0 & 0 & 0 & x^4_{15} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Clearly, $D \in \mathcal{D}_\infty(R)$ and $X \in \mathcal{N}T_\infty(R)$. Since $Y \in \mathfrak{h}_1$, it must be

$$\text{ad}X(Y) = XY - YX = -YX \quad \text{and} \quad \text{ad}D(Y) = DY - YD = -YD.$$

Easy calculations give us that

$$\text{ad}X(Y) = -YX = -\begin{pmatrix} 0 & 0 & (y_{12}x_{13}^2) & (y_{12}x_{14}^2 + y_{13}x_{14}^3) & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \tag{2}$$

$$\text{ad}D(Y) = -YD = -\begin{pmatrix} 0 & y_{12}x_{12}^2 & y_{13}x_{13}^3 & y_{14}x_{14}^4 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{3}$$

From the equations $(1), (2)$ and (3), we obtain that

$$(\varphi - \text{ad}X - \text{ad}D)(Y) = 0 \text{ for any } Y \in \mathfrak{h}_1.$$

In order to complete proof, say $\psi = \varphi - \text{ad}X - \text{ad}D$. Accordingly,

$$\varphi = \psi + \text{ad}X + \text{ad}D \quad \text{with} \quad \psi(\mathfrak{h}_1) = \{O\}.$$

$\square$

Lemma 5.2.4. *Let ψ be a derivation of $\mathcal{N}T_\infty(R)$ with $\psi(\mathfrak{h}_1) = \{O\}$. Then $\psi(X) \in \mathfrak{h}_1$ for any $X \in \mathcal{N}T_\infty(R)$.*

Proof. Let $X \in \mathfrak{h}_1, Y \in \Omega_1 = \mathcal{N}T_\infty(R)$. Of course, $YX = 0$ since $X \in \mathfrak{h}_1$. Thus, it must be $[X, Y] = XY - YX = XY$. Also, $[X, Y] = XY \in \mathfrak{h}_1$ because $\mathfrak{h}_1$ is an ideal of $\mathcal{N}T_\infty(R)$. Thereby, it follows from hypothesis that $\psi([X, Y]) = O$. Hence:

$$\begin{aligned} \psi([X, Y]) &= [\psi(X), Y] + [X, \psi(Y)] = [O, Y] + [X, \psi(Y)] \\ O &= O + [X, \psi(Y)] \\ O &= X\psi(Y) - \psi(Y)X. \end{aligned}$$

As can be seen above, we get the following equality

$$X\psi(Y) = \psi(Y)X.$$

It also follows from $X \in \mathfrak{h}_1$ that $\varphi(Y)X = 0$, and so $X\psi(Y) = 0$. As X and Y are chosen randomly, we deduce that $\mathfrak{h}_1\psi(\Omega_1) = \{O\}$. We are now ready to prove our lemma thanks to these critical examinations.

Let $X \in \mathcal{N}T_\infty(R)$. We just observed that $\mathfrak{h}_1\psi(\Omega_1) = \{O\}$. Thus, $Y\psi(X) = 0$ for any $Y \in \mathfrak{h}_1$. Of course, $\psi(X) = (u)_{i,j\in\mathbb{Z}^+} \in \mathcal{N}T_\infty(R)$ can be written as

$$\psi(X) = \begin{pmatrix} 0 & u_{12} & u_{13} & u_{14} & ... \\ 0 & 0 & u_{23} & u_{24} & ... \\ 0 & 0 & 0 & u_{34} & ... \\ 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Now, as a motivating example, observe the matrices $Y_2 \in \mathfrak{h}_1$ and $Y_3 \in \mathfrak{h}_1$, which are defined as

$$Y_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad Y_3 = \begin{pmatrix} 0 & 0 & 1 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

It is easy to compute that

$$Y_2\psi(X) = \begin{pmatrix} 0 & 0 & u_{23} & u_{24} & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \text{ and } Y_3\psi(X) = \begin{pmatrix} 0 & 0 & 0 & u_{34} & u_{35} & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

As you see, $u_{2k} = 0$ for any $k \geq 3$ because $Y_2\psi(X) = O$. Similarly, $Y_3\psi(X) = O$ implies that $u_{3k} = 0$ for any $k \geq 4$. In fact, through similar arguments one can generalize these results. Let's consider the matrices $Y_k \in \mathfrak{h}_1$ defined as below for $k \geq 2$:

$$Y_k = \begin{cases} y_{1j} = 1 & \text{if } j = k \\ y_{1j} = 0 & \text{if } j \neq k \end{cases}$$

Then, we have

$$Y_k\psi(X) = O = \begin{pmatrix} 0 & 0 & ... & 0 & u_{k,k+1} & u_{k,k+2} & u_{k,k+3} & ... \\ 0 & 0 & ... & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

As a result, we find that all coefficients of $\psi(X)$ in the k-th row are zero for $k \geq 2$, which means $\psi(X) \in \mathfrak{h}_1$. $\square$

Lemma 5.2.5. *If ψ is a derivation of our Lie algebra $\mathcal{N}T_\infty(R)$ such that $\psi(\mathfrak{h}_1) = \{O\}$, then there must exist $Y \in \mathfrak{h}_1$ such that $\psi = adY$.*

Proof. Since $\psi(\mathfrak{h}_1) = \{O\}$, by the above lemma it must be $\psi(\xi_{k,m}) \in \mathfrak{h}_1$ for any $k \geq 2$ with $m > k$. Thus, one can write $\psi(\xi_{k,m})$ as

$$\psi(\xi_{k,m}) = \sum_{j\geq 2} y_{km}^j \xi_{1j} = \begin{pmatrix} 0 & y_{km}^2 & y_{km}^3 & y_{km}^4 & y_{km}^5 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

For any $l \neq k-1, m$ note that $[\xi_{k,m}, \xi_{l,l+1}] = O$. So, $\psi([\xi_{k,m}, \xi_{l,l+1}]) = O$ by the linearity of ψ. Also, $\xi_{l,l+1}\psi(\xi_{k,m}) = O = \xi_{k,m}\psi(\xi_{l,l+1})$ as $\psi(\xi_{k,m})$ and $\psi(\xi_{l,l+1})$ are both in $\mathfrak{h}_1$. Thus,

$$\begin{aligned} \psi([\xi_{k,m}, \xi_{l,l+1}]) &= [\psi([\xi_{k,m}), \xi_{l,l+1}] + [\xi_{k,m}, \psi(\xi_{l,l+1})] \\ O &= \psi(\xi_{k,m})\xi_{l,l+1} - \psi(\xi_{l,l+1})\xi_{k,m} \\ O &= \left(\sum_{j\geq 2} y_{km}^j \xi_{1j}\right)\xi_{l,l+1} - \psi(\xi_{l,l+1})\xi_{k,m} \end{aligned}$$

Now, let's take a look at the matrices in the last equation above.

$$X = (x_{ij})_{i,j\in\mathbb{Z}^+} = \psi(\xi_{k,m})\xi_{l,l+1} = \begin{pmatrix} 0 & 0 & ... & 0 & x_{1,l+1} & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where $x_{1,l+1} = \psi(\xi_{km})_{1l} = y^l_{km}$ and

$$Z = (z_{ij})_{i,j\in\mathbb{Z}^+} = \psi(\xi_{l,l+1})\xi_{k,m} = \begin{pmatrix} 0 & 0 & ... & 0 & z_{1,m} & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where $z_{1,m} = \psi(\xi_{l,l+1})_{1k}$. Since $X - Z = 0$, we find that $y^l_{km} = 0$ if $l + 1 \neq m$. Of course, we must examine what happens if $m = l + 1$. In such a case, it must be $l \neq k - 2$ by our assumption. (At the beginning of this investigation, we assumed that $l \neq k - 1, m$.) Here, the thing making us happy is that $y^l_{km} = 0$ even though $l = m + 1$. Don't worry about how we can prove it. That can also be shown easily using the fact that $[\xi_{k,m}, \xi_{l,l+2}] = 0$. In short, what we obtain is

$$y^l_{km} = 0 \text{ for any } l \neq k - 1, m.$$

Now, let us assume that $l \neq k - 2, m$. Then, $[\xi_{k,m}, \xi_{l,l+2}] = 0$ for any $l \neq k - 2$, m. With an approach similar to the above, one can conclude that

$$y^l_{km} = 0 \text{ for any } l \neq k - 2, m.$$

Accordingly, $y^l_{km} = 0$ if $l \neq m$. (That is, we don't care what k is.) Therefore,

$$\psi(\xi_{k,m}) = y^m_{k,m}\xi_{1m} = y_{k,m}\xi_{1,m}. \qquad (*)$$

Moreover, for every $m \neq k + 1$

$$\psi(\xi_{km}) = \psi([\xi_{k,k+1}, \xi_{k+1,m}]) = [\psi(\xi_{k,k+1}), \xi_{k+1,m}] + [\xi_{k,k+1}, \psi(\xi_{k+1,m})]$$
$$\psi(\xi_{km}) = \psi(\xi_{k,k+1})\xi_{k+1,m} - \psi(\xi_{k+1,m})\xi_{k,k+1}$$
$$\psi(\xi_{km}) = \psi(\xi_{k,k+1})\xi_{k+1,m} = y_{k,k+1}\xi_{1m}. \qquad (**)$$

Combining the equations $(*)$ and $(**)$, we get an extremely important relationship between y_{km}'s

$$y_{i,i+1} = y_{i,i+2} = y_{i,i+3} = y_{i,i+4} = y_{i,i+5}... \text{ for any } i \geq 2.$$

This result will make things easier as you will now see. Now, put

$$Y = \sum_{k=2}^{\infty} y_{k,k+1}\xi_{1k} = \begin{pmatrix} 0 & y_{23} & y_{34} & y_{45} & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \in \mathfrak{h}_1.$$

As you guess, from now on, our goal is to show that ψ in an inner derivation induced by Y. Let $Z = (z_{ij})_{i,j\in\mathbb{Z}^+} = \sum_{1\leq i<j} z_{ij}\xi_{ij} \in \mathcal{N}T_\infty(R)$, then

$$\psi(Z)_{1j} = \sum_{i=2}^{j-1} z_{ij}\psi(\xi_{ij})_{1j} \text{ for } j = 3, 4, 5...$$

To be more precise,

$$\psi(Z) = \begin{pmatrix} 0 & 0 & z_{23}y_{23} & z_{24}y_{24} + z_{34}y_{34} & z_{25}y_{25} + z_{35}y_{35} + z_{45}y_{45} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Now, let's find out what ad$Y(Z)$ equals. Clearly,

$$\text{ad}Y(Z) = [Y, Z] = YZ - ZY = YZ \quad (\text{since } Y \in \mathfrak{h}_1.)$$

The matrix YZ also can be computed easily

$$YZ = \begin{pmatrix} 0 & 0 & y_{23}z_{23} & (y_{23}z_{24} + y_{34}z_{34}) & (y_{23}z_{25} + y_{34}z_{35} + y_{45}z_{45}) & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$YZ = \sum_{j=3}^{\infty} \left(\sum_{i=2}^{j-1} z_{ij}y_{i,i+1} \right) \xi_{1j}.$$

Now, it is time to make the final push. Using the fact that "$y_{i,i+1} = y_{i,i+2} = y_{i,i+3} = \dots$" for any $i \geq 2$, one can easily see that

$$\psi(Z) = \begin{pmatrix} 0 & 0 & z_{23}y_{23} & (z_{24}y_{23} + z_{34}y_{34}) & (z_{25}y_{23} + z_{35}y_{34} + z_{45}b_{45}) & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$
$$= \mathrm{ad}Y(Z).$$

Because Z was taken arbitrarily, we are done. $\square$

With these lemmas in hand we are now ready to prove our main result.

5.3 Proof of the Main Result

Theorem 5.3.1. *Let φ be a derivation of $\mathcal{N}T_\infty(R)$. Then it can be written as*

$$\varphi = adZ + adD,$$

where $Z \in \mathcal{N}T_\infty(R)$ and $D \in \mathcal{D}_\infty(R)$. The derivation adZ is determined uniquely and adD is determined uniquely up to scalar matrix.

Proof. From Lemma 5.2.3, we can express φ as

$$\varphi = \mathrm{ad}D + \mathrm{ad}X + \psi,$$

where $\psi(\mathfrak{h}_1) = \{O\}$. Then, Lemma 5.2.5 implies that $\psi = \mathrm{ad}Y$ for some $Y \in \mathfrak{h}_1$. Thus, we obtain

$$\varphi = \mathrm{ad}D + \mathrm{ad}Z,$$

where $Z = X + Y$. Of course, in order for such an expression to make any sense, it must be unique. Now, let us check "uniqueness." Suppose that there are matrices $D' \in \mathcal{D}_\infty(R)$ and $Z' \in \mathcal{N}T_\infty(R)$ satisfying

$$\varphi = \mathrm{ad}D + \mathrm{ad}Z = \mathrm{ad}D' + \mathrm{ad}Z'.$$

For any $k > 1$,

$$D\xi_{1k} = \begin{array}{c} \qquad\qquad\qquad \text{k-th column} \\ \begin{pmatrix} 0 & ... & 0 & d_{11} & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \end{array} \text{ and } \xi_{1k}D = \begin{array}{c} \qquad\qquad\qquad \text{k-th column} \\ \begin{pmatrix} 0 & ... & 0 & d_{kk} & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \end{array}.$$

Thus,

$$\text{ad}D(\xi_{1k}) = \begin{array}{c} \qquad\qquad\qquad \text{k-th column} \\ \begin{pmatrix} 0 & ... & 0 & d_{11} - d_{kk} & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \end{array}.$$

Also, easy calculations show that

$$\xi_{1k}Z = (v_{ij})_{i,j\in\mathbb{Z}^+} = \begin{pmatrix} 0 & 0 & ... & 0 & v_{1,k+1} & v_{1,k+2} & \cdots \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \text{ with } v_{1,k+i} = z_{k,k+i}, i \geq 1.$$

Clearly, ad$Z(\xi_{1k}) = -\xi_{1k}Z$ since $Z\xi_{1k} = 0$. So, we get $\varphi(\xi_{1k}) = X = (x_{ij})_{i,j\in\mathbb{Z}^+}$, where $x_{1k} = d_{11} - d_{kk}$, and $x_{1j} = -z_{kj}$ for $j > k$. Moreover, from our assumption it must be $x_{1k} = d'_{11}$ - d'_{kk}, and $x_{1j} = -z'_{kj}$ for $j > k$. Accordingly, we find that $z_{kj} = z'_{kj}$ for $k > 1$ and $j > k$. For $j \leq k$, it is clear that we have $z_{kj} = 0 = z_{kj}{}'$. Thus, $z_{kj} = z'_{kj}$ for all j and $k > 1$. There is only one thing left for us to see the equality of Z and Z'. Let's compare the first rows of these two matrices: For $k > 1$, one can easily check that $\varphi(\xi_{k,k+1}) = Y$, where $y_{1,k+1} = z_{1k} = z'_{1k}$. By the way, of course, $z_{11} = 0 = z'_{11}$. Hence, we get $z_{1k} = z'_{1k}$ for all k. Consequently, $Z = Z'$.

The other statement comes from the fact that $d'_{11} - d'_{kk} = d_{11}$-$d_{kk}$. Using this, we can directly write d'_{11} - d_{11} as $d'_{kk} - d_{kk}$, which implies that $D - D' = \lambda I_\infty$, where $\lambda \in R$. Hence, we are done. $\square$

Chapter 6

CONCLUSION

In this thesis, some infinite matrix rings and algebras have been examined, and description of derivations of these matrix rings and algebras have been presented by compiling papers [20] and [11]. We're not limited to just that. It has also been demonstrated with striking examples that the idea of extending many basic theorems in the theory of finite matrices to infinite matrices is exceptionally wrong. In doing so, we benefited from the articles [4] and [12].

As previously mentioned, studies on the context of infinite matrices are ongoing, and there are still too many unanswered questions in the literature. The exciting thing here is that we had contact with one of them. The ring $\mathcal{N}T_\infty(R)$ of all infinite (upper) niltriangular matrices, whose rows and colums are indexed by $\mathbb{Z} \times \mathbb{Z}$, over a commutative ring R with identity provides an excellent example of reasonable open problems in the sense that its Lie automorphisms and Jordan automorphisms are not yet known.

You appreciate knowing all derivations and automorphisms of a mathematical object helps us understand its structure, and that way we can reach our goal more easily. In this sense, we think that the topics we have compiled in this thesis will be a guiding resource for those studying infinite matrix rings and algebras. This is exactly why we dedicate our thesis to them.

Bibliography

[1] Benkart, G.M., Osborne, J.M., Derivations and automorphisms of non-associative matrix algebras, Trans. Am. Math. Soc., 263, 411-430, **1981**.

[2] Benkovic, D., Lie derivations of triangular matrices, Linear Multilinear Algebra, 55, 619–626, **2007**.

[3] Benkovic, D., Lie triple derivations of unital algebras with idempotents, Linear Multilinear Algebra, 63, 141-165, **2015**.

[4] Bossaller, D.P., López-Permouth, S.R., On the Associativity of Infinite Matrix Multiplication, The American Mathematical Monthly, 126:1, 41-52, **2019**.

[5] Cao, Y., Wang, J., A note on algebra automorphisms of strictly upper triangular matrices over commutative rings, Linear Algebra and its Applications, 311, 187–193, **2000.**

[6] Chun, J.H, Park, J.W., Derivations on subrings of matrix rings, Bulletin-Korean Mathematical Society, 43, 635-644, **2006**.

[7] Cooke, R.G., Infinite Matrices and Sequence Spaces, New York: Dover Publications, **1966**.

[8] Dubisch, R., Perlis, S., On total nilpotent algebras, American Journal of Mathematics, 73, 439-452, **1951**.

[9] Erdmann, K., Wildon, M.J, Introduction to Lie Algebras, Springer-Verlag London, **2006**.

[10] Herstein, I.N., Noncommutative Rings, The Mathematical Association of America, **1968**.

[11] Hołubowski, W., Kashuba, I., Żurek, S., Derivations of the Lie algebra of infinite strictly upper triangular matrices over a commutative ring, Communications in Algebra, 45(11), 4679-4685, **2017.**

[12] Hołubowski, W., An inverse matrix of an upper triangular matrix can be lower triangular , General Algebra and Applications, 22 ,161-166, **2002.**

[13] Jondrup, S., Automorphisms of upper triangular matrix rings, Archiv der Mathematik, 49, 497-502, **1987**.

[14] Kolesnikov, S.G., Mal'tsev, N.V., Derivations of a matrix ring containing a subring of triangular matrices, (Russian) Izvestiya Vysshikh Uchebnykh Zavedenii Matematika, 11, 23-33, **2011**; translation in Russian Mathematics (Iz. VUZ) 55, 18-26, **2011.**

[15] Kuzucuoğlu, F., Jordan derivations on strictly triangular matrix rings, Algebra Colloquium, 18, 519-522, **2011**.

[16] Levchuk, V.M., Connections between the unitriangular group and certain rings,II. Groups of automorphisms, Siberian Mathematical Journal, 24, 543- 557, **1983.**

[17] Ou, S., Wang, D., Yao, R., Derivations of Lie algebra of strictly upper triangular matrices over a commutative ring, Linear Algebra and its Applications, 424, 378-383, **2007.**

[18] Ou, S., Wang, D., Yu, Q., Derivations of the intermediate Lie algebras between the Lie algebra of diagonal matrices and that of upper triangular matrices over a commutative ring , Linear Multilinear Algebra, 54, 369-377, **2006.**

[19] Semrl, P., Maps on matrix spaces, Linear Algebra and its Applications, 413, 364-393, **2006.**

[20] Slowik, R., Derivations of Rings of Infinite Matrices, Communications in Algebra, 43, 3433-3441, **2015.**

[21] Sushkevich, A. K., On an infinite algebra of triangular matrices (Russian), Har'kov. Gos. Univ. Uc. Zap., 34, 77-93, **1950.**